QUANTIFIERS, QUESTIONS AND QUANTUM PHYSICS

Quantifiers, Questions and Quantum Physics

Essays on the Philosophy of Jaakko Hintikka

Edited by

DANIEL KOLAK
William Paterson University, Wayne, NJ, U.S.A.

and

JOHN SYMONS
University of Texas at El Paso, TX, U.S.A.

Springer

A C.I.P. Catalogue record for this book is available from the Library of Congress.

ISBN 1-4020-3210-2 (HB)
ISBN 1-4020-3211-0 (e-book)

Published by Springer,
P.O. Box 17, 3300 AA Dordrecht, The Netherlands.

Sold and distributed in North, Central and South America
by Springer,
101 Philip Drive, Norwell, MA 02061, U.S.A.

In all other countries, sold and distributed
by Springer,
P.O. Box 322, 3300 AH Dordrecht, The Netherlands.

Printed on acid-free paper

Printed in the Netherlands.

Contents

Foreword and Acknowledgements

Jaakko Hintikka is one of the most creative figures in contemporary philosophy. He has made significant contributions to virtually all areas of the discipline (with the exception of moral philosophy) from epistemology and the philosophy of logic to the history of philosophy, aesthetics and the philosophy of science. In our view, part of the fruitfulness of Hintikka's work is due to its opening important new lines of investigation and new approaches to traditional philosophical problems.

In this volume we have gathered together essays from some of Hintikka's colleagues and former students exploring his influence on their work and pursuing some of the insights that we have found in his work. While the book does contain some criticism of Hintikka's views, this certainly does not purport to be a fair and balanced look at his work. We are unabashedly partisan in our admiration for the man and his work and have put this volume together in a collaborative spirit as a celebration of Hintikka's many contributions to philosophy.

In this volume we have included an annotated bibliography of Hintikka's work. We gratefully acknowledge the *Philosopher's Information Center, The Philosopher's Index* and Dick Lineback in particular for permission to reprint some of the abstracts included in the bibliography. By itself, this would serve as an important resource for philosophers and scholars. 'Prolific' is too modest an adjective for Hintikka, as readers can see for themselves from the size of this annotated bibliography. His massive and diverse body of work poses a real challenge for scholars who hope to find a single philosophical agenda or view that we can associate with Hintikka.

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 1-2.

300+ articles, many of them groundbreaking, overwhelm and in a certain sense eclipse his 35+ books. There are a number of ways that one can approach the scale and variety of this work. Our purpose in including the bibliography is to permit others to glean what they will from Hintikka's prodigious philosophical output. We eagerly anticipate the publication of a current bibliography of Hintikka's work, including all reprint and translation details in the *Library of Living Philosophers* volume dedicated to Hintikka. That task, unfortunately, was beyond us. Heartfelt thanks also to Anthony E. Nelson for expert assistance with the grueling task of typesetting.

When we considered the importance and impact of Hintikka's work, it occurred to us that its philosophical consequence is not the additive property of the sum of its parts. We struggled for a way to think about the proliferation of research programs, counterarguments and Ph.D. dissertations that Hintikka's work inspires and settled in the end on the awkward analogy of the powerset. Hintikka's philosophical legacy will be something like the philosophical powerset of his publications and lines of research. The powerset of a set *S*, is the set of possible subsets of *S,* and by analogy, rather than attempting to synthesize Hintikka's work into well-defined themes or bumper-stickers, our goal here is to represent the proliferation of different ways one can construe his work and the variety of lines of inquiry that it suggests.

We are very grateful to the distinguished group of colleagues who have contributed to this volume. We are a diverse group, from recent students of Hintikka to some of his most distinguished peers. While we are far from agreement on all the issues discussed in this volume, we are all united by a great fondness for this remarkable man. We see him as a central and pivotal figure in our individual and collective pursuits of wisdom.

Anyone who is even remotely aware of what Hintikka may be working on at the moment will have the impression that his next greatest achievement, his next greatest result, is just down the road ahead of us, just around the next bend. Those of us who have the privilege of knowing Hintikka cannot help feeling the intensity and excitement of philosophical discovery. Unlike so many of the cynical, world-weary philosophers who figured so prominently in recent decades, Hintikka's energy, optimism and mental agility are unparalleled. In that respect, he is the most refreshingly immature mature philosopher in our midst. To put it simply, among philosophers Hintikka is youngest at heart, and boldest of mind.

Daniel Kolak and John Symons

HINTIKKA ON EPISTEMOLOGICAL AXIOMATIZATIONS

Vincent F. Hendricks
Department of Philosophy and Science Studies
Roskilde University, Denmark

1. INTRODUCTION

Among the many intellectual accomplishments for which Jaakko Hintikka is recognized is his pioneering work in epistemic logic. Although epistemic logic was studied somewhat in the Middle Ages the real break-throughs are to be found in the work of von Wright [59] and most notably Hintikka's seminal book *Knowledge and Belief: An Introduction to the Logic of the Two Notions* from 1962 [24]. There has hardly been an article or book published on the logic of knowledge and belief since that has not made reference to this exquisite treatise.

For the past 40 years epistemic and doxastic logics have developed into fields of research with wide ranges of application. They are of immanent importance to theoretical computer science, artificial intelligence, linguistics, game theory, economics and social software. Be that as it may, epistemic and doxastic logics are still in an awkward philosophical position today. Computer scientists, linguistics and other formally minded researchers utilizing the means and methods do not necessarily have an epistemological ambition with their use of epistemic logic. At the same time it is a discipline devoted to the logic of knowledge and belief but alien to epistemologists and philosophers interested in the theory of knowledge.

Hintikka from the very beginning had a strong epistemological ambition with his development of epistemic logic however. It was not to be another technical spin-off of advances in modal and other intensional logics. Its purpose was, and still remains, to elucidate various epistemic notions and reason about knowledge and belief. Epistemic logic is to serve as a *logical epistemology* for mainstream and formal epistemological approaches alike.

Despite Hintikka's original intentions, ambitions and own work the epistemological significance of epistemic logic has in general been neglected and perhaps even sometimes intentionally ignored by both formal and

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 3-32.

mainstream epistemologists. Epistemology is in the business of dealing with skepticsm and the possibility of error—logical epistemology may actually be viewed as being much in the same business. Modal concepts of knowledge quantify over other possible worlds to secure the robustness and streadfastness of knowledge. But the classical conception of infallibilism is taken to require, that for an agent to have knowledge of some hypothesis or proposition,[1] he must be able to eliminate *all* the possibilities of error associated with the hypothesis in question. The set of *all* worlds is considered. This set of possible worlds is too big for knowledge to have scope over. The set includes some rather bizarre worlds inhabited by odd beasts from demons to mad and malicious scientists who have decided to stick your brain in a tank of nutritious fluids to systematically fool you. Or worlds in which contradictions are true. If these worlds were to be considered relevant all the time skepticism would have the upper hand all the time. There may not be a way for an agent to determine that he is not in the world of the beast or the brain. If infallibilism is to be a viable reply to the skeptic, then infallibilism cannot be defined with respect to all possible worlds. Hintikka may be read as saying something similar when it comes to epistemic logic:

> What the concept of knowledge involves in a purely logical perspective is thus a dichotomy of the space of all possible scenarios into those that are compatible with what I know and those that are incompatible with my knowledge. This observation is all we need for most of epistemic logic. [31], p. 2.

This way of battling the skeptic by limiting the set of citable possible worlds carrying potential error has been referred to as ‘forcing’ in Hendricks [17], [18] and in particular [19]:

Whenever knowledge claims are challenged by alleged possibilities of error, the strategy is to show that the possibilities of error fail to be genuine in the relevant sense

[1]‘Hypothesis’ and ‘proposition’ will be used interchangably.

Logical epistemology or epistemic logic pays homage to the forcing strategy as the partitioning of the space of possible worlds compatible with knowledge attitude determines a certain set over which the epistemic operator is to have scope. Contemporary mainstram epistemologists choose to speak of the *relevant* possible worlds as a subset of the set of all possible worlds.[2] The epistemic logician considers an *accessibility* relation between worlds in a designated class out of the entire universe of possible worlds. There is no principled difference between relevance and accessibility. Informal epistemologies differ by the way in which relevance is forced given, say, perceptual equivalence conditions, counterfactual proximities or conversational contexts circumscribing the possible worlds. Formal epistemologies differ by the way in which the accessibility relation is defined over possible worlds.

Epistemic logicians obtain different epistemic modal systems valid for a knowledge operator by varying (adding, dropping or relativizing) the properties of the accessibility relation from, say, reflexive and transitive to a reflexive, symmetric and transitive relation. Algebraic constraints on the accessibility relation are the forcing foundation for a formal approach to the theory of knowledge like logical epistemology. Constraints on accessibility relations between possible worlds is a way of demonstrating some of the epistemological significance of Hintikka's philosophical program in epistemic logic already present in *Knowledge and Belief* and of course beyond.

2. EPISTEMIC LOGIC AND SKEPTICISM

For a proper syntactic augmentation of the language of the propositational logic with two unary operators $K_\Xi A$ and $B_\Xi A$ such that

$K_\Xi A$ reads 'Agent Ξ knows *A*' and $B_\Xi A$ reads 'Agent Ξ knows *A*'

for some arbitrary proposition *A*, Hintikka came up with the following semantic interpretations of the epistemic and doxastic operators [24], [25]:

$K_\Xi A$ ≈ *in all possible worlds compatible with what* Ξ *knows, it is the case that A*

[2]Explicit forcing proposals in the epistemological literature are sometimes referred to as '*relevant alternatives proposals*'. Cf. Bernecker and Dretske [1].

$B_\Xi A$ *in all possible worlds compatible with what* Ξ *knows, it is the case that* A

The basic assumption is that any ascription of propositional attitudes like knowledge and belief, requires *partitioning* of the set of possible worlds into two compartments: The compartment consisting of possible worlds compatible with the attitude in question and the compartment of worlds incompatible with it. Based on the partition the agent is capable of constructing different 'world-models' using the epistemic modal language. He is not necessarily required to know which one of the world-models constructed is the real world-model. All the same, the agent does not consider all these world-models equally possible or accessible from his current point of view. Some world-models may be incommensurable with his current information state or other background assumptions. These incompatible world-models are excluded from the compatibility partition. This is a variation of the forcing strategy. In logical epistemology, as in many mainstream epistemologies, it is typically stipulated that the smaller the set of worlds an agent considers possible, the smaller his uncertainty, at the cost of stronger forcing assumptions.

The set of worlds considered accessible by an agent depends on the actual world, or the agent's actual state of information. It is possible to capture the forcing dependency by introducing a relation of accessibility, R, on the set of compatible possible worlds. To express the idea that for agent Ξ, the world w' is compatible with his information state, or accessible from the possible world w which Ξ is currently in, it is required that R holds between w and w'. This relation is written Rww' and read 'world w' is accessible from w'. The world w' is said to be an *epistemic alternative* to world w for agent Ξ. Given the above semantical interpretation, if a proposition A is true in all worlds which agent Ξ considers possible then Ξ knows A.

Formally, a *frame* F for an epistemic system is a pair (W, R) for which W is a non-empty set of possible worlds and R is a binary accessibility relation over W. A *model* **M** for an epistemic system consists of a frame and a denotation function φ assigning sets of worlds to atomic propositional formulae. Propositions are taken to be sets of possible worlds; namely the set of possible worlds in which they are true. Let *atom* be the set of atomic propositional formulae, then φ: *atom* $\rightarrow P(W)$ where P denotes the powerset operation. The model $\mathbf{M} = \langle W, R, \varphi\rangle$is called a Kripke-model and the resulting semantics Kripke-semantics [34]: An atomic propositional formulae, **a**, is said to be true in a world w (in **M**), written $\mathbf{M}, w \models \mathbf{a}$, iff w is in the set of possible worlds assigned to **a**, i. e. $\mathbf{M}, w \models \mathbf{a}$ iff $w \in \varphi(\mathbf{a})$ for all **a** $\in$ *atom*. The formula $K_\Xi A$ is true in a world w, i.e. **M**, $w \models K_\Xi A$, iff $\forall w' \in W$: if Rww', then **M**, $w \models A$. The semantics for the Boolean connectives are given in the usual recursive way. A modal formula is said to be *valid* in a frame iff the formula is true for all possible assignments in all worlds admitted by the frame.

A nice feature of possible world semantics is that many common epistemic axioms correspond to certain algebraic properties of the frame in the following sense: A modal axiom is valid in a frame if and only if the accessibility relation satisfies some algebraic condition. For an example, the axiom

$K_\Xi A \rightarrow A$ (1)

is valid in all frames in which the accessibility relation is *reflexive* in the sense that every possible world is accessible from itself. (1) is called axiom T,[3] or the *axiom of truth* or *axiom of veridicality*, and says that if A is known by Ξ, then A is true in accordance with the standard tripartite definition of knowledge as true justified belief.

Similarly if the accessibility relation satisfies the condition that

$\forall w, w', w'' \in W: Rww' \wedge Rw'w'' \rightarrow Rww''$

then the axiom

$K_\Xi A \rightarrow K_\Xi K_\Xi A$ (2)

is valid in all *transitive* frames. (2) is called axiom 4 and is also known as the *axiom of self-awareness*, *positive introspection* or *KK-thesis*. The labels all refer to the idea that an agent has knowledge of his knowledge of A if he has knowledge of A. Other axioms require yet other relational properties to be met in order to be valid in all frames: If the accessibility relation is reflexive, symmetric and transitive, then

$\neg K_\Xi A \rightarrow K_\Xi \neg K_\Xi A$ (3)

is valid. (3) is called axiom 5 also better known as the *axiom of wisdom*. It is the much stronger thesis that an agent has knowledge of his own ignorance: If Ξ does not know A, he knows that he doesn't know A. The axiom is sometimes referred to as the *axiom of negative introspection*.

As opposed to (1)–(3) there is a formula or axiom which is valid in all possible frames

$K_\Xi(A \rightarrow A') \rightarrow (K_\Xi A \rightarrow K_\Xi A')$ (4)

The axiom amounts to the contentious closure condition for knowledge and is also referred to as axiom K, or the *axiom of deductive cogency*: If the agent Ξ knows $A \rightarrow A'$, then if Ξ knows A, Ξ also knows A'. One rule of

[3]This nomenclature due to Lemmon [36] and later refined by Bull and Segerberg [4] is helpful while cataloguing the axioms typically considered interesting for epistemic logic.

inference which is valid in all possible frames is the rule of *necessitation* or *epistemization* (N)

$A \,/\, K_{\Xi}A$ (5)

which says that if A is true in all worlds of the frame, then so is $K_{\Xi}A$.

Logical epistemology unproblematically accepts (4)–(5) but for formal reasons. Neither (4) nor (5) require any assumptions to be made pertaining to the accessibility relation between the possible worlds considered compatible with the knowledge attitude. It actually turns out that (4) together with (5) comprise the characterizing axiom and rule for possible world semantics with binary accessibility relations. All modal logics in which (4) and (5) are valid are called *normal* modal logics.

These axioms in proper combinations make up epistemic modal systems of varying strength depending on the modal formulae valid in the respective systems given the algebraic properties assumed for the accessibility relation. The weakest system of epistemic interest is usually considered to being system **T**. The system includes T and K as valid axioms. Additional modal strength may be obtained by extending **T** with other axioms drawn from the above pool altering the frame semantics to validate the additional axioms. Reflexivity is the characteristic frame property of system **T**, transitivity is the characteristic frame property of system **S4**, equivalence the characteristic frame property of **S5**, etc. From an epistemological point of view, the algebraic properties of the accessibility relation are really forcing conditions.

The cognitive rationale of logical epistemology must be something like this: The more properties the accessibility relation is endowed with, the more access the agent has to his epistemic universe, and in consequence the more epistemic strength he will obtain. The stronger knowledge, the stronger forcing clauses.[4]

Modal epistemic axioms and systems may be viewed as measures of infallibility and replies to skepticism. For instance, knowing your own knowledge is a way of blocking the skeptic, but knowledge of your own ignorance in terms of axiom 5 is better still. One motivation for the plausibility of axiom 5 is in data-base applications: An agent examining his own knowledge base will be let to conclude that whatever is not in the knowledge base he does not know and hence he will know that he does not.

The axiom of wisdom or negative introspection is a sort of closed world assumption. A closed world assumption is a forcing assumption if anything is, ‘shutting the world down’ with the agent, leaving the skeptic nowhere to go. To know the truth, to know of your knowledge, and to know of your own

[4] Attention is currently restricted to Kripke-semantics and the forcing clauses restricted accordingly.

ignorance as in **S5** requires 'full' epistemic access which is exactly why the accessibility relation must be an equivalence relation. A theorem of **S5** is the following

$\neg A \to K_\Xi \neg K_\Xi A$ (6)

which states that if A is not the case, then Ξ knows that he does not know A—the 'truly Socratic person' as Girle explains ([13], p. 157) knowing exactly how ignorant he is.

A bit more ignorance, a bit more skepticism and accordingly a bit more fallibility is allowed in **S4**. Since axiom **5** is dropped and (6) is no longer a theorem, $\{\neg A, \neg K_\Xi \neg K_\Xi A\}$ and $\{\neg K_\Xi \neg A, \neg K_\Xi \neg K_\Xi A\}$ are not inconsistent in **S4**. It is possible for an agent to be ignorant of the fact that he does not know when actually he does know. Put differently, the agent is allowed false beliefs about what is known. Yet more ignorance and skepticism are allowed in system **T** because while $\{K_\Xi \neg A, \neg K_\Xi \neg K_\Xi A\}$ is inconsistent in **S4**, this set of epistemic statements is not inconsistent in **T**. The agent may thus know something without knowing that he does.[5]

What Hintikka recently dubbed 'first generation epistemic logic' in [30] is characterized by the ambition that cataloguing the possible complete systems of such logics would allow for choosing the most 'appropriate' or 'intuitive' ones(s).[6] Hintikka himself settled for **S4** in *Knowledge and Belief*, but he had very strong epistemological arguments for doing so.

3. THE LOGIC OF AUTOEPISTEMOLOGY

Hintikka stipulated that the axioms or principles of epistemic logic are conditions descriptive of a special kind of general (strong) *rationality* from a first person perspective.[7] The statements which may be proved false by application of the epistemic axioms are not inconsistent meaning that their truth is logically impossible. They are rather rationally 'indefensible'. Indefensibility is fleshed out as the agent's epistemic laziness, sloppiness or

[5] All the same, a restricted kind of positive introspection is still prevalent in system **T**. Given the rule of necessitation (5), Ξ knows all the theorems of the epistemic logic. By iteration, $K_\Xi K_\Xi A$ is also known. Thus if A is a theorem, Ξ knows that he knows A.

[6] Hintikka's 'second generation epistemic logic' is discussed under the rubric 'active agenthood' in Hendricks [18], [19], and [23]. For excellent surveys of epistemic logic and its contemporary themes see also van Benthem [2] and Gochet and Gribomont [14].

[7] For a systematic discussion of logical epistemology from first and third person perspectives refer to Hendricks [19].

perhaps cognitive incapacity whenever to realize the implications of what he in fact knows. Defensibility then means not falling victim of 'epistemic neglience' as Chisholm calls it [5], [6]. The notion of indefensibility gives away the status of the epistemic axioms and logics. Some epistemic statement for which its negation is indefensible is called 'self-sustaining'. The notion of self-sustenance actually corresponds to the concept of validity. Corresponding to a self-sustaining statement is a logically valid statement. But this will again be a statement which is rationally indefensible to deny. So in conclusion, epistemic axioms are descriptions of rationality.

There is an argument to the effect that Hintikka early on was influenced by the autoepistemology of G.E. Moore [47] and especially Malcolm [46] and took, at least in part, their autoepistemology to provide a philosophical motivation for epistemic logic. Moore's common-sense considerations on which autoepistemology is founded deflates the skeptical possibilities of error from various dialectic angles of which one is particularly pertinent to the current discussion. It is called the argument from *incoherence*. The idea is to demonstrate that skepticism has severe difficulties in formulating its own position coherently. As with any argument, a skeptical conclusion presupposes knowledge of a set of premisses. Moore then points to the fact that merely *asserting* these premisses imply at least a doxastic commitment, but most likely an epistemic commitment. The skeptics cannot be retreating to a statement like

> '*There are 9 planets in our solar system but it is not the case that I believe it.*' (7)

The statement in (7) is an instance of what later has become known as the *Moore-paradox*. Let it be granted that (7) only involves an error of omission. All the same it still sounds self-contradictory simply given mere assertion. No formulation of skepticism without incoherence, or in Hintikkian terms, skepticism is an irrational or indefensible epistemological position.

The argument from incoherence is a first person point argument. Skepticism is thus rejected along these lines. A first person perspective is one of the very characteristics of autoepistemology. This is also suggested in the label 'autoepistemology' attaching the Moore-paradox to it: Whatever an agent may know or believe is partly fixed by the concern whether the epistemic or doxastic claim advocated by the inquiring agent fall victim of a Moore-paradox or not. As long as a thesis concerning epistemic commitments does not pan out in a Moore-paradox the inquiring agent is free to adopt it. As an autoepistemologist one may, by way of example, say

> '*If I believe that A, then I believe that I know that A.*' (8)

which has later been called the *Moore-principle* and sometimes the *principle of positive certainty*.[8] Formalized (8) amounts to:

$B_{\Xi}A \rightarrow B_{\Xi}K_{\Xi}A$ (9)

According to Moore's theory, there is nothing self-contradictory or incoherent about asserting the principle. No more Moore paradox to the Moore principle than to the widely adopted principle that one knows that one knows if one does the plausibility of which Malcolm argues for below and elsewhere [46].

From Moore's first person autoepistemological perspective a statement like

'*A is the case, but I don't believe whether A.*' (10)

is a paradoxical Moorean statement. There is however nothing paradoxical about

'*A is the case, but Ξ doesn't believe whether A.*' (11)

from a third person perspective. In consequence, what for sure may sound quite implausible from the first person perspective, may sound very plausible from the third person perspective on inquiry and vice versa.

The epistemic and doxastic commitments that an agent may hold in the course of inquiry are sensitive the epistemic environment and what the agent in these local circumstances is both willing to and capable of defending or maximizing. He does not necessarily have an over-all skepticism defeating method at his disposal: You are doing the best you can, so is the skeptic, but he is probably not doing as well as you are due to incoherence. Forcing in autoepistemology then means:

Whenever knowledge claims are challenged by alleged possibilities of error, the strategy is to show that on an individual basis one can do no better than what is being done in the current epistemic environment and attempt to show that the skeptic is doing at least as bad as you are but probably even worse

Epistemic axioms may be interpreted as principles describing a certain strong rationality congruent with autoepistemology. First of all, neither Malcolm nor Moore would object to the idea that knowledge validates axiom T (1). Secondly, in Hintikka's logical system knowledge is closed in the sense of (4), and the argument cited by Hintikka in favor of closure has the flavor of autoepistemology:

[8] Lamarre and Shoham explain: 'To the agent, the facts of which he is certain appear to be knowledge', [35].

> In order to see this, suppose that a man says to you, ‘I know that *p* but I don’t know whether *q*’ and suppose that *p* can be shown to entail logically *q* by means of some argument which he would be willing to accept. Then you can point out to him that what he says he does not know is already implicit in what he claims he knows. If your argument is valid, it is irrational for our man to persist in saying that he does not know whether *q* is the case. [24], p. 31.

Not accepting (4) is irrational, but the acceptance of (4) does not entail that the agent in question has to be immediately aware of his own rationality, let alone able to immediately compute it from Hintikka’s first person perspective on inquiry.

The autoepistemological inspiration is vindicated while Hintikka argues for the plausibility of the *KK*-thesis as a governing axiom of his logic of knowledge. Approximately a decade after the publication of *Knowledge and Belief*, the *KK*-thesis came under heavy attack. *Synthese* dedicated an issue to the matter where especially Ginet and Castenada were on the offensive, while Hintikka and Hilpinen defended.[9] And while defending, Hintikka refers to Malcolm:[10]

> Many of the things Malcolm says fall flat if it is not the case that I in fact know what I claim to know. For instance, if I am the victim of a clever optical trick when I believe that there is an ink-bottle in front of me—and even believe that I know it in the strong sense—then exposing the trick will provide conclusive evidence against claiming that the ink-bottle is there ... More generally, we might perhaps say that if one knows in the strong sense that *p*, then it is the case that one will refuse (if acting rationally) to consider any experience compatible with what he in fact knows as evidence against one’s knowing that *p*. ([26]), p. 153.

From this Hintikka concludes that Malcolm’s position is sufficiently close to Hintikka’s own for a behavioral identity between the strong knowledge á la Malcolm á la Hintikka:

[9] *Synthese* **21**, 1970.

[10] For a thourough discussion of Hintikka’s conception of the *KK*-thesis, refer to Hendricks [17], pp. 253.

> This is especially interesting in view of the fact that Malcolm himself uses his strong sense of knowing to explain in what sense it might be true that whenever one knows, one knows that one knows. In this respect, too, Malcolm's strong sense behaves like mine. [26], p. 154.

Besides the requirement of closure and the validity of the *KK*-thesis, axiom T is also valid so the suggestion is that a logic of autoepistemology is philosophically congruent with Hintikka's suggestion for an **S4** axiomatization describing strong rationality.

Although the epistemic logic of autoepistemology may be **S4**, the doxastic logic is another matter, and the affinities with autoepistemology end. Moore's principle above (8) is a kind of introspection axiom for rational belief or *subjective certainty*. In a combined epistemic and doxastic logical system in which knowledge and belief are approximately equally strong (save for a truth-condition) the agent will (while subjectively reflecting upon his own state of mind with respect to what he believes) be led to believe that he knows the proposition in question if he certainly believes it. Some contemporary logical epistemologists embrace Moore's principle (e.g. Halpern [15]). Hintikka denies Moore's principle in *Knowledge and Belief*:

> Hence ... and (C.BK) [Moore's principle] are acceptable only when an unrealistically high standard of defensibility is imposed on one's beliefs. The conditions would make it (logically) indefensible to suppose that anyone would have given up any of his present beliefs if he had more information than he now has. And this is clearly too stringent a requirement. [24], p. 52.

To Hintikka belief is a significantly weaker commitment than knowledge. For good reason too it turns out: Consider a combined epistemic and doxastic logic in which belief is understood as subjective certainty such that (9) holds. Assume also that positive doxastic introspection

$B_{\Xi}A \rightarrow K_{\Xi}B_{\Xi}A \quad (12)$

holds for belief together with negative doxastic introspection

$\neg B_{\Xi}A \rightarrow K_{\Xi}\neg B_{\Xi}A. \quad (13)$

Even subjective certainty, as strong as it may seem in this system, implies a margin of error: The fact that Ξ is subjectively certain of A does not necessarily imply that A is true. Accordingly axiom T will be dropped for subjective certainty and replaced by the consistency axiom D

$B_{\Xi}A \rightarrow \neg B_{\Xi}\neg A. \quad (14)$

On the standard definition of knowledge, knowledge implies belief

$K_{\Xi}A \rightarrow B_{\Xi}A \quad (15)$

which is also an uncontroversially accepted assumption for knowledge and subjective certainty. The logic of subjective certainty is **KD45**. Knowledge will obviously have to be stronger than subjective certainty, so it must validate **S5**. On assumptions (9), (12)–(15) Lenzen was able to show that $B_\Xi A$ in the end is equivalent to $K_\Xi A$ [37]. So knowledge and belief collapse into each other![11]

Many contemporary epistemic logics do nevertheless consider strong belief, rational belief or subjective certainty to be approximately as strong as knowledge. Assuming belief is taken to be approximately as strong as **S5** knowledge with the equivalence relation over worlds implies some attractive formal features like readily epistemic and doxastic partitions. This does not by itself make up for the result that the logic of knowledge and belief coincide.

Hintikka denies the axiom of wisdom because introspection alone should not license agents to ascertain whether some proposition in question is known. Other objections to (3) include the following: Under special circumstances axiom 5 suggests that agents can even decide intractable problems as Binmore reveals in [3], and Shin in [53]. Williamson has launched two objections to models of knowledge and belief validating axiom 5. For **S5** knowledge Williamson disagrees with the ones interpreting knowledge in a data-base like fashion to justify the closed world assumption of axiom 5. Even under the closed world assumption it does not follow in general that an agent can 'survey the totality of its knowledge'.[12] Secondly, Williamson recently noted that the result to the effect that knowledge and belief collapse under the strong understanding of belief in a combined system points to the untenability of axiom 5, not to the unacceptable nature of subjective certainty *per se*. Moore's principle is not too extravagant an assumption for rational belief, neither are axioms (12), D, (15) nor axioms T, 4 for knowledge. That leaves axiom 5 as the culprit responsible for collapsing the two notions and besides entails the infallibility of the agent's beliefs: Whatever Ξ believes is true. On these grounds, Williamson abandons axiom 5 rather than any of the other principles used in the derivation [61]. Voorbraak makes the unusual move of sacrificing (15) accordingly challenging the intuitions of philosophers since antiquity [58]. In Hendricks [17] it is shown how limiting convergent knowledge and (3) conflict, and in Hendricks [19] it is demonstrated how the axiom of wisdom gives rise to both conceptual and technical problems in multi-agent systems.

[11] Stalnaker also discusses this issue in [56].

[12] See [60], p. 317.

4. 'EPISTEMOLOGICS'

If **S5** assumptions about knowledge and belief are dropped ideal rationality descriptions and autoepistemological considerations may supply a philosophical foundation and motivation for logical epistemology.[13] The treatment of logical epistemology as a branch of modal logic is still quite costly also for much less ambitious logics than **S5**. The principle of closure (4) is enough to generate problems, and worse, skeptical problems. Nozick for instance emphatically denies closure for epistemic operators given his subjunctive definition of knowledge, and a whole range of other epistemic axioms likewise have to go [48].

4.1 Counterfactuality

According to Nozick, epistemology is not going to get off the ground before the skeptical challenge is met. It must be demonstrated that knowledge is at least possible. The often cited premiss in favor of the skeptical conclusion that agents do not know much of anything is this: If the agent cannot be guaranteed to be able to know the *denials* of skeptical hypotheses, then the agent cannot be ascribed knowledge on any other issues. The traditional understanding of infallibilism counting every possible world as relevant supports the pessimistic premiss presented. Some arbitrary skeptical hypothesis is a possibility of error the falsity of which must be known to the agent for him to acquire knowledge of some other common hypothesis in question. The inability to know the denials of skeptical hypotheses suffice for lacking knowledge of the ordinary hypotheses.

The classical thesis of infallibilism supports the skeptical premiss by the demand that Ξ should be capable of knowing the denials of all the possibilities of error. The closure condition (4) demands that Ξ only is knowledgable of the denials of those possibilities of error which in effect are known logical consequences of Ξ's knowledge.[14] Suppose Ξ knows the hypothesis that he is currently sitting reading this article on forcing

[13]From the point of view of autoepistemology, one also suspects that Moore himself would be disinclined to advocate the axiom of negative introspection (axiom 5). Either because it could amount to a Moorean sentence or because it imposes too much rationality on the part of the singular agent—there is a difference between doing the best you can, and then outdoing yourself.

[14]... or perhaps rather known logical consequences of Ξ's knowledge – including denials of all possibilities of error (the so-called contrast consequences, Dretske [9]).

epistemology. Let it also be the case that Ξ knows that if he is sitting reading this paper, then he is not being fooled by the Cartesian demon. Then Ξ must also know that he is not being fooled by the demon. If Ξ does not know that he is not being deceived by the demon then, given Ξ knows the implication, Ξ in turn lacks knowledge of the hypothesis that he is sitting reading forcing epistemology. Now this is exactly what the pessimistic premiss pushes for. But Ξ can know that he is sitting reading this article without knowing that there is no demon of deception seducing him into the false belief that he is sitting reading this paper. Being seated reading this paper implies that no Cartesian demon is leading Ξ to falsely believe that he is reading this very article.

Two things follow from this reasoning: (1) Everyday knowledge is secured, but (2) knowledge is not closed in the sense of (4) according to Nozick's counterfactual epistemology. If knowledge was to be closed it could fly far away into skepticism.

Having denied the condition of closure the epistemological mission is still not completed. An explanation must still be provided describing how knowledge of common hypotheses is possible joined with an explanation of the failure to know the denials of skeptical hypotheses. This also goes for the situations in which it is known that the common hypothesis at issue implies relevantly rejecting the skeptical hypothesis.

Dretske's solution is to install a modal condition for knowledge imposing truth-conduciveness by *sensitivity* [9]:

'*If A were not true,* Ξ *would not believe A.*' (16)

A belief qualifying as knowledge is a belief which is sensitive to the truth: The proposition A is true in accordance with the standard definition of knowledge. Had A which is believed been false, the agent would not be led to the belief that A.

Condition (16) readily explains why closure fails. Proximity relations between possible worlds are introduced due to the semantics for the inserted subjunctive conditional. One may know both antecedents A and $A \rightarrow A'$ relative to one set of relevant worlds accessible from the actual world, and yet fail to know the consequent A' relative to a different set of possible worlds. Now relative to a set of possible worlds with proximity 'close' to the actual world one knows A and simultaneously knows that A implies the denial of the skeptical hypothesis, say A. But one may all the same fail to know the consequential denial of the skeptical hypothesis itself for knowledge of the skeptical hypothesis is relative to possible worlds with a 'way-off' proximity to the actual world . These possible worlds are radically different from the actual world by all means. 'Way-off' worlds are accordingly forced out, skepticism far away because closure fails, but the possibility of knowledge prevails.

In the monumental monograph on knowledge, skepticism, free will and other pertinent philosophical issues [48], Nozick completes a definition of counterfactual knowledge along the Dretskian lines:[15]

Ξ *knows A iff*
A is true,
Ξ *believes that A*,
$\neg A \Rightarrow \neg(\Xi$ *believes that A*),
$A \Rightarrow (\Xi$ *believes that A*)

To see how the definition works, the possible world semantics provides the following account of the truth-conditions for the subjunctive conditional: A subjunctive $A \Rightarrow B$ for arbitrary statements A and B, is true, insofar, in all those worlds in which A is true that are in proximity 'closest' to the actual world, B is also true in these 'closest' worlds. More specifically of three worlds w, w', w'' if w' is closer to w than w'', then $A \Rightarrow B$ will be true in w iff A is not true in any world or there exist a world w' in which A and B are true which is closer to w than any world w'' in which A is true but B is false.[16]

For knowledge possession, one does not have to consult all possible worlds as the skeptic would insist: Given the standard semantical analysis of the subjunctives it is enough that the consequent B holds in those possible worlds which are closest to the actual world such that the antecedent A holds. Speaking in terms of forcing a subjunctive conditional is true just in case the consequent is forced among the closest worlds to the actual world in which the antecedent holds.

The third condition of the definition above is there to avoid error. The fourth is there to gain truth. The two conditions are collapsible into one condition: Ξ'*s belief tracks the truth of A*:

> To know is to have a belief that tracks the truth. Knowledge is a particular way of being connected to the world, having a specific real factual connection to the world: tracking it. [48], p. 178.

The idea of introducing the proximity relation is that the agent's local epistemic environment normally suffices for the truth witnessing Nozick's first person stance. Although everyday knowledge is possible in many

[15]'$\Rightarrow$' denotes the subjunctive conditional.

[16]This semantic account of the subjunctive follows rather closely Lewis in [42]. Nozick is however not committed to a particular understanding of the semantics and also discusses Stalnaker's subjunctive semantics from [54]. See furthermore [48], p. 680, footnote 8.

contexts, some contexts are just beyond reach: It is impossible for Ξ to know that he is not this brain in a vat. Assuming the brain receives the same sensory patterns as it would was it not dumped in the vat, there would not be anything in the input revealing to Ξ that he was not a brain in a vat. In this devious scenario Ξ is also barred from knowing that he is sitting reading this paper on forcing. If Ξ claims to know that he is sitting reading this article, it must follow that he as a prerequisite tacitly approves of the hypothesis that he is not a brain in a vat. Given this prerequisite and *modus tollens* as Ξ does not know that he is not sunk into the vat he does not know that he is sitting reading this paper either.

Now the possible world in which Ξ is a brain in a vat is *ceteris paribus very* distant from the actual world. Failure of knowledge in these cases is not devastating to counterfactual epistemology. It hinges on the relevant possibilities of error. True beliefs are only required in possibilities closer to actuality that any $\neg A$-possibilities: Picture a physicist measuring the voltage drop over some LRC-circuit. A student from epistemology class comes to him and asks whether a relevant possibility of error could be that the voltmeter is calibrated incorrectly. The physicist would probably answer 'yes' as calibration problems could lead to a measurement error. Then asking the scientist whether being a brain in a vat is a relevant possibility of error would likely result in the physicist asking the student to go back to his course and stop bothering him with silliness.

By his definition of counterfactual knowledge, Nozick accepts the axiom of veridicality (1), and the rule of necessitation (5) also seems to hold: A is true, Ξ believes A, and since A is true in all possible worlds, A is also true in close worlds so Ξ knows A.[17]. But he rejects both closure and the *KK*-thesis (2) for counterfactual knowledge:

> Some writers have put forth the view that whenever one knows, one knows that one knows. There is an immediate stumbling block to this, however. One may know yet not believe one knows; with no existing belief that one knows to do the tracking of the fact that one knows, one certainly does not know that one knows. [48], p. 246.

An agent may be tracking the truth of A without tracking the fact that he is tracking the truth of A. For much the same reason chances are also that Nozick would dismiss the axiom of wisdom (3) because if an agent is not tracking the truth of A it does not follow that he will be tracking the fact that he is not tracking A. The first person logic of counterfactual epistemology is

[17]I'm indebted to Robert Stalnaker for bringing this to my attention.

thus very weak and *not* normal in the technical sense in contrast to Hintikka's logical epistemolology.

The counterfactual epistemology in general accommodates elements of the contextualistic epistemology of the next section. Dretske's view of the closure lets knowledge transfer work across known implications insofar as the implications in question are close or relevant. Knowing that one is sitting down reading this article transfers immediately through the known implication to the 'close' hypothesis that one is not standing on a street corner doing the same. This knowledge will at the same time not run through the known implication to the 'way-off' hypothesis that one is not being fooled by a malicious demon. Dretske's point seems to be that knowledge acquisition of a hypothesis in some common *context* assumes by default the very falsity of particular 'way-off' and irrelevant possibilities of error [9]. These possibilities of error are skirted, or their falsity presupposed in many everyday knowledge acquisition contexts. Lewis strongly subscribes to this contextualistic forcing feature in his modal epistemology – so does Hintikka.

5. CONTEXTUALITY

Lewis' new 'modal epistemology' [45] is an elegant variation of contextualism which has many (forcing) features in common with Hintikka's formal theory of knowledge.

Contextualistic epistemology starts much closer to home than counterfactual epistemology. Agents in their local epistemic environments have knowledge—and plenty of it in a variety of (conversational) contexts. Knowledge is not only possible as counterfactual epistemology will have it, it is real human condition. The general contextualistic template for a theory of knowledge is crisply summarized in DeRose's description of the attribution of knowledge. The description also embodies many of the epistemological themes central to the contextualistic forcing strategy:

> Suppose a speaker *A* says, '*S* knows that *P*', of a subject *S*'s true belief that *P*. According to contextualist theories of knowledge attributions, how strong an epistemic position *S* must be in with respect to *P* for *A*'s assertion to be true can vary according to features of *A*'s conversational context. [7], p. 4.

The incentive to take skeptical arguments to knowledge claims seriously is based on an exploitation of the way in which otherwise operational epistemic concepts, notably knowledge, can be gravely disturbed by sudden changes of the linguistic context in which they figure.

The standards for the possession of knowledge vary from context to context depending on what is at stake. In a course on epistemology the standards for knowledge possession fixed by the interlocutors (teacher and students) are usually very high. The conclusions that we know very little, if anything at all, may by the end of class be true. In a discussion after class a fellow student says ‘I know that *Matrix Revolutions* plays in the Park & 86th Street Theater on 125 E. 86th St.’. The circumstances have now changed and the standards for knowledge possession in this new, presumably, non-skeptical conversational context are lower. The relatively lower standards put us in the comfortable position of maintaining that we know most of what we think we know. It is admittedly to this low epistemic standard but it surely suffices for going to the movies.

Not only may knowledge attributions fluctuate with contexts, they may also be sensitive to who ascribes knowledge to whom. As indicated by DeRose there is a delicate issue to be addressed pertaining to the strength of the position an agent has to be in order for the epistemic commitment to truthfully pan out. This position is context-sensitive, not only to the agent in the environment, but also to possible ascribers of knowledge to the very agent in question. The first-third person dichotomy is immanent in contextualistic epistemologies.

Finally, the strength of the epistemic position is responsible for turning the contextualistic theory of knowledge into a modal account according to DeRose. For every local environmental ‘time-slice’ the epistemic position of the agent remains constant. The epistemic position the agent however *were* to be in to warrant possession of knowledge is a subjunctively defined spatio-temporal function of the context. A strong epistemic position with respect to some hypothesis A is to have belief as to whether A is the case and tracking this fact not only through the actual world but through close worlds as well. Maintaining that one’s belief still tracks the truth at long distances increases the strength of the epistemic position with respect to the hypothesis in question. For belief to become knowledge it should be ‘non-accidentally’ true in the actual world and in close ones as well.[18] This way of realizing the forcing relation resembles the construction advanced by the counterfactual epistemology of the previous section using sensitivity or tracking.

Lewis’ modal epistemology as a contextualistic theory of knowledge is particularly engaging as it balances elegantly between mainstream and formal *modi operandi*. This is not too surprising since Lewis through his career was concerned with modal logic, in particular the logic of counterfactuals [42], modal ontology [44] and almost consequently modal

[18]See further [7], p. 34.

epistemology [45]. Modal logics, epistemic logics in particular, are much about partitioning the set of all possible worlds into classes that are in close proximity, similar, relevant or accessible from the actual world and into classes which are not.

As humans we force for knowledge on a daily basis and obtain it. This means partitioning the set of all possible worlds into relevant, irrelevant and extravagant possibilities of error determined by the current context. To obtain knowledge eliminate the relevant possibilities of error, ignore the extravagant ones, and succeed over the remaining possible worlds where the hypothesis in question is true. Everything dictated by the current context. There are rules for elimination, ignoring and success. On a new definition of knowledge yet to be formulated, these rules are what Lewis' modal epistemology is about. Only a selected few of them will be discussed here.[19]

Taking infallibility as a basic epistemological condition, for an agent to know a hypothesis, all possibilities of error must be eliminated given the agent's available information. That is, all the possible worlds in which the negation of the hypothesis is the case must be eliminated. This forcing relation is given by different measures. One measure is simply to ignore possibilities extravaganza, another is to use the available evidence to force such that the uneliminated possible worlds are determined by perceptual equivalences over these alternatives with the actual world as the fix-point. The perceptual experience (and memory) the agent has in the actual world fixes the set of uneliminated possible worlds insofar the agent's cognitive apparatus functions the same in these worlds. Suppose that a perceptual experience has the propositional content A. The perceptual experience with content A (memory included) eliminates a certain world w' if and only if the content of the experience the agent has in w' differs from A.

Quantifiers are usually restricted to run over some domain of interest. This also goes for the universal quantifier over possible worlds that would lead to error. Every uneliminated world in which the hypothesis holds is restricted to a sub-domain of properly all uneliminated worlds. Saying that the surface is 'clean' in a certain conversational context is to properly ignore the microscopic dust particles laying on the surface. If somebody was to disagree it would have to be because the new interlocutor in the conversational context means clean in a more restrictive sense. The microscopic dust balls in this case suffice for making the assertion about the

[19]See Hendricks [19] for a complete exposition of the rules from a forcing perspective.

clean surface false. Words like ‘flat’ or ‘round’ behave in the same way, as does the word ‘knowledge’. They are context-sensitive.[20]

Alterations of the conversational context occur when a new hypothesis is introduced which for its part is more demanding than any of the other hypotheses currently explicit in the particular context. Such a *non-uniform* introduction implies an increase in the range of possible worlds to be considered for attribution of knowledge. The strength of the required epistemic position mentioned above is increased accordingly. In a context where the usage of ‘knowledge’ remains *uniform* throughout the conversation, the range of possible worlds to be considered remains stable. Given the context-sensitive nature of knowledge, in every context where knowledge attribution is at stake some uneliminated possible worlds are not rendered relevant by the *current* context. The universal quantifier is restricted accordingly. This restriction is very similar to the quantifier restriction on knowledge in logical epistemology. In epistemic logic, knowledge claims are circumscribed by the compartment of possible worlds in accordance with the epistemic attitude, not the incompatible compartment and not the set of all possible worlds.

These considerations essentially pave the way for the colloquially stated but forceful knowledge definition of modal epistemology:

> *S* knows that *P* iff *S*’s evidence eliminates every possibility in which not-*P*—Psst!—except for those possibilities that *we* are properly ignoring. [45], p. 378.

During the individual knowledge attribution process, the possible world which the agent takes to be the actual state of affairs is never ignored. Actuality is by reflexivity always a relevant possible world alternative although indexical. It follows that falsity may not properly be supposed. If falsity is never to be presupposed whatever in the end will turn up knowledge must be true, so the classical condition of truth for knowledge is derived. Never ignoring the actual world is referred to as the *rule of actuality*.

Turn next to the ascription of knowledge to others. The way in which the modal knowledge definition is stated italicizes ‘*we*’. What *we* may properly ignore is going to be dependent on whose actuality is being referred to in the context in question. Assuming that there is only one actual world-index in play in non-modal contexts one should expect that the world considered actual by the agent coincides with the world indexed ‘actual’ by the ascribers.

[20]The context-sensitivity of various words including ‘knowledge’ was noted by Lewis much earlier in [43].

In counterfactual situations referring for instance to what an agent would have known today had he read the paper yesterday, or whether an agent knew yesterday who he was then, fixing the index of actuality is trickier. Had the agent read the paper yesterday he would presumably have known more than he in fact knows today. The agent is ascribing knowledge and ignorance to himself now as the one not having read the paper last night. The ascriber, say Ξ*'*, of knowledge to agent Ξ has an index of actuality demonstratively different from Ξ's index. The index on actuality for Ξ*'* is what Ξ*'* would have been like knowledge-wise had he read the paper yesterday. Actuality indices differ for Ξ and Ξ*'* in this situation. Similarly for the attribution of knowledge to Ξ knowing yesterday who he was. For Ξ's reality is defined for his spatio-temporal history up until yesterday; for Ξ*'* reality is defined for his spatio-temporal history up to today when the question is popped whether Ξ knew yesterday who he was. The two world stories are different. Ξ's actuality yesterday is different from Ξ*'*'s actuality today. Similarly for a host of other situations involving say iterated modal constructions like knowledge of knowledge etc.

The rule of actuality applies both to the ascriber and the ascribed. What may not be properly ignored is the local agent's actuality. Epistemologists considering what Ξ knows from a third person perspective will attend to whatever possible worlds that Ξ himself attends to as possible and then some. The set of possible worlds ignored by a third person knowledge attributor for Ξ will properly be a superset of the possible worlds Ξ ignores. An agent may know more than what may be ascribed to him because his actuality in some cases differs from the ascribers and his range of viable worlds does as well. Applying the principle of 'epistemic' charity means that while attributing knowledge to an agent in his local epistemic environment, the third person ascriber may ignore fewer possibilities than Ξ.

Next, a world *w'* which 'salient resembles' another world *w* enforces a kind of symmetry. If *w* may not be properly ignored in virtue of the other rules neither may *w'* and vice versa. This accessibility clause is referred to as the *rule of resemblance*. The rule is dangerous and powerful at the same time.

The rule of resemblance is dangerous because not applied carefully invites skepticism and global underdetermination back in. The actual world is left uneliminated by the agent's available evidence. It follows that any other uneliminated world resembles the agent's actual world in one important respect, namely, with respect to the agent's evidence. This will continue to hold even in worlds which otherwise are radically different from the agent's actual world including the demon world. By application of the rule of actuality together with the rule of resemblance leads to the conclusion that these worlds are relevant alternative worlds as well!

There is no obvious remedy to this problem and it reappears with respect to knowledge closure. Agreeing with counterfactual epistemology that closure over arbitrary contexts amounts to a fallacy driving skeptical arguments, modal epistemology holds that closure is possible locally without skepticism. Knowledge is closed for a fixed context. Knowing that you are reading this paper implies that you are reading this paper and not being deceived (by a demon or a mad scientist) in this particular uniform context c_1. If the context is non-uniformly changed right after the antecedent conditions obtain to a new context c_2, 'all bets are off' [45], p. 382. Closure fails because the strength of the epistemic position now required in c_2 to attribute knowledge has been increased way beyond c_1 by the increase in possible worlds at issue dictated by c_2. The range of possible worlds may now include the demon world which is a whole different context. Knowledge is closed under implication because implication preserves truth in a fixed context not over arbitrary contexts.

If knowledge is closed in uniform contexts, then this seems to be exactly what Hintikka could say when presented with the closure challenge and the skeptical invitation. The argument for closure so far rests on autoepistemological and rationality considerations but does not necessarily escape Nozick's argument against closure. Since *Knowledge and Belief* Hintikka has emphasized the importance of partitioning the set of worlds into the two distinct compartments consisting of the worlds in accordance with the attitude and the ones not. The worlds in accordance with the epistemic attitude may be read in accordance with Lewis' context-sensitive quantifier restriction on knowledge above. Then, the demon world, brain-in-a-vat world and other derivatives of global underdetermination are simply excluded from the compatibility partition; these extravagant worlds are not in accordance with the epistemic attitude.[21] Thus, these error-possibilities will not disturb the context, or in Hintikkian terms, will not pass over into the compatibility partition, so knowledge is closed for a given compatible partition, i.e. uniform context.[22]

[21]Global underdetermination amounts to the impossibility of reliable knowledge acquisition anyway as Kelly has argued in [33].

[22]There is however not any obvious way to ensure that such a contextual change is not taking place in Lewis' modal theory of knowledge. The rules of actuality and resemblance combined immediately permit for such a change to occur. The demon world resembles saliently the actual world with respect to agent's evidence and should accordingly not be ignored. Lewis readily admits to an *ad hoc* modification of the rule as to exclude this resemblance. Observe that this does not immediately apply to Hintikka's logical epistemology.

One of Lewis' rules seem trivial, and yet it furnishes insight as to Lewis' view of the situation in epistemology today. Knowledge attribution is partly a socially determined process forced by conventional means to be taken seriously. This seriousness is reflected in the *rule of attention*. Which worlds are ignored is context-dependent. When ignored in a specific context these worlds are *really*, not only counterfactually so, tossed out and not to be considered. Attending to even far-fetched possible worlds in a different context make them relevant possibilities again. Relevant possibilities of error undercut infallible knowledge claims and knowledge flies away—becomes elusive.[23]

Buying into too many uneliminated possibilities of error often makes epistemologists end up with buyers regret. Potential counterexamples to knowledge ascriptions are waiting everywhere in the wings of rich domains making the required epistemic position impossible to reach for anybody. No first persons have knowledge in these particularly demanding contexts, no third persons either. Unfortunately, as a discipline epistemology is one such demanding context. The foe of epistemology is not really skepticism but epistemology itself:

> That is how epistemology destroys knowledge. But it does so only temporarily. The pastime of epistemology does not plunge us forevermore into its special context. We can still do a lot of proper ignoring, a lot of knowing, and a lot of true ascribing of knowledge to ourselves and others the rest of the time. [45], p. 377.

Modal epistemology concedes to skepticism the high epistemic standards on which the skeptical position operates. These epistemic standards are exceedingly harder to meet than those required for everyday attributions of knowledge. Admitting this much to skepticism licences the concern that these elevated standards are in fact the correct standards to be met for genuine knowledge ascriptions and acquisitions. When push comes to shove, the everyday knowledge attributions do not stand up to these standards, so knowledge attributions on a daily basis are bogus as discussed by Pritchard [49]. Skepticism can never be dodged. The rules may conflict in such a way that skeptical possibilities like hallucinations become relevant. Applying the prohibitive rule of resemblance merely escapes skepticism by *ad hoc* qualifications. This leaves us again 'caught between the rock of fallibilism,

[23]Ignoring worlds may from this perspective be seen as a necessary last resort because the available evidence may always be insufficient to block global underdetermination. Ignoring is a precondition for knowledge—love it or leave it.

and the whirlpool of skepticism' as Lewis puts it [45], p. 367. Modal epistemology was supposed to come to the rescue.

As bogus as these ascriptions may seem, they may also be as good as it gets. A similar response to skepticism following 'smooth' lines may be found in Levi's formal epistemology [39], [40]. To gain truth and avoid error beliefs should be chosen carrying the highest 'epistemic utility'. The epistemic utility embodies truth as well as content. Significant possibilities of error are forgivable just the agent settles for the belief with the highest epistemic utility in the particular context. This may not exactly add up to real knowledge but it is good enough for decision and action. The elevation of the skeptical standards for knowledge is immaterial for common epistemic practice. Infallibilism with respect to all worlds cannot be reached anyway and agents are doing the best they can quantifying over less reaching at least a workable impasse with skepticism. That is the epistemic balance; Agents can act on their 'discount' infallible knowledge, but skeptics can do very little with their high standards. Turning the tables, skeptics are the *real* epistemologists.

Denials of skeptical hypotheses cannot be known on the modal conception of knowledge trans-contextually. So an objection would be that knowledge is not even possible, much less real. A defense would be to simply admit that the logics of knowledge are rather weak at least for the third person knowledge operator and in case of contextual changes. As opposed to counterfactual epistemology's denial of closure it holds for a first person operator in a uniform context in Lewis' modal epistemology. Closure may fail from the third person perspective because the set of worlds to be considered is strictly a superset of the set of worlds the first person operator has to consider leaving room for radical context change, and a failure. There is support to be found for such a defense.

Levi's epistemological program is a version of a first person perspective emphasizing a distinction between *the logic of truth* and *the logic of consistency* and not the first and third person perspectives [41]. Even though related the two distinctions are not exactly the same. Levi denies the validity of various epistemic axioms as axioms of an epistemic logic of truth. This crudely means to reject these axioms as axioms for a third person knowledge operative. An axiom like the *KK*-thesis found to be invalid in counterfactual epistemology is here valid as an axiom serving regulative purposes of maintaining consistency for a rational epistemic agent. The logic of truth for an epistemic agent on the other hand is not necessarily regulated by a principle like the *KK*-thesis. Lewis seems to follow suit because knowledge of knowledge introduces a discrepancy of actualities for the first and the third person operator. Because of the subject-based contextualism enforced by the rule of actuality, the third person operator is to ignore fewer worlds leaving more room for error. The agent may perhaps know that he knows, the third person may not necessarily be able to determine that the *KK*-thesis holds for the agent, nor that it holds for himself pertaining to the agent in question. The agent in the local environment may have more knowledge than a third person is able to ascribe to him or to the third person himself. If there is a trans-contextual third person logic of knowledge, such a logic is probably rather weak seems to be the suggestion of Levi and Lewis.

While Lewis may consider a universal third person logic rather weak there is nothing in the way of arguing for a much stronger first person logic. This is in stark contrast to the counterfactual proposal of the previous section in which the first person logic was quite weak. On the modal epistemological account all of (1)–(4) may be valid in uniform contexts for a first person knowledge operator.

6. 'ELUSIVE' LOGICAL EPISTEMOLOGY

There is a feature of Hintikka's logical epistemology which may make it become as 'elusive' as any careless mainstream theory of knowledge. The principle of closure, axiom K (4), can under the certain circumstances be generalized to a stronger closure property of an agent's knowledge considered still more unacceptable than (4) itself. *Logical omniscience*:

Whenever an agent Ξ *knows all of the formulae in a set* Γ *and A follows logically from* Γ*, then* Ξ *also knows A.*

In particular, Ξ knows all theorems (letting $\Gamma = \varnothing$), and he knows all logical consequences of a formula which he knows (letting Γ consist of a single formula). Logical omniscience incorporates some generally weaker forms of omniscience like knowledge of valid formulae: Agent Ξ knows all logical truths (given rule 5) etc.[24]

Technical solutions to logical omniscience are either facilitated on the syntactical or semantical level. On the syntactical level, Hintikka recently suggested [28] to place suitable syntactical constraints on deductive arguments which preserve knowledge. Interesting philosophical solutions are to be found on the semantical level. The idea is here to introduce some semantical entities which account for why the agent could be accused of

[24] See [15], [14] for a full list of logical omniscience forms.

logical omniscience but by the end of the day is not guilty of logical omniscience. These entities are called 'impossible, possible worlds' by Hintikka [27]. Similar entities called 'seemingly possible' worlds represented by urn-models are introduced by Rantala [52]. Allowing impossible possible worlds in which the semantic valuation of the formulas in a certain sense is arbitrary provide the necessary means for dodging logical omniscience: The logical laws do not pass muster in the impossible possible worlds. When knowledge is evaluated with respect to all possible worlds but the logical laws do not hold in some of them, logical omniscience is simply out. In an impossible possible world a tautology $A \rightarrow A$ may, as odd as it admittedly sounds, be false. Now the agent Ξ may all the same view that very world a possibility, so universally $K_{\Xi}(A \rightarrow A)$ fails. In consequence, the rule of necessitation (5) is invalid in impossible possible world models. Axiom K is the victim of failure as well. In the impossible possibility both A and $A \rightarrow A'$ may be true while simultaneously A' is false. The failure of axiom K would satisfy Nozick although he probably would consider impossible possible worlds as weird as demon worlds if not weirder. From a strictly logical point of view the epistemic logics specified by impossible worlds models are not very exciting. No real epistemic statement is valid in a universal way. The validity of the various epistemic principles may however be obtained by imposing suitable constraints on the impossibly possible models.

From a forcing perspective the introduction of impossible possible worlds is a rather curious strategy. The idea is to first inflate the local circumstances of the agent in the sense that the agent may regard some models of the (real) world possible. Then afterwards deflate the local situation because of the limited reasoning capacities of the agent. The worlds in question are really logically impossible. For example, a logical contradiction cannot be true. An agent may nevertheless not have enough resources to determine the truth-value of that contradiction and simply assume it to be true. He will consider some worlds possible, although logically they are impossible. To avoid logical omniscience let more worlds in, worlds worse than the demon worlds since the latter are at least logically possible whereas the former impossible possible worlds are not.

7. EPISTEMOLOGICAL AXIOMATIZATIONS

There is a distinct formal feature to both Nozick's counterfactual and Lewis' contextual theories of knowledge. They are in a sense 'formal mainstream' theories as they both observe the significance of epistemic axioms drawn from Hintikka's logical epistemology and their intimate relations to the algebraic properties of the accessibility relation between possible worlds. Nozick considers the accessibility relation to be reflexive while Lewis takes it to be at least reflexive and a sort of symmetric given the rule of actuality and the rule of resemblance respectively. Now, closure holds in uniform contexts, the *KK*-thesis holds, and the rule of necessitation will also immediately hold for a first person contextual epistemological logic. Using the sliding scale devised by logical epistemology to determine validity will make the first person modal epistemological logic at least have epistemic strength on the order of **S4**, perhaps even **S5** is acceptable to Lewis under certain conditions although not discussed. The third person logic of Lewis' contextualism seems to be no stronger than Nozick's first person logic validating (1) and rule (5) which by being so weak is a non-normal modal logic.

Table 1 below summarizes the results pertaining to the validity of common epistemic axioms given the first and third person perspectives on inquiry for logical, counterfactual and modal epistemology

	CE	ME	LE
N: $A / K_{\Xi}A$	1	1/3	1
K: $K_{\Xi}(A \to A') \to (K_{\Xi} A \to K_{\Xi} A')$		(1) / (3)	(1)
T: $K_{\Xi}A \to A$	1	1/3	1
4: $K_{\Xi}A \to K_{\Xi} K_{\Xi}A$		1	1
5: $\neg K_{\Xi}A \to K_{\Xi} \neg K_{\Xi}A$		(1)	

Table 1. CE: Counterfactual Epistemology, ME: Modal Epistemology, LE: Logical Epistemology. 1: First person perspective, 3: Third person perspective. (,): context-sensitive validity

The axioms are in turn answers to skepticism as their validity is sensitive to the forcing restrictions entertained by the various paradigms of knowledge considered above. Nozick's strategy to combat the skeptic is to impose very little relational structure on the universe of possible worlds leaving the skeptic with very little room to manoeuver, thus limiting the skeptic's movement. The strategy of modal and logical epistemology is the opposite: To impose much more relational structure on the universe of worlds (in uniform contexts) leaving the agent with much room to manoeuver, thus enhancing the agent's movement. To combat skepticism, force the skeptic out, either by not giving him a chance to cite distant possibilities of error as relevant, or by making sure that whatever he cites you can reach truthfully at least from the first person perspective.

The common epistemic axioms now furnish a challenging meeting point for mainstream and formal epistemologies ... and there are many others. Some more are to be found in *Forcing Epistemology* [19] others yet uncovered. Let's join the forces and continue what Jaakko Hintikka pursued from the very beginning: To create an interactive epistemology of value to the interdisciplinary study of knowledge.

REFERENCES

[1] Bernecker, S. and Dretske, F. (eds.) (2000). *Knowledge: Contemporary Readings in Epistemology*. Oxford: Oxford University Press.

[2] Benthem, J.F.A.K. v. (2003). 'Fifty Years: Changes and Constants in Logic', in [20]: 35-56.

[3] Binmore, K. and Shin, H. S. (1992). 'Algorithmic Knowledge and Game Theory', in *Knowledge, Belief, and Strategic Interaction,* Bicchieri, C. and Dalla-Chiara, M.-L. (eds). Cambridge: Cambridge University Press: 141–154.

[4] Bull, R. and Segerberg, K. (1984). 'Basic Modal Logic', in ([12]): 1–88.

[5] Chisholm, R. (1963). 'The Logic of Knowing', *Journal of Philosophy* **60**: 773–795.

[6] Chisholm, R. (1977). *The Theory of Knowledge*, 2nd edn. First edition 1966. Englewood Cliffs, NJ: Prentice Hall.

[7] DeRose, K. (1995). 'Solving the Skeptical Problem', *Philosophical Review* **104**: 1–52.

[8] DeRose, K. (2000). 'How Can We Know that We're not Brains in Vats?', *The Southern Journal of Philosophy* **38**: 121–148.

[9] Dretske, F. (1970). 'Epistemic Operators', *The Journal of Philosophy* **67**(1970): 1007–1022.

[10] Dretske, F. (1981). *Knowledge and the Flow of Information*. Cambridge: MIT Press.

[11] Fagin, R., Halpern, J. Y., Moses Y. and Vardi, M. Y. (1995). *Reasoning about Knowledge*. Cambridge: MIT Press.

[12] Gabbay, D. and Guenthner F. (1984). *Handbook of Philosophical Logic*, vol. II: *Extensions of Classical Logic*. Dordrecht: D. Reidel Publishing Company.

[13] Girle, R. (2000). *Modal Logics and Philosophy*. Acumen Publishing.

[14] Gochet, P. and Gribomont, P. (2003). 'Epistemic Logic,' forthcoming in *Handbook of the History and Philosophy of Logic*, edited by Gabbay, D.M. and Woods, J. Amsterdam: Elsevier Science.

[15] Halpern, J. Y. (1995). 'Should Knowledge Entail Belief?', *Journal of Philosophical Logic* **25**: 483–494.

[16] Hendricks, V. F. (1999). 'Methodology in Epistemology', *Danish Yearbook of Philosophy* **34**: 43–64.

[17] Hendricks, V. F. (2001). *The Convergence of Scientific Knowledge—a view from the limit*. Trends in Logic: Studia Logica, vol. 9. Kluwer Academic Publishers.

[18] Hendricks, V. F. (2003). 'Active Agents', ΦNEWS (2002), vol. **2**: 5–40. A revised version of the paper is published in a special volume of the *Journal of Logic, Language, and Information*, van Benthem, J. and van Rooy, R. (eds.), vol. **12**, Autumn 2003: 469–495.

[19] Hendricks, V. F. (2005). *Forcing Epistemology*, forthcoming. New York: Cambridge University Press.
[20] Hendricks, V.F. and Malinowski, J. (eds.) (2003). *Trends in Logic: 50 Years of Studia Logica*. Trends in Logic: Studia Logica Library Series, Kluwer Academic Publishers. Contributions from J. van Benthem, W. Buszkowski, M.L. Dalla Chiara, M. Fitting, J.M. Font, R. Giuntini, R. Goldblatt, V. Marra, D. Mundici, R. Leporini, S.P. Odintsow, H. Ono, G. Priest and H. Wansing.
[21] Hendricks, V. F. and Pedersen, S. A. (2000). *The Companion to Modal Operator Theory—A Program in Philosophy Online*. Department of Philosophy and Science Studies, Roskilde University, Denmark: `http://www.mot.ruc.dk`.
[22] Hendricks, V. F. and Pedersen, S. A. (2002). *Moderne elementær logik*. Forlaget Høst & Søn.
[23] Hendricks, V.F., Pedersen, S.A. og Jørgensen, K.F. (eds.) (2003). *Knowledge Contributors*. Synthese Library, vol. 322. Kluwer Academic Publishers. Contributions from H.v. Ditmarsch, R. Fagin, J. Halpern, J. Hintikka, W. v. d. Hoek, B. Kooi, W. Lenzen, Y. Moses, H. Rott, J. Sowa, M. Vardi and R. Wójcicki.
[24] Hintikka, J. (1962). *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell: Cornell University Press.
[25] Hintikka, J. (1969). 'Semantics for Propositional Attitudes', in *Models for Modalities*, Dordrecht: D. Reidel Publishing Company: 87–111.
[26] Hintikka, J. (1970). "'Knowing that One Knows' Revisited', *Synthese* **21**: 141–162.
[27] Hintikka, J. (1975). 'Impossible Possible Worlds Vindicated', *Journal of Philosophical Logic* **4**: 475–484.
[28] Hintikka, J. (1989). 'Reasoning about Knowledge in Philosophy: The Paradigm of Epistemic Logic', in *The Logic of Epistemology, and the Epistemology of Logic*, J. and M. Hintikka (eds.). Kluwer Academic Publishers: 17–35.
[29] Hintikka, J. (1999). *Inquiry as Inquiry: A Logic of Scientific Discovery*. Kluwer Academic Publishers.
[30] Hintikka, J. (2003). 'A Second Generation Epistemic Logic and its General Significance', in [23].
[31] Hintikka, J. (2003). 'Epistemology without Knowledge and Belief,' *draft*: 35 pp.
[32] Huang, Z. and Kwast, K. (1991). 'Awareness, Negation and Logical Omniscience', *Logics in AI, Proceedings JELIA'90,* Eijck, J. van (ed.). New York: Springer-Verlag: 282–300.
[33] Kelly, K. (1996). *The Logic of Reliable Inquiry.* New York: Oxford University Press.
[34] Kripke, S. (1963). 'Semantical Analysis of Modal Logic', *Zeitschrift für Matematische Logik und Grundlagen der Matematik* **9**: 67–96.
[35] Lamarre, P. and Shoham, Y. O. (1994). 'Knowledge, Belief and Conditionalization', *Principles of Knowledge Representation and Reasoning. Proceedings of the 4. International Conference (KR '94)*, Doyle, J., Sandewall, E. and Torraso, P. (eds.). Morgan Kaufman: 415–424.
[36] Lemmon, E. J. (1977). *An Introduction to Modal Logic*, in collaboration with D. Scott. Oxford: Basil Blackwell Publishers.
[37] Lenzen, W. (1978). *Recent Work in Epistemic Logic*, in *Acta Philosophica Fennica* **30**: 1–219.
[38] Lenzen, W. (2003). 'Knowledge, Belief, and Subjective Probability', forthcoming in [23].
[39] Levi, I. (1983). *The Enterprise of Knowledge.* Cambridge: Cambridge University Press.

[40] Levi, I. (1991). *The Fixation of Belief and its Undoing*. Cambridge: Cambridge University Press.

[41] Levi, I. (1997). 'The Logic of Full Belief', *The Covenant of Reason: Rationality and the Commitments of Thought*, Peters, A. K. (ed.), Cambridge: Cambridge University Press: 40–69.

[42] Lewis, D. (1973). *Counterfactuals*. Cambridge: Harvard University Press.

[43] Lewis, D. (1976). 'Scorekeeping in Language Games', *Journal of Philosophical Logic* **8**: 339–359.

[44] Lewis, D. (1984). *On the Plurality of Worlds*. Oxford: Basil Blackwell Publishers.

[45] Lewis, D. (1996). 'Elusive Knowledge', *The Australian Journal of Philosophy* **74**: 549-567. Reprinted in [1]: 366–384.

[46] Malcolm, N. (1952). 'Knowledge and Belief', *Mind* LXI, 242. Reprinted in *Contemporary Readings in Epistemology*, Goodman, M. F. and Snyder, R. A. (eds.). Prentice Hall. New Jersey: Englewood Cliffs (1993): 272–279.

[47] Moore, G. E. (1959). 'Certainty'. *Philosophical Papers*. London: Allen and Unwin/Unwin Hyman. Reprinted in *Contemporary Readings in Epistemology*, Goodman, M. F. and Snyder, R. A. (eds.). Prentice Hall. New Jersey: Englewood Cliffs (1993): 257–271.

[48] Nozick, R. (1981). *Philosophical Explanations*.Cambridge: Harvard University Press.

[49] Pritchard, D.H. (2001). 'Contextualism, Skepticism, and the Problem of Epistemic Descent', *Dialectica* **55**: 327–349.

[50] Pritchard, D.H. (2002a). 'Contemporary Skepticism', *Internet Encyclopedia of Philosophy*: `http://www.utm.edu/research/iep/s/skepcont.htm`.

[51] Pritchard, D.H. (2002b). 'Recent Work on Radical Skepticism', *American Philosophical Quaterly* **39**(3): 215–257.

[52] Rantala, V. (1975). 'Urn Models: A New Kind of Non-Standard Model for First-Order Logic', *Journal of Symbolic Logic* **4**: 455–474.

[53] Shin, H.-S. and Williamson, T. (1994). 'Representing the Knowledge of Turing Machines', *Theory and Decision* **37**(1): 125–146.

[54] Stalnaker, R. (1968). 'A Theory of Conditionals,' in *Studies in Logical Theory*, Rescher, N. (ed.). Oxford: Basil Blackwell: 98–112.

[55] Stalnaker, R. (1983). *Inquiry*. Cambridge: MIT Press.

[56] Stalnaker, R. (1996). 'Knowledge, Belief and Counterfactual Reasoning in Games', *Economics and Philosophy* **12**: 133–163.

[57] Stalnaker, R. (1999). 'Extensive and Strategic Form: Games and Models for Games,' *Research in Economics* **53**: 291–293.

[58] Voorbraak F. (1991). 'The Theory of Objective Knowledge and Rational Belief, *Logics in AI, Proceedings of the European Workshop JELIA '90*. New York: Springer Verlag: 499–515.

[59] Von Wright, G., H. (1951). *An Essay on Modal Logic*. Amsterdam: North-Holland Publishing Company.

[60] Williamson, T. (2000). 'Tennant on Knowable Truth', *Ratio* 13, **2**: 99-114.

[61] Williamson, T. (2001). 'Some Philosophical Aspects of Reasoning about Knowledge', *Proceedings of the 8th Conference on TARK*, van Benthem, J. (ed.). San Francisco: Morgan Kaufman Publishers: 115–121.

[62] Williamson, T. (2002). *Knowledge and Its Limits*. Oxford: Oxford University Press.

HINTIKKA ON THE PROBLEM WITH THE PROBLEM OF TRANSWORLD IDENTITY

Troy Catterson

It is now almost an established canon in the philosophical literature on modality that there is no problem of transworld identity. Even Kaplan, one of the first to give a precise expression to the problem, has long ago repented of the views which led to his worries.[25] Indeed a survey of the literature on transworld identity reveals that almost nothing has been written on the question since the early 80's. The emphasis, however, should be on the word 'almost' in the last sentence. There has been one philosopher who has continued to insist against conventional philosophical wisdom that there is a problem with the notion of transworld identity. That philosopher is Jaakko Hintikka. I would like to accomplish two things in this paper: 1) I would like to outline the reasoning that has led philosophers to believe that there is no problem of transworld identity. 2) Then I hope to show that Hintikka is right; there is a problem of transworld identity. It is a problem because one cannot decide which theory of metaphysical necessity is correct without first determining the correct theory of transworld identity. Every viable theory of metaphysical necessity will assume some substantive theory of identity.

1. THE ORIGINAL PROBLEM

The minute I take a realistic stance toward possible worlds, and I want to use these alternative scenarios to explain the possibilities with respect to one particular object, a problem arises. Let us use the example of rolling a dice to illustrate the conundrum;.

When I roll the dice, it has a one in six chance of landing on two. This involves an implicitly counterfactual claim, a claim that can be explicated in terms of classes of possible worlds. To say that the dice in my hand has a

[25] See David Kaplan, "Transworld Heir Lines," In *The Possible and the Actual*, ed. Michael Loux (Ithaca: Cornell University Press, 1979), 88-109. For his recantation see the note on the bottom of page 88.

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 33-47.

one in six chance of landing on two is to say that there are six types of possible worlds; in each of these types of worlds, the dice lands on distinct face; and in one of these types it lands on two. For simplicity of exposition, let's imagine that there are six distinct possible worlds: one where the dice lands on one; one where the dice lands on two etc… In each of these distinct worlds we are talking about the same dice. For if the dice in each of these worlds were distinct from the ones in the other worlds then we would cease to be talking about the thing I originally sought to explain: the possibilities with respect to this particular dice in my hand. But now a problem arises: what is the basis for this identification? The minute I take a realistic view of these six distinct worlds, it becomes very difficult to say I have just one dice. In each possible world I have a distinct manifestation of a dice, and I have to find some way of linking them up as manifestations of one and the same dice. I cannot use a mere coincidence of their properties to identify them with each other, since, by hypothesis, they all landed on distinct faces, and hence have distinct properties. Thus I have a problem. The minute I use possible worlds to explain the various possibilities of this dice in my hand, I must assume that this dice exists in a variety of possible worlds. But the minute I assume this I no longer seem to be talking about the one dice. Possible worlds semantics thus seems to undercut the very possibilities it was meant to explain.

Of course, this is not exactly the manner in which Kaplan himself frames the problem. His line of reasoning would go something like this: consider the statement, 'It is possible that Quine never went into philosophy.' According to the possible worlds analysis of the truth of modal statements, this statement is true just in case there is a possible world w where the statement, 'Quine never went into philosophy.' is true. But in order to verify that w is indeed such a world, we must not only show that there is someone in w who is very much like Quine in many important respects but is not a philosopher, we must also be able to identify that person as Quine himself. But that's the problem; certainly this person will differ from Quine in at least one respect: he will not be a philosopher. So we cannot identify them by a mere coincidence of properties. Hence the question arises: How do we identify individuals across possible worlds? If there is no way of identifying individuals across possible worlds, how can statements of possibility or necessity specifically concerned with individuals make sense?

2. KRIPKE'S FAMOUS RESPONSE

Kripke believes that this problem is the product of a misleading way of picturing possible worlds. He states:

One thinks, in this picture, of a possible world as if it were like a foreign country. One looks upon it as an observer. Maybe Nixon has moved to the other country and maybe he hasn't, but one is given only qualities. One can observe all his qualities, but, of course, one doesn't observe that someone is Nixon…So we better have a way of telling in terms of properties when we run into the same thing as we saw before; we had better have a way of telling, when we come across one of these other possible worlds, who was Nixon. (Kripke 1972, 43)

But "A possible world isn't a distant country that we are coming across, or viewing through a telescope…A possible world is *given by the descriptive conditions we associate with it.*" That is to say, possible worlds are suppositions that are constructed, and not realms that are discovered. Our access to a possible world is via our ability to construct scenarios that might have happened. As long as the supposition is possible, then there is no reason why it "can't be a part of the *description* of a possible world that it contains *Nixon* and that in that world *Nixon* didn't win the election" (Kripke 1972, 44).

So there is no problem of trans-world identity because, in constructing counterfactual situations about an individual, we already assume that the situation includes the individual; we are supposing what would have happened to *this very individual*. A possible world, therefore, cannot be given in totally qualitative terms. *Pace* Wittgenstein, the world is not just 'the totality of facts.' We must also include what things exist in it, if we are to have a genuine possible world.

In arguing thus, Kripke is taking his cues from the model theory he originally formulated to provide the semantics for quantified modal logic. Any time one constructs a model for a theory one has to supply the domain of individuals over which the quantifiers in the theory will range. Thus the members of this domain are already assumed by the logician to be completely individuated; the identity relations must already be fixed or given. In the case of a possible worlds model this will mean that the cross world identities must already be given, otherwise one just doesn't have a bone fide model. In order to see this consider the following proposition:

(S) It is possible that someone succeeded in assassinating Hitler.

S can be taken in one of two ways: It can be interpreted as a statement concerning the possible truth of a particular proposition, or it can be taken as a claim concerning an existing individual. This distinction can be made clearer if we utilize the existential quantifier and a possibility operator. The former interpretation of S would look like this:

(S_1) $\Diamond\exists x$(x succeeded in assassinating Hitler)

In possible worlds language S_1 states that there is some possible world where someone succeeds in assassinating Hitler. It could be anyone. Therefore, the truth of S_1 does not depend on any facts concerning cross world identities. But now consider the latter way of taking S:

(S_2) $\exists x$($\Diamond$ x succeeded in assassinating Hitler)

Here we are asserting that someone in the actual world meets a condition in some other merely possible world. So we are assuming that there is a possible world where some individual that exists in the actual world assassinated Hitler. Hence we must assume that there is some individual in the above mentioned possible world which is identical to the individual that satisfies the formula '$\Diamond$ x succeeded in assassinating Hitler' in the actual world; we must assume the truth of a cross-world identity in order for S_2 to be true.

To see this, let us try to construct a model of S_2 that is neutral with respect to cross world identities. To do this we must assume that each possible world's domain is specified internally; that is to say, we only refer to each member of the domain of a possible world with the resources available in that possible world. By doing this we avoid *building* any cross world identities into the specification of the domains of the model because we do not incorporate into our referential apparatus any ways of referring to individuals *across* possible worlds. With this restriction in mind, let us construct a simple model M consisting of $K = \{w_1, w_2\}$, $\psi(w_1) = \{a, b\}$, $\psi(w_2) = \{c, d\}$, $V(A, w_1) = \varnothing$, and $V(A, w_2) = \{c\}$, where A abbreviates the predicate: "x assassinated Hitler." Whether or not we suppose that w_1 is the actual world in this model, it is clear that $M \models S_1$. For the truth of S_1 only requires that some individual in some possible world assassinated Hitler, and c in w_2 fulfills that requirement. And yet it should be equally clear that M is neither a model of S_2 nor its negation. It is the case that S_2 is true at w_2, since there is some individual that exists in that world which assassinates Hitler; hence it also has the property of possibly assassinating him. But, given M's neutrality concerning cross-world identities, we have no way of fixing the truth of either S_2 or its negation at w_1. For in order to know whether there is an individual in w_1 that fulfills the condition, we would have to know whether or not c is identical to some individual in $\psi(w_1)$. This follows directly from the truth conditions for a sentence beginning with an existential quantifier. In order for such a sentence to be true in the proposed model at some world w, there must be some assignment to the variable under the scope of the quantifier that makes the formula without the quantifier true. That is to say, if '$\exists x(P(x))$' is the sentence under consideration, then '$\exists x(P(x))$' is true, if there is some assignment r to x from the domain of w such that '$P(r(x))$' is true. But in the case of S_2, '$P(x)$' is '$\Diamond A(x)$'. Hence, S_2 can be true only if '$\Diamond A(r(x))$' is true. This, however, can be the case only if there is some w` where '$A(r(x))$' is true. Notice that A has to be true of r(x) *in this other world*; it has to be true of the same individual. In M, r(x) is at most identical to a or b. Hence, it is the case that the truth of S_2 at w_1 depends on whether or not c is identical to one of the members of $\psi(w_1)$. But that's just the problem: by hypothesis M doesn't give us this information. If w_1 is the actual world in M, then M is a model of neither S_2 nor its negation. We conclude that we cannot have a full fledged model of sentences like S_2 without assuming certain facts concerning cross-world identities.

Hence, the intuitive plausibility behind Kripke's remarks rests in our very way of specifying the truth conditions for a particular sentence at a particular world. We cannot just state the general facts that will hold in that world and hope thereby to derive the facts that hold of the individuals that exist in that world. For in stating those general facts we must have recourse to the quantifiers in our theory, and the truth conditions for these quantifiers will assume those very individuals and facts. There therefore can be no problem of transworld identity in the object language, because by the time we get there the problem must already be assumed to be solved.

My interpretation of Kripke's response gives the lie to one popular objection to his way of handling the problem. Some have said that Kripke's way of answering the question of what grounds transworld identifications is no answer at all; it just leaves these identifications a mystery. But to think this is to fail to see what Kripke is getting at in saying that there is no problem. He states: "even if there were a purely qualitative set of necessary and sufficient conditions for being Nixon, the view I advocate would not demand that we find these conditions *before* we can ask whether Nixon might have won the election…"[26] In other words, Kripke, in this response, is not claiming that there is no principled way of identifying individuals across possible worlds. Rather he is claiming that, even if there is, we can do modal logic, we can formulate the semantics, without first having to get clear about what these principles are. The modal logician, in constructing a formal theory of metaphysical necessity, can remain neutral as to what theory of identity correctly grounds the cross world identities he assumes in the

[26] Ibid., 47.

stipulation of his domains. Thus as I stated above, there might be a problem with identity, but it is not a problem for modal logic or possible worlds.

3. HINTIKKA'S RESPONSE TO KRIPKE'S RESPONSE

The interesting thing is that Hintikka agrees with all of the data that Kripke utilizes to draw his conclusion. He states that one of Kripke's major insights is perceiving that "quantifying in presupposes that criteria of cross-identification have been given."[27] Thus it cannot be required that these criteria be specified by means of descriptions in the object language, because they too must use quantifiers which presuppose these identifications. But Hintikka draws the opposite conclusion from all of this. It is precisely because quantification into modal or intensional contexts presupposes transworld identification that this becomes such an urgent question. In Hintikka's eyes "we do not have well-defined individuals to attribute properties to without criteria for identification."[28] One might wonder how two equally competent logicians and philosophers of logic could draw opposing conclusions from the same premise. One sees the primitive nature of identity with respect to quantification into modal contexts as evidence that there is no need to worry over how identities are constituted. The other sees this as evidence of the transcendental nature of the question. I think it points to an even deeper level of disagreement between them. This is brought out in Hintikka's diagnosis of Kripke's dismissal of the problem of transworld identity. He states "all we get is explanation by postulation. A fixed store of individuals is first postulated, and then another store of proper names is postulated to enable us to refer to them. But no account is really provided of the constitutive question of what counts as identity between possible worlds…It reduces this conceptual and transcendental question to naïve speculative metaphysics…"[29] In this statement Hintikka is accusing Kripke of two things: first, he thinks Kripke is ignoring the transcendental nature of the question of cross-identification; second, he charges him with naively and gratuitously postulating a fixed store of individuals. Thus as long as we have

[27] Jaakko Hintikka and Gabriel Sandu, "The Fallacies of the New Theory of Reference," *Synthese* 104(2) (August 1995) : 249.
[28] Ibid., 274.
[29] Ibid., 266.

a means to refer to these pre-given individuals across possible worlds, we can skirt the issue of what constitutes their modal persistence.

At first glance the second charge seems rather unwarranted. Kripke admits many times in *Naming and Necessity* that there are individuals that exist contingently if at all. At one point he explicitly states that "we don't require that the objects [referred to by rigid designators] exist in all possible worlds."[30] But that is not what Hintikka is getting at. Later on in the same page he merely states that Kripke's theory implies the existence of a class of individuals that exist necessarily, and are such that "it is logically impossible that there should exist other ones."[31] He never charges Kripke, at least, with holding to the doctrine that all individuals exist necessarily if they exist at all. Instead Hintikka's point seems to be that Kripke is advocating a fixed store of primitive individuals. And this seems to me to get at the heart of one of Kripke's basic assumptions: to Kripke it is nigh well impossible to cash out the identity of individuals, even in one world, in terms of their qualities or relations.[32] Hence, the identity relation is not only primitive with respect to possible worlds semantics, it is metaphysically rock bottom tout court. Our conceptual scheme must start with the individual already individuated, and then go on and theorize about what qualities this individual possesses or could possess, and what relations said individual enters into or could enter into.[33] Thus Kripke is at heart an haecceitist.[34] Of course, if his view of the ontological primacy of the individual is true, then the relation of identity cannot be analyzed into more basic terms, and the whole question of cross-

[30] Kripke, *Naming and Necessity*, 48. Hence his distinction between weak and strong rigid designators.

[31] Hintikka and Sandu, *The Fallacies of the New Theory of Reference*, 266.

[32] See Saul Kripke, "Identity through Time," paper delivered at the Seventy Sixth Annual Meeting of the APA Eastern Division New York 1979. Here he states that we must take for granted the notion of an enduring individual that persists through time. We can readily see how he extends the analogy to modal persistence, or existence across counterfactual situations.

[33] As Kripke, himself, states on page 53 of *Naming and Necessity*: "we do not begin with worlds…on the contrary we begin with the objects which we *have*, and can identify, in the actual world. We can then ask whether certain things might have been true of the objects."

[34] To say that Kripke is an haecceitist is not to say he believes in 'bare particulars'. It is obvious that he is an essentialist. Rather it is to say that he does not believe that the identity relation across possible worlds can be analyzed in terms of properties that uniquely specify an individual in all worlds in which it exists. The identity relation is ontologically primitive, and must be assumed even when we adjudicate the essential properties of an individual. As he states: "Some properties of an object may be essential to it, in that it could not have failed to have them. But these properties are not used to identify the object in another possible world, for such an identification is not needed." Ibid.

identification becomes moot. But it doesn't eliminate the original need to answer the question, because to espouse haecceitism *is* to answer the question, and thus acknowledge the question's legitimacy. Thus, Hintikka is right for thinking Kripke's response rather naïve.

But one has to look elsewhere to see that this is what Kripke is really assuming. In *Naming and Necessity* he does not rely on his haecceitism. His claim is rather that the question as to what constitutes identity across possible worlds is not a question that needs to be dealt with *in order* to do modal logic and possible worlds semantics. He thus denies that it is a transcendental question. Hintikka obviously demurs. In the following section I shall argue that the question of transworld identity is transcendental after all.

4. THE TRANSCENDENTAL NATURE OF THE PROBLEM OF TRANSWORLD IDENTITY

In order to determine who is right, we must first understand what Hintikka means when he calls the question of transworld identity a *transcendental* question. It is obvious from the passages I quoted above that he means to say that formulating a viable theory as to how identity is constituted across alternative scenarios is the precondition for using possible worlds to explain the truth conditions for modal sentences. I viable theory of transworld identity is thus *the condition for the possibility* of a possible worlds semantics.

Now let us formulate this notion a little more precisely. I shall characterize a theory as the union of a set of axioms A with the set of A's consequences . And I shall say that a theory T is distinct from another theory T* just in case the set of T's axioms is distinct from the set of T*'s axioms. T implies another theory T* just in case any model of T is also a model of T*. We may thus define the notion of being transcendental as follows:

(θ) Theories of D_1 are *transcendental* with respect to theories of D_2 *if and only if* there exist theories of D_1 T and T* such that $T \neq T^*$ and for all theories of D_2 T_δ exactly one of the following conditions hold: either (1) T_δ implies T or (2) T_δ implies T* .

θ implies that deciding which theory is the correct account of D_2 forces a decision with respect to theories of D_1, which means that theorizing with respect to D_2 cannot be done independently of, or in isolation from, theorizing with respect to D_1.

By articulating θ we can now see why both Kripke and Hintikka reach different conclusions concerning the urgency of the problem of transworld identity. For one can only say that transworld identity is a problem for the theory of modality if θ holds for theories of modality and theories of identity. It is not enough to say that Modal Logic presupposes these identifications. Thus Kripke's real argument for there being no problem of transworld identity would look something like this:

> Quantifying into modal contexts presupposes that the relevant cross-world identifications have already been made.
>
> Theories of cross-world identity are not transcendental with respect to theories of modality.
>
> Therefore, the problem of transworld identity is not a problem for modal logic and possible worlds semantics.

Now I know of nowhere where Kripke actually justifies his denial of θ. But I think it is fair to view him as thinking along the lines of classical first order logic and its relation to theories of identity. If we assume that we are dealing with first order logic with identity, and we restrict theories of identity to those which make identity an equivalence relation and affirm Leibniz's law, then theories of identity will not be transcendental with respect to theories of first order logic. The laws of logic will be indifferent to one's choice of theories of identity. Of course models of first order logic will assume that the identity relations between members of the domain are already fixed, but how they are fixed will have no impact on our determination of the scope of logical validities. No particular candidate will stand or fall with our choice of some particular theory of identity. Thus, any theory of first order logic will be compatible with any theory of identity that meets the above constraints. If the same kind of situation holds in the case of theories of metaphysical modality and their relation to theories of identity, then Kripke is right; there is no real problem of transworld identity. The only problem is he is not right. What particular theory of identity one affirms makes an important difference in determining what theory of metaphysical necessity is correct, even when we put the same kind of constraints on theories of identity.

But before we argue for the transcendental nature of the problem of transworld identity, we must restrict our domain of discourse. I am claiming this for theories of *metaphysical necessity*,[35] that is to say, theories of

[35] "What I am concerned with here is a notion [of necessity] which is not a notion of epistemology but of metaphysics…consider Goldbach's conjecture…if this is true it is presumably necessary." Ibid., 36. This statement clearly shows that Kripke is

modality that affirm the necessity of the truths of arithmetic and set theory, not restricted systems like S5. I will, however, be assuming that any viable theory of metaphysical necessity is an extension of S5. I am also assuming possible worlds realism. But this should be of no surprise, for the problem of identity across possible worlds would not arise if there were in fact no possible worlds. Now for the argument:

Consider the Barcan formula. Surely no theory of modality worth its salt can afford to remain neutral about this proposition's validity. For its assertion (or denial) has implications for the relationship between unrestricted quantification and unrestricted modality, something about which every viable theory of metaphysical modality should have something to say. The formula is:

(BF) $\Diamond\exists x\, S \supset \exists x\, \Diamond S$,

or, equivalently, in terms of the universal quantifier and necessity:

(BF) $\forall x\, \Box S \supset \Box\forall x\, S$.

Now suppose that there could have been more individuals than there actually are, that the actual world does not contain all possible individuals. Then we have a handy way of falsifying BF. We let α designate the actual world, and we define the one place predicate 'A(x)' as: x has A if and only if x exists in α. Since there could have been more individuals than there actually are, we have the truth of

$\Diamond\exists x\, (\forall y(Ay \supset x\neq y))$.

But the falsity of

$\exists x\, \Diamond(\forall y(Ax \supset x\neq y))$,

since 2 expresses the manifestly absurd proposition that something in α is possibly not identical to anything in α.

This alleged counterexample demonstrates at least one thing: models that validate BF are such that for any w, $z\in K$, $\psi(w) = \psi(z)$. Hence any theory of identity that implies $\psi(w) \neq \psi(z)$, for some w, $z\in K$, will imply the invalidity of BF.

---

presupposing a rather thick notion of necessity, one where mathematical truths are necessary. Thus he cannot be confining himself to a purely logical or conceptual notion.

Parsons, in his paper defending BF, observes that 1 is not self-evident (Parsons 1995). How do we know that this world does not exemplify as many individuals as it possibly could? Recent investigations into large cardinals have explored axioms of set theory which demand largeness without end. All we have to do to make this plausible is acquiesce to the existence of pure sets. And if the existence of pure sets is a matter of necessity, then every world contains as many individuals as it possibly could. "This, together with a large dose of anti-essentialism, undercuts the claim that there might have been things in addition to the things that there are" (Parsons 1995, 10).

In pointing out the need for an anti-essentialist stance in order to undermine this counterexample, Parsons has put his finger on something important, something, as far as I can discern, that has been left entirely implicit in the current literature. There is a close logical connection between the truth of essentialism and BF. In a nutshell, that connection is this: Essentialism is true in metaphysical theories of necessity if and only if BF is false.[36] In characterizing essentialism I shall be sticking pretty close to Parsons' original formulation in his landmark article, "Essentialism and Quantified Modal Logic," subject to a few modifications. A theory of identity is essentialist just in case it affirms a proposition of the following form:

$$(\varepsilon)\quad \exists x_1\ldots x_n(\pi_n x_n \wedge \Box(Ex_1 \wedge \ldots \wedge Ex_n \supset F)) \wedge \Diamond\exists x_1\ldots x_n(\pi_n x_n \wedge \sim\Box(Ex_1 \wedge \ldots \wedge Ex_n \supset F))$$

where $\pi_n x_n$ is a conjunction of formulas of the form $x_i=x_j$ or $x_i \neq x_j$, for every $1 \leq i \leq j \leq n$, but which does not include both formulas,[37] and E is the existence

[36] Strictly speaking BF is consistent with non-trivial essentialism as I define it below. That is to say, there is a model where both come out true. Let $K = \{w_1, w_2\}$, $\psi(w_1) = \psi(w_2) = \{a, b\}$, and $V(P, w_1) = \{a\}$, $V(P, w_2) = \{a, b\}$. It is obvious that BF is valid in this model, since the antecedent is necessarily false. And yet it is also the case that in every possible world in this model there is something that has a property in every world in which it exists, while it is possible that there exists something that does not have said property in every world in which it exists. Thus non-trivial essentialism is valid in this model. However, this result only indicates my argument won't work for theories of *logical* necessity. Certainly metaphysical theories of necessity require the truth such statements as 'there could have been more raccoons'. And certainly the truth of non-trivial essentialism opens up the possibility of beings with mutually exclusive essential properties.

[37] Terence Parsons, "Essentialism and Quantified Modal Logic," *Philosophical Review* 78 (January, 1969) : 52. One might think that the results in Parsons' article already show that Modal Logic is independent from essentialism. But that is to misunderstand Parsons' argument. Parsons, in constructing his anti-essentialist model, is only assuming the truth of S5. Moreover, since the domains of the possible worlds in his model are coextensive, it

predicate. Thus ε has the effect of saying that there are individuals that fulfill certain non-identity conditions F in every world in which they exist and there could have been individuals that do not fulfill F in every world in which they exist.

Once we have essentialism, it is very easy to construct a myriad of plausible counterexamples to BF. Plantinga (1974, 59) undermines the validity of BF by claiming that there could have been a possible world containing all and only abstract or immaterial objects such as propositions, pure sets, and numbers. If essentialism is true, it seems obvious that immaterial objects are essentially immaterial; for surely no proposition could have been a material object. It makes no sense to say that the proposition expressed by the Gödel sentence could have been a tricycle. Hence, in a world where only immaterial objects exist it is true that

Everything is necessarily immaterial.

But, in view of the possibility that there are indeed material objects –a possibility that is amply confirmed by their existence in the actual world, it is false that.

Necessarily, everything is an immaterial object.

Hence, BF is, at the very least, not valid in essentialist models of modality. But it can also be shown to be false. Remember that BF assumes that the actual world is such that it contains all possible objects or individuals. Thus it presupposes that all possible individuals are compossible; there are no individuals whose existence precludes the existence of other possible individuals. This is a very plausible assumption on an anti-essentialist model, but easily falsifiable if essentialism is true. Let Amida be a being who is essentially loving, that is to say, he loves everything in every possible world in which he exists. Let Beelzebub be a being who is essentially hateful; he is hated by every being that can love or hate in every possible world in which he exists. It is obvious, if essentialism is true, that both of these beings are possible in the metaphysical sense. Their existence does not involve any incoherence, and does not violate any clearly necessary truth. But it is also obvious that the existence of the one excludes the existence of the other. For if Amida and Beelzebub were both actual, then Amida would

---

validates BF. Thus his model is not neutral with respect to any theory that affirms or denies BF. His conclusions are therefore not in conflict with what I am attempting to argue.

love Beelzebub. But, since Beelzebub must be hated by everyone, then Amida would also hate Beelzebub –that is to say, not love him, which is a contradiction. Now, if not all possible beings are compossible, then not all possible beings are actual. These are the possibilities: either one exists but not the other, or both do not exist. If only one does not exist, then there is at least one possible being that is not actual. If both, then there are at least two. Either way BF is false.

Suppose you do not find the above example convincing. Clearly many essentialists claim that there are certain properties that are had essentially by anything that has them at all. These are the so-called 'sortals', properties that serve to individuate and provide identity and persistence conditions for whatever is qualified by them. Let *being human* be such a property. Then

$$\forall x(x \text{ is human} \supset \Box(Ex \supset x \text{ is human}))$$

is true. 5 has the consequence that whatever is not in fact human could not have been human. For if there were some possible world in which some existing non-human is human, then that being is human in every possible world in which it exists. Hence it would be human in the actual world, which is a contradiction. Therefore, 11, together with the obviously true assumption that there could have been more humans than there actually are, implies that this world does not contain every possible being.

Indeed, one does not need something as strong as 5 in order to show that essentialism implies the falsity of BF. All one needs to assume is that there is some property which is possibly exemplified essentially but which is not actually exemplified at all. Consider the fact that not every mereological sum is exemplified. Take, for example, the two pennies I have in my hand at the moment. I could have welded the two pennies together to form a new body, which I shall call for convenience sake A. A is simply the fusion of the two pennies in my hand, nothing more, nothing less. From this two facts are evident: First, A, at least, has the property of being the fusion of these two pennies essentially; for that is how we defined A. And second, nothing in the actual world has this property; for these pennies are in fact not fused together. These two facts imply that A does not exist. But it is certainly possible that A exists; these pennies could have been welded together. Hence, there is something possible that is not actual, and BF is false. We have thus shown that the truth of BF depends on the falsity of a well-known theory concerning what it is that constitutes the identity of an individual, essentialism.

If we grant that there is at least one necessary being, that is, a being that exists in every possible world, then the denial of BF will also entail the truth of essentialism. For the denial of BF admits that there is at least one individual that does not exist in the actual world. Hence this being does not exist in every possible world. It therefore does not have the property of *existing necessarily*. But then consider our assumption that there is at least one being that exists necessarily. Certainly, this being has the property of *existing necessarily* in every world in which it exists. Hence, the property of existing necessarily is such that it is had necessarily by some but not all individuals. We thus have a clear instance of essentialism –that is, of course, if the property of *existing necessarily* is a non-identity property in Parson's sense of the term. To demonstrate this it will be enough to show that the assumption of the existence of at least one necessary being and the falsity of BF makes an instance of ε true. The instance of ε that we are concerned with is:

$$\exists x(x=x \wedge \Box(Ex \supset \Box\exists y(y=x))) \wedge \Diamond\exists z(z=z \wedge \sim\Box(Ez \supset \Box\exists u(u=z)))$$

Because we are assuming that there is at least one individual that exists necessarily, the first conjunct is obviously true. The second conjunct states that it is possible that there exists an individual that is identical to itself and is such that necessary existence is not essential to it. The falsity of BF gives us just such a possibility: Since not everything that is possible is actual, there is a possible world w where there is an individual z that does not exist in the actual world. In this world z will obviously be self-identical. But the fact that z does not exist in the actual world will also preclude z from having the property of existing in every possible world in w. Hence z does not have this property in every world in which z exists. Thus the second conjunct is true, and thereby non-trivial essentialism.

But must every viable theory of metaphysical necessity admit the existence of a necessary being? As long as we admit the necessity of the axioms of Zermelo-Fraenkel set theory and possible worlds realism, I think so. Consider the fact that in every possible world at least one world exists, namely the world that is actual from the perspective of that world. This guarantees that it is a necessary truth that at least one thing exists. But if one thing exists then the singleton of that thing exists, by the singleton axiom, from which it follows that the null set exists; for the null set is just the set of all members of the singleton that are self-distinct. Thus the null set exists in every possible world and there is at least one necessary being.

Someone might object that all that follows from this is that something plays the role of the null set in every possible world, not that one thing exists in all possible worlds. But in order to make this objection stick, one has at least to agree that the property of being the null set is instantiated in every possible world. But then doesn't it follow that this property exists in every possible world? How can a property that doesn't exist be instantiated? One cannot respond that this is only the case if one accepts property realism, because by eliminating proper names in favor of descriptions like *the unique x such that x is the null set* one must recognize the existence of the corresponding attribute,[38] otherwise we must commit ourselves to the even more audacious claim that the linguistic entity, the predicate, exists in every possible world. So either way we must conclude that there is at least one necessary being. Thus any theory of metaphysical necessity that admits the existence of possible worlds and repudiates BF must espouse some form of essentialism.

We may thus summarize our argument for the transcendental nature of the problem of transworld identity as follows:

> There are at least two theories of identity I_h, I_e such that $I_h \neq I_e$, and, for all T, $T \cup BF \models I_h$ and $T \cup {\sim}BF \models I_e$.
>
> Every viable theory T of metaphysical modality is such that either $T \models BF$ or $T \models {\sim}BF$.

Hence, there are at least two theories of identity I_h, I_e such that $I_h \neq I_e$, and, for all T, exactly one of the following conditions hold:

1) T implies I_h;

2) T implies I_e.

Therefore, by θ, theories of identity are transcendental with respect to theories of metaphysical modality.

Hintikka was right after all. There is a problem of transworld identity, and it is a problem despite the fact that we must assume cross-world identities in order to get possible worlds semantics off the ground. Indeed, as a consequence of the transcendental nature of the question, it is a problem precisely because of this assumption.

[38] Even Quine admits that his strategy of eliminating proper names in favor of descriptions 'seems to commit us to recognizing that there is a corresponding attribute…' See Willard Van Orman Quine, "On What There Is," in *From a Logical Point of View* (Cambridge: Harvard U.P., 1980), 8.

WHAT IS EPISTEMIC DISCOURSE ABOUT?

Radu J. Bogdan
Department of Philosophy
Tulane University

1. INTRODUCTION

In the 1960's and 70's Jaakko Hintikka has written extensively about epistemic logic, epistemic concepts and ordinary epistemic discourse. As a (graduate) student of Jaakko's toward the end of that period, I was somewhat familiar with that body of work and even discussed some fragments of it in my dissertation on the pragmatics of knowledge. Since then my interests developed in different directions, toward philosophy of mind and cognitive science in general and commonsense or naive psychology in particular. This paper looks back at Jaakko's work on epistemic discourse from the vantage point of my later work on naive psychology.

The title question, about the subject matter of epistemic discourse, is not an easy one to answer, for several reasons, surveyed in the next section. These reasons bear on the tight yet not fully identified and understood links between epistemic discourse, on the one hand, and naive psychology, ordinary language, and epistemological expectations, on the other hand. Section 3 focuses on the relation between naive psychology and epistemic discourse, surveys some empirical data about the development of epistemic discourse in the context of a more general mental development, and suggests the need for a top-down explanatory approach to our competence for epistemic discourse. In section 4 I propose to read Jaakko Hintikka's work on epistemic discourse in cognitive-scientific terms and view it as contributing to such an explanatory approach. A concluding section 5 suggests deeper evolutionary reasons why epistemic discourse would likely work along the lines such as those suggested by Hintikka's account.

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 49-60.

2. PROBLEMS

The notion of epistemic discourse is meant here to include epistemic terms, such as those for perception, belief, memory or knowledge, and epistemic locutions, such as 'she knows (believes, perceives) that p.' Epistemic discourse is part of a broader mentalist discourse that also contains terms and locutions for various other sorts of mental states and intentional relations, such as desire, intention, fear, regret, and so on. The mentalist discourse and its epistemic subset express in ordinary language a body of knowledge or a competence variously known as naive or folk or commonsense psychology or, somewhat more technically, theory of mind. Naive psychology is credited with enabling us to conceptualize, explain and predict what is going on in other minds and our own, and how these mental goings on translate into behaviors. Since naive psychology is the home base of our epistemic concepts and discourse, our understanding of the former is bound to affect our understanding of the latter. Whence the first problem: in different ways and terms, philosophers and psychologists disagree rather sharply over the nature, job and modus operandi of our naive psychology, and this disagreement is inevitably echoed by disagreements over what epistemic discourse is all about and how it works. I will return to this problem in the next section.

But even if this first problem were solved, there is a second problem to be faced. Ordinary epistemic discourse appears to have duties that go much beyond the basic job description of naive psychology—or, if we choose to look at it somewhat differently, duties that would expand the job description of naive psychology in directions that are even more elusive and controversial than those of its basic job description. These other duties, emanating from the nature of linguistic communication and social interactions, bring a host of contextual and pragmatic parameters into our epistemic discourse and weave them—in ways still poorly understood -- into the fabric of naive-psychological concepts and attributions.

The problems just surveyed are problems of substance that will be in focus throughout this paper. There are also methodological problems of how to approach and study epistemic discourse. Philosophers have done more work on epistemic discourse, and ventured more explanations of it, than either psychologists or linguists. Since I will review some psychological data and hypotheses in the next section, I should say a few words now about the philosophical approaches.

Among the ways in which philosophers and logicians have gone about studying ordinary discourse in general, one stands out. It consists in making sense of and explicating the linguistic intuitions of the speakers of a natural language. In the case of epistemic discourse, we are talking about the epistemic intuitions that speakers have about knowledge or belief or memory claims and attributions. I use the notion of intuition here in a loose analogy to Noam Chomsky's notion of grammatical intuitions that speakers of a natural language have in distinguishing grammatically correct from incorrect uses of words and sentences in that language. They can do so without necessarily knowing how they do it, by what rules—at least not until properly schooled. The same is thought to be true of the intuitions speakers have about epistemic discourse.

One serious problem with basing philosophical reconstructions on epistemic intuitions is that it is not always clear whether the intuitions in question are used to support an epistemological analysis or, alternatively, an epistemic-discourse analysis—that is, whether the intuitions bear on what the concepts of knowledge or belief ought to be or, alternatively, on elucidating our ordinary discursive practices involving knowledge or belief claims. To see what the distinction is and why it matters, suppose we ask, for example, what the concept of knowledge is and when it is attributed. On the epistemological reading, we are asking what knowledge is or what it takes to have knowledge in general and in what conditions knowledge can be said to be instantiated in some organism or system. This is the question that philosophers have asked forever but whose answer is still rather elusive. Since Plato, most answers have taken the form of conceptual analyses that unpack a typically idealized or normative concept. The standard analysis is that of knowledge as justified true belief, but there are other analyses as well, including naturalist accounts that replace justification with causation or reliable information processing. In the heyday of Gettier-like games that epistemologists used to play with gusto (but mercifully, no more), the examples of and counterexamples to some definition of knowledge were checking epistemological intuitions with an eye to this conceptual project, even though many of these intuitions may have originated in the use of epistemic discourse.

The epistemic-discourse reading, on the other hand, is concerned solely with the epistemic terms and locutions used normally, rather descriptively and usually contextually by language users. The analytic project here is to figure out and explain the rules of and constraints on the ordinary use of such terms and locutions. The distinction between the two readings allows us to say, without contradiction, that one can make a knowledge claim to the effect that he knows that p even though, on some epistemological analysis, he does not really have that knowledge because, for example, he lacks appropriate justification. Perhaps the best known and most insightful analyses of epistemic discourse can be found in Wittgenstein's later writings, in some of Gilbert Ryle's, John Austin's and those of other philosophers of ordinary language (Urmson, Malcolm, etc.).

Yet even when the territory is clearly demarcated, and epistemic discourse is seen for what it is, independently of epistemology, there is no guarantee that the linguistic intuitions associated with it would or could reveal how it works and why. As the main source of insights about and analyses of epistemic discourse, ordinary language philosophy is programmatically descriptive and rather uninterested in explanation. Explanation requires taking linguistic intuitions and practices at best as data or clues pointing to deeper causes or reasons why they work as they do. If epistemic discourse is handled by a psycholinguistic competence or expertise, then explanation requires a theory of that competence or expertise. The first order of business, in the next section, is to establish the antecedent of this conditional and get a sense of the competence in question. In the section following it, I propose to look at Jaakko Hintikka's work on epistemic discourse as contributing to an explanatory theory of that competence.

3. NAIVE PSYCHOLOGY AND EPISTEMIC DISCOURSE

Making epistemic attributions in ordinary contexts by employing an ordinary epistemic discourse requires the resources of naive psychology, which is our competence to recognize and represent how other minds and our own relate to the world in perception, desire, intention, thinking, or memory. The most systematic empirical study of naive psychology and its language has been so far undertaken by developmental psychologists.

Despite inevitable and often sharp disagreements over the nature of naive psychology as a mindreading competence, most researchers agree on some innate and prelinguistic basis for the competence and also agree on several age-related milestones in the development of the competence. The earliest abilities to detect and represent basic intentional relations, such as looking at, seeing, trying to, and wanting, begin to be exercised by children before the age of 1 and thus before the onset of language. It is likely that these early and prelinguistic naive-psychological and epistemic insights might influence and perhaps constrain the later development of epistemic discourse and of the concepts based on it. Thus, some psychologists think that the concept of belief may be modeled on the earliest and prelinguistic concept of perception or gaze. In close analogy, an early version of the concept of knowledge may be modeled on the prelinguistic concept of seeing, in the sense of successful perception or, more generally, successful access to information (see Perner 1991 for a general survey; also Bogdan 1997).

The main implication for our discussion is that the earliest and most central concepts of naive psychology are prelinguistic and thus owe nothing to the rules and practices of linguistic communication in general and mentalist and epistemic discourses in particular—although they may owe much to prelinguistic interpersonal interactions (Bogdan 2000, chapter 3). Epistemic discourse itself seems to have its own developmental schedule, although it has been less investigated than the developmental schedule of naive psychology. Still, there some pertinent data, which I will report telegraphically and then weave into our discussion. The main sources are Bartsch and Wellman (1995) and Nelson (1996).

The former authors distinguish several phases in the child's acquisition of epistemic terms. The first terms, acquired a few months after the age of 2, are WANT and LIKES. Around 3 or soon afterwards emerge DESIRE, BELIEF and THOUGHT. Around 4 the term for BELIEF is used for explaining actions, first those of others, before those of self. Only around 4 do children begin to distinguish between KNOW and THINK or GUESS, although they do not seem to distinguish between THINK and GUESS before the age of 8. What do these data mean? They clearly show an emerging mastery of epistemic terms and attributions. But what exactly do these terms and attributions represent at each developmental stage? This is a difficult question that psychological research has not yet answered fully. There are two main reasons for the difficulty. The first and most important is that there is no agreement among philosophers or psychologists over what epistemic discourse in general is about. The nature and function of this discourse are still to be plausibly defined. Do epistemic terms and locutions represent what is going on in the minds of those targeted by them? Or do such terms and locutions represent something entirely different, such as conversational appropriateness, evaluation of information and evidence, prediction of behavior, and the like? Or are epistemic representations targeting a mixture of mind and outside factors in combinations still to be figured out?

This indeterminacy reflects a parallel but deeper uncertainty and hence disagreement about the nature and function of the more basic and broader competence for naive psychology. Again, is naive psychology directed at minds, our own and those of others, or at something else, of which minds may be only a part? What is clearer already is that the first intentional relations to be represented by very young children (and possibly great apes), such as gaze, attention, and behavioral postures and motions indicating goals or simple desires, are the most visible, informative and causally manipulable relations. There is also a rather wide consensus that the naive-psychological categories of these relations are likely to be innate and modularized or at least to reflect innate predispositions. These properties make evolutionary sense (Bogdan 1997). It is also likely that the meanings of the first mentalist words, such as SEE, WANT or DESIRE, track closely the relevant perceptual experiences that activate these primordial categories.

The more difficult problem is determining what happens in the second and later phases of the acquisition of epistemic and, more generally, mentalist vocabularies, when the higher-level and more abstract categories of belief, opinion, thought, and knowledge no longer track perceptual experiences and depend increasingly on linguistic descriptions and other social and contextual factors. The child's naive psychology now becomes inextricably linked—indeed woven into—the mentalist and specifically epistemic discourses. As a result, it is during these later phases that the child assimilates most of the adult epistemic meanings and the language games in which they are involved. It appears that this assimilation process is complex, difficult and protracted, with comprehension emerging earlier than production. It is symptomatic that the child's meaning of KNOW remains different from the adult meaning, and keeps changing, until about the age of 12.

So what is going on? Hard to tell, but a few developmental facts point to a dramatic mental revolution that affects naive psychology during the second, nonperceptual phase that begins around the age of 4, when it gradually moves beyond representing here-and-now situations and tracks more abstract attitudes, such as nonperceptual belief, intention or knowledge. Two contrasting metaphors may help clarify this transformation. Until around the age of 3 to 4 the young mind operates on a single screen, where perceptual inputs of current events are displayed and constantly updated by new inputs. It is a mind largely confined to current motivation and perception, controlled by a unique focus of attention, and representing things on a single mental screen. After 4, the young mind (mostly its prefrontal cortex) is shaken by several mental commotions, executive as well as cognitive, and revolutionary in their cumulative impact. Chief among them are the inhibition of current perception, the linguistic recoding and representational explicitation of earlier procedural competencies, such as counting, mental imagery, and naive theories of various domains, including naive psychology, and the ability to deploy multi-layered clusters of thoughts and to embed thoughts into other thoughts. These changes liberate the young mind from the captivity of single-screen mentation and enable it to entertain simultaneously, on different and interconnected mental screens, nested sets of alternative and often conflicting representations of actual and nonactual, current, past and counterfactual situations. The single screen of early childhood is replaced by a multi-screen or multiplex mentation.

Among the new mental activities made possible by the emerging multiplex mind, two are relevant to our discussion. One is the imaginative and often counterfactual access to nonactual, future as well possible situations or worlds. As a result, it becomes now possible to reconfigure earlier naive-psychological categories and to conceptualize new propositional attitudes, such as thought, intention or knowledge, in terms of possible worlds -- at least when the default attributions fail. Hintikka's possible worlds semantics for epistemic attributions thus has some psychological bite. The point is not that the young or adult epistemic attributor envisages possible worlds whenever she makes a belief or knowledge claim for herself or others. The point is that she could do that when the need arises—for example, in contexts of uncertainty, doubt, criticism, incomplete evidence, high stakes, rigorous inquiry, and so on. And the further related point is that she would not recognize epistemic ascriptions and the concepts behind them for what they are, if, for some reason, thinking and talking in terms of possible worlds would be always unavailable. The other new mental activity made possible by multiplex mentation is integrating many sources of information across several modes of representation -- language, memory, perception, imagination, inference, and so on. Epistemic attributions require such integration—for example, in iterating attributions involving different attitudes, such as belief, perception and memory, as in 'he believes that she remembered seeing him going home.'

The point of these remarks is that the naive-psychological and epistemic concepts and attributions available to and employed by a multiplex mind are vastly different and more complex from those of the younger uniplex mind. I think that neither observations nor experiments nor analyses of linguistic intuitions are sufficient to reveal what the former are all about. What is needed is a theory of the cognitive tasks that the new naive-psychological and epistemic concepts and attribution abilities are designed to carry out -- in other words, a theory in the spirit of recent cognitive science. From a logical and semantic angle, this is how I read Jaakko Hintikka's work on epistemic logic and how I think it can contribute to a better understanding of epistemic discourse.

4. A THEORY OF EPISTEMIC MEANINGS

I begin with a familiar picture of cognitive-scientific explanation, in whose terms I want to frame my discussion. To understand the mind as a system of mechanisms that act on information in pursuing its goals, the theorist must figure out the programs run by the mechanisms and thus the competencies involved in processing information and acting on it; but to figure out the programs, the theorist must antecedently identify the tasks the programs execute and the problems encountered and solved in the execution. The analysis and explanation thus proceed top-down: from Tasks to Programs to Mechanisms to Ware (hard, wet, whatever). I abbreviate it as the T->P->M->W method. Thanks to the influential work of Noam Chomsky on language, David Marr on vision, and Allen Newell on artificial intelligence (to cite pioneers), the T->P->M->W method has achieved classic status in cognitive science.

One indication that Hintikka's angle on epistemic discourse fits the explanatory methodology of cognitive science is that he is no friend of intuitions as the primary basis for philosophical or formal analysis. According to Hintikka, when not wrong, which they may often be, intuitions tend to lead not to the concepts or abilities they purport to illuminate but to some collateral relations or indirect associations. This is why a deeper analysis is needed. It is in this spirit that Hintikka conceives of epistemic logic as an explanatory model of the workings of ordinary language. It brings out the "deep logic" (which I read as: core tasks) underlying epistemic discourse (Hintikka 1969, 3-5). It should be noted that Hintikka's work in many other domains—such as inductive logic, the logic of questions, mathematical reasoning, and game-theoretical semantics -- is also intuition-free, theory-driven and task-oriented.

Hintikka writes that, "as the case is with theoretical models in general, it [the explanatory model] does not seem to be derivable from any number of observations concerning ordinary language. It has to be invented rather than discovered" (1969, 5; with, significantly, a footnote reference to Chomsky). Hintikka is thinking in the same spirit as Chomsky about the tasks of epistemic attributions. Hintikka points out that the explanatory model embodied in epistemic logic does not reproduce what is found observationally or intuitively in ordinary discourse as surface phenomena. The latter at best point to the tasks of the "deep logic" of our epistemic-language competence, just as the surface grammars of English or Chinese point to the computational tasks of the "deep grammar" that characterizes our grammatical competence in general. In the case of epistemic discourse, not only are the surface phenomena distinct from the deep-logic tasks but they are constantly influenced by all sorts of collateral interests and pressures, such as conflicting goals, pragmatic considerations, cognitive limitations, and contextual factors.

Given all these considerations, Hintikka's suggestion is to treat as basic the meaning of an epistemic expression captured by the explanatory model and then view its modifications and variations by the collateral factors just cited as residual meanings (Hintikka 1969, 6-7). The actual use of the expression reflects the specific relation between the basic and the residual meanings, which is the relation between what the expression (through its terms and operations) is designed to convey according to its deep logic (i.e., its core tasks) and the collateral conditions of its use.

This, quite roughly, is the line of metaepistemic analysis pursued by Hintikka in his classic Knowledge and Belief (1962) and many other works, including the collection of historical essays, Knowledge and the Known (1974). The distinction between basic and residual meanings is best revealed in his equivalence claim that for one to know is to know that one knows. Call this the KK equivalence. Its critics, according to Hintikka, failed to see that the equivalence concerns the basic meaning of a knowledge claim (what its deep logic conveys) and not its surface variations in ordinary discourse, due to collateral interests. In its different surface manifestations the KK equivalence is bound to break down most of the time, precisely because of collateral interferences. 'Being certain' or 'being aware' or 'having enough evidence' are expressions of residual epistemic meanings that often defeat the KK equivalence for contextual and pragmatic reasons. But these are not the reasons why the KK equivalence holds fundamentally. This is why, according to Hintikka, the ordinary language analyses of epistemic terms and locutions highlight variety and diversity but fail to address their deep logic or basic meanings or core tasks. Hintikka is an essentialist realist about the deep logic of ordinary epistemic discourse whereas most ordinary-language analysts are postmodernist impressionists.

Having sketched the broader picture, we can now ask the key question: What is the deep logic or basic meaning of a knowledge claim and hence of the KK equivalence? Recall that the question is not, epistemologically speaking, about having knowledge or instantiating it in some form. The question is about an item of epistemic discourse, in particular a knowledge (or some other epistemic) claim or description made explicitly by a speaker of a language. The answer is that if one knows something, one ipso facto knows that one knows, because the same circumstances that would justify one in saying 'I know that I know' would justify one in saying 'I know' simpliciter (Hintikka 1962, chapter 5). There are demonstrably no situations or possible worlds in which one claim would be true and the other false. The reason is that one always knows what one is thinking when saying something, such as making an epistemic claim; for not knowing it would be epistemically indefensible or inconsistent. This, then, seems to be the core task of a knowledge claim: to ascertain that the one making the claim has as good a justification as there can be (in terms of all possible eventualities) and that further doubt or criticism are beyond the point. It is the task of discussion (criticism, inquiry)-stopper (Hintikka 1962, 111; 1969, 13).

Notice that this account fits the main joints of the standard epistemological analysis. When one makes a knowledge claim, one presupposes that one has a true belief that is justified. The difference is that in the case of epistemic discourse and its knowledge claims, the question of evidence and justification (which frustratingly eludes most epistemological analyses) is settled, as it were, by the deep-logic design of the enterprise. That is the very point of making a knowledge claim, its core task -- to indicate an end to inquiry and to the pursuit of evidence and justification.

It may appear that in the first-person case, the KK equivalence entails that mental states are transparent to self or self-intimating or introspectable. But Hintikka does not take self knowledge to be a mental state and therefore one's self knowledge claim does not say anything substantive about one's own mind -- except that it made up its mind, so to speak, to conclude an inquiry or the search for evidence, and so declares publicly. For Hintikka, the deep logic of epistemic discourse has no room for privileged access or incorrigible authority. Although the focus here has been on knowledge—perhaps the epistemic-discourse notion most systematically investigated by Hintikka—I expect similar conclusions to be drawn, mutatis mutandis, about the deep logic of other epistemic terms and attributions. Contrary to many historically and recently fashionable views, Hintikka's analysis of its deep logic or core tasks suggests that epistemic discourse in general is not about the mind, nor about the vagaries of context and conversation. My reading of his analysis is that the basic job of epistemic discourse is to inform publicly about the range of actual and possible situations or worlds compatible with a given intentional attitude (or a sequence of attitudes) of the person discoursed about. I find this reading congenial to a larger picture I draw of naive psychology and mentalist discourse, as I explain in the next and concluding section.

5. A DEEPER 'WHY?'

Why would the deep logic of epistemic discourse work the way Hintikka proposes, at least according to my reading? To answer this question, we need to go beyond the confines of the classic top-down method of explanation in cognitive science, the T->P->M->W method mentioned at the beginning of the previous section. To see what I mean, consider a methodological distinction that is familiar in evolutionary biology. It is the distinction between proximate and ultimate explanations of biological traits, in particular competencies. A proximate explanation tells us how a competence works, according to what program, executing which tasks. But it does not explain the reason for the tasks themselves and hence for the design of the program and the modus operandi of the mechanism running the program. It is the business of an ultimate or evolutionary explanation to identify the deeper reason. The implication, then, is that the T->P->M->W method organizes and provides an order to the proximate explanations of mental capacities but does not tell us why these capacities exists and why they evolved. The why question must be answered at a higher level of theorizing, that of the evolutionary function (E) of the tasks and programs under scrutiny. The classic method must be augmented to take the E->T->M->P->W form.

In proposing this addition, I argued in an earlier work that, unlike the more transparent tasks of our competencies for vision or grammar, the tasks of naive psychology are not obvious without a careful look at their evolutionary function (Bogdan 1997, chapters 3 and 5). When their evolutionary function is factored in, it becomes apparent and plausible that the tasks of naive psychology are to detect, represent and categorize those relational (not intrinsic) properties of other individuals, which are mentally intentional as well as behavioral, and which the naive psychologist can use causally to engage, influence or otherwise exploit in order to pursue her goals in a variety of social and communicational contexts. In other words, naive psychology is an evolved mental tool kit that services the active goals of the naive psychologist when interacting with conspecifics.

This analysis works best in the case of simple, most visible and informative as well as causally manipulable intentional relations, such as gazing, noticing, seeing, or wanting. The categories of such relations may indeed have evolved by natural selection. To get a flavor of the analysis, consider the competence to represent gaze -- a basic pillar of primate naive psychology (Bogdan 1997, 137-138). The metaintentional category of gaze, underlying this competence, contains instructions and procedures to represent those aspects of someone’s gaze that predict behavior and allow causal interventions that meet the goals of the naive psychologist. This idea can be unpacked and illustrated in the following analysis of the gaze category:

(a) eyes open → alertness and propensity for behavior → involvement

(b) eyes open + line of regard → interest and its general direction → involvement

(c) eyes open + line of regard tracked → goal to be identified or the direction of a behavior to be initiated or something happening somewhere → involvement

(d) eyes open + line of regard tracked + the target of the line of regard identified → specific goal → involvement

A human child or a chimpanzee can use these different components of the gaze category to find out about others and the situations they are in or will be in, and also to interfere with those situations or the actions of others. For example, the (b) instructions enable a naive psychologist to obstruct the line of regard of a gazer and prevent him from seeing something of interest to the naive psychologist. Apes are known to do this trick quite often.

Consider now the much more sophisticated naive psychology of human adults tracking complex, invisible and linguistically expressed propositional attitudes, such as opinions, memories, intentions, thoughts, and claims to knowledge. Epistemic discourse becomes the main avenue not only to the identity of such attitudes but also and crucially to what the attitudes inform about—mainly people and situations, actual, past or possible—and to the opportunities of interference, manipulation or utilization afforded by this informativeness of the attitudes. Looked at from the perspective of its initial evolutionary rationale, the naive-psychological game played with the epistemic discourse about propositional attitudes is not that different in its basic tasks from the much primitive game played with the much simpler metaintentional categories, such as gaze. As I read it, Jaakko Hintikka's analysis portrays the basic epistemic meanings and their deep logic in the same instrumental light, as providing information on attitudes and the possible situations in which they hold—information that can be exploited in some way or put to some use.

REFERENCES

Bartsch, K. and Wellman, H. (1995) Children Tall about the Mind, Oxford: Oxford University Press.

Bogdan, R. (1997) Interpreting Minds, Cambridge, MA: MIT Press.

Bogdan, R. (2000) Minding Minds, Cambridge: MA, MIT Press.

Hintikka, J. (1962) Knowledge and Belief, Ithaca, NY: Cornell University Press.

Hintikka, J. (1969) Models for Modalities, Dordrecht: Reidel.

Hintikka, J. (1974) Knowledge and the Known, Dordrecht: Reidel.

Hintikka, J. (1975) The Intentions of Intentionality, Dordrecht: Reidel.

Nelson, K. (1996) Language in Cognitive Development, Cambridge: Cambridge University Press.

Perner, J. (1991) Understanding the Representational Minds, Cambridge, MA: MIT Press.

INTERROGATIVE LOGIC AND THE ECONOMIC THEORY OF INFORMATION

Raymond Dacey
University of Idaho
Moscow, ID 83844-3161

1. INTRODUCTION

It is with great pleasure that I return to earlier work based directly (Dacey 1979b) and indirectly (Dacey 1979a, 1981a, 1981b, 1985) on the fruits of Jaakko Hintikka's labors. In particular, it is a great pleasure to work in the area of Hintikka's interrogative logic, which provides a systematic approach to the many subfields of philosophy (Hintikka 1968/1999, 1973, 1974, 1975, 1976, 1985, 1993; Hintikka and Hintikka 1989; Hintikka, Halonen, and Mutanen forthcoming/1999; Hintikka and Kulas 1983). One open area within interrogative logic pertains to situations where the inquirer receives "merely probable answers [from] Nature" (Hintikka 1988/1999, p. 156). One approach to this open area is the economic theory of information as advanced by Jacob Marschak (1971, 1974). Indeed, Marschak shows that the economic theory of information can be employed within the context of Hintikka's general interrogative logic to account for probable answers as they appear in both pragmatic (i.e., economic) and inductive (i.e., epistemic) decision making (Marschak 1974, particularly pp. 145-149 and pp. 167-169).

The present paper attempts to reveal the robustness of Hintikka's interrogative logic by applying the union of interrogative logic and the economic theory of information to examine an aspect of the sexual selection problem of biology. The two-part claim made and supported here is that Hintikka's interrogative logic is strikingly robust and renders testable one of the two biological theories of male ornamentation.

The biological decision problem is both fundamental and structurally simple. Suppose a female is about to select a mate. What questions can the female ask a contending male that will move her to choose or refuse that male as a mate? Similarly, what answers can a contending male provide in order be selected as the mate? The answers to these questions, when asked

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 61-74.

about the human species, are remarkably difficult to answer. I will therefore focus on a "lesser," albeit more colorful, species. Consider the peacocks. The females are dull and drab in coloration, whereas the males are ornate and flamboyant in coloration. The fundamental question a female asks of a male is "Do you have good genes?" The fundamental answer provided by a male is "I do not know, but I do have this marvelous tail of many colors."

The female faces one of two decision problems. In the <u>pragmatic decision problem</u>, the female must decide between mating and not mating with the male. In the <u>inductive decision problems</u>, the female must decide between believing that the male has good genes and believing that the male does not have good genes, or believing the generalization that all beautiful males have good genes. The pragmatic decision problem involves a payoff table with traditional gains and losses, and can be resolved via the decision rules of Pascal (Arnauld and Nicole 1662), von Neumann and Morgenstern (1947), or Kahneman and Tversky (1979, 1992). The inductive decision problems involve information-theoretic payoffs, and can be resolved via the pure truth seeker decision rule (Marschak 1974, pp. 170-173) and the maximum expected content decision rule (Hintikka 1968/1999, p. 219).

In what follows, we will show that both pragmatic and inductive decision problems, however they are resolved, yield a result that renders testable the good genes view of male ornamentation. Thus, we can conclude that Hintikka's interrogative logic is very strikingly robust.

2. MALE ORNAMENTATION

Sexual selection in almost all species involves choice problems for the females, typically based on competition among the males. Male competition often involves ornamentation that is costly.

> One large class of apparently costly characters are those found usually only in males and which Darwin called secondary sexual characters. … The peacock's tail (or, more exactly, train) is an example. … The peacock's tail almost certainly reduces the male's survival (though this has never been demonstrated): the tail reduces maneuverability, powers of flight, and makes the bird more conspicuous; its growth must also impose an energetic cost. (Ridley, Mark 1993, p. 283).

Note that the notion of 'costly' employed here involves a cost borne by the male and not by the female. In Marschak's terms, this is a cost of communicating. See Marschak (1971, 1974) for the formal treatment of the costs pertaining to information systems and rational inquiry. We will ignore these refinements here.

There are two views on secondary sexual characters. The first view, associated with Sir Ronald A. Fisher, is commonly called the "sexy-son" view, and holds that the secondary sexual characters do not carry information about the quality of the male's genes. This view holds

> that females choose males according to the gaudiness of their colors, the length of their plumes, the virtuosity of their songs, or whatever, because the species is ruled by an arbitrary fashion for preferring beauty that none dares buck (Ridley, Matt 1993, pp. 142-143).

The alternative view, associated with Alfred Russel Wallace, is commonly called the "good-genes" view, and holds that the secondary sexual characters do indeed carry information about the quality of the male's genes. This view holds that "Ornaments and displays are designed to reveal the quality of the genes" (Ridley, Matt 1993, p. 143).

The claim made here is that the good-genes view is testable given that we employ the economic theory of information as a particular form of Hintikka's interrogative logic. The testable proposition, and thereby the warrant for the claim, is derived in each of the two decision-theoretic models of the female's choice problem.

3. THE PRAGMATIC DECISION PROBLEM

The good-genes view posits that male ornamentation carries information about the quality of genes. As such, the presence or absence of the ornamentation can be taken as a partial (or uncertain) answer, sent by the male to the female, to the female's question "Do you have good genes?" To construct the payoff table for the pragmatic decision problem, presume the female has two choices – mate with the male at hand or do not mate with this male – and faces two states of nature – the male at hand has good genes and the male at hand does not have good genes. If the female selects the male and the male has good genes, then she is successful, and wins an amount W (i.e., the value of offspring with good genes); and if the female selects the male and the male does not have good genes, then she is unsuccessful, and loses an amount L (i.e., the value of offspring without good genes). We presume only that $W, L > 0$, so that $W > 0 > -L$. If the female does not select the male, then she neither wins nor loses. Presume that the female has probabilities over the states of nature, and also presume that she has reliability probabilities over the signals sent by the male. As we shall see, the specific values of W and L play no role in the analysis of the good genes view.

The female's choice problem can be characterized by the table on the following page:

Payoff Table	a1 mate	a2 not mate	P(s)
s1 good genes	W	0	p
s2 ¬good genes	-L	0	1-p

Note that this decision problem is an instance of the Becker-DeGroot-Marschak (1964) reference lottery. In Marschak's (1974) formulation of Hintikka's interrogative logic, the female's question *is* the information system represented by the reliability probabilities P(z/s), i.e., by the system

Information System	P(z z1 beautiful	/s) z2 ugly
s1 good genes	b	1-b
s2 ¬good genes	1-u	u

and the male's answer is represented by the signal he sends to the female, i.e., by either z1 or z2 (Marschak 1974, especially p. 152). Note that we have adopted a particular form of inquiry by specifying the female's question as an information system and the male's answer as a signal from that system. Of course, other forms of inquiry could be specified (Kolak 2001, especially pp. 55-58). However, the form of inquiry adopted here captures "merely probable answers." Interestingly, this form of inquiry has other applications within the context of Hintikka's interrogative logic (e.g., Dacey 1981b).

As per traditional decision theory, we presume the female behaves as if she has posterior probabilities, generated via Bayes' rule, as presented in the following extended table:

pragmatic decision problem				information system		posterior probabilities	
			P(s)	P(z	/s)	P(s	/z)
	a1	a2		z1	z2	z1	z2
	mate	not mate		beautiful	ugly	beautiful	ugly
s1 good genes	W	0	p	b	1-b	$\frac{pb}{pb+(1-p)(1-u)}$	$\frac{p(1-b)}{p(1-b)+(1-p)u}$
s2 ¬good genes	-L	0	1-p	1-u	u	$\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}$	$\frac{(1-p)u}{p(1-b)+(1-p)u}$

Informed mate selection occurs if and only if the female selects a male who sends the signal 'beautiful' and rejects a male who sends the signal 'ugly'. Under Pascal's decision rule, these choices are made via maximizing expected value. In decision-theoretic terms, informed mate selection occurs if and only if the Bayes strategy is <mate, not mate>, i.e., if and only if

$$\text{E[V(a1)/z1]} = \frac{pb}{pb+(1-p)(1-u)}W + \frac{(1-p)(1-u)}{pb+(1-p)(1-u)}(-L) > 0 = \text{E[V(a2)/z1]}$$

And

$$\text{E[V(a1)/z2]} = \frac{p(1-b)}{p(1-b)+(1-p)u}W + \frac{(1-p)u}{p(1-b)+(1-p)u}(-L) < 0 = \text{E[V(a2)/z2]}.$$

The foregoing conditions reduce to

$$\frac{p}{1-p} > \frac{(1-u)L}{bW}$$

and

$$\frac{p}{1-p} < \frac{uL}{(1-b)W},$$

respectively.

Together, these conditions yield

$$\frac{(1-u)L}{bW} < \frac{uL}{(1-b)W}.$$

Canceling W and L and rearranging terms yields b+u > 1. This inequality makes the good-genes view testable. If females select beautiful males and reject ugly males, then if b+u is not greater than unity, then the good-genes view is incorrect. Note that the condition b+u > 1 does not involve p, W, or L. Also note that the reliability probabilities b and u are objectively determinable, by a human researcher, as relative frequencies.

Recall that b = P(beautiful/good genes) and u = P(ugly/not good genes). These probabilities, called reliability probabilities, make a direct connection between the male's ornamentation (beautiful or ugly) and the state of the male's genes (good, not good). This connection is what makes b+u > 1 a tool for testing the good genes view associated with Wallace, and not a tool for testing the sexy-son view advanced with Fisher.

An information system is said to be informative if and only if b+u > 1 (Marschak 1971). Note that if b > .5 and u > .5, as we would expect for any plausible information system, then b+u > 1. However, we have the curious result that an information system can be informative even if either b < .5 or u < .5 (but not both), as long as the larger reliability probability is sufficiently large so that b+u > 1.

Now consider the von Neumann-Morgenstern decision rule. Under this rule, the female has a utility function U, defined over the payoffs, that captures her attitude toward risk. Without loss of generality, let U(0) = 0, so that U(G) > 0 > U(-L). The Bayes strategy is <mate, not mate> if and only if

$$\text{E[U(a1)/z1]} = \frac{pb}{pb+(1-p)(1-u)}U(W) + \frac{(1-p)(1-u)}{pb+(1-p)(1-u)}U(-L) > 0 = \text{E[U(a2)/z1]}$$

And

$$\text{E[U(a1)/z2]} = \frac{p(1-b)}{p(1-b)+(1-p)u}U(W) + \frac{(1-p)u}{p(1-b)+(1-p)u}U(-L) < 0 = \text{E[U(a2)/z2]}.$$

The foregoing conditions reduce to

$$\frac{p}{1-p} > \frac{-(1-u)U(-L)}{bU(W)}$$

and

$$\frac{p}{1-p} < \frac{-uU(-L)}{(1-b)U(W)},$$

respectively.

Together, these conditions yield

$$\frac{-(1-u)U(-L)}{bU(W)} < \frac{-uU(-L)}{(1-b)U(W)}.$$

Canceling U(W) and –U(-L) and rearranging terms again yields b+u > 1.

Finally, consider the Kahneman-Tversky decision rule. Under this rule, the female has a valuation function, v, defined over the payoffs, that is concave over gains, convex over losses, and more steeply sloped over losses than over gains. Without loss of generality, let v(0) = 0, so that v(G) > 0 > v(-L). Also under this rule, the female has a probability weighting function, w, that is reverse S-shaped so that w(p) > p for low values of p and w(p) < p for medium and high values of p. (Empirically, w(p) = p when p is approximately 1/3, and, of course, when p = 0, 1.) Since w(p) is sub-additive, i.e., w(p) + w(1-p) < 1, the functions based on w are not expected values. We will use a lower case e to denote the relevant functions.

Under the Kahneman-Tversky decision rule, the Bayes strategy is <mate, not mate> if and only if

$$\text{e[v(a1)/z1]} = w\left(\frac{pb}{pb+(1-p)(1-u)}\right)v(W) + w\left(\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}\right)v(-L) > 0 = \text{e[v(a2)/z1]}$$

and

$$\text{e[v(a1)/z2]} = w\left(\frac{p(1-b)}{p(1-b)+(1-p)u}\right)v(W) + w\left(\frac{(1-p)u}{p(1-b)+(1-p)u}\right)v(-L) < 0 = \text{e[v(a2)/z2]}$$

The simplified form of the Kahneman-Tversky probability weighting function is

$$w(p) = \frac{p^{\gamma}}{(p^{\gamma} + (1-p)^{\gamma})^{1/\gamma}},$$

where 0< γ < 1. Employing the simplified form of this probability weighting function yields :

$$\left(\frac{\left(\frac{pb}{pb+(1-p)(1-u)}\right)^{\gamma}}{\left[\left(\frac{pb}{pb+(1-p)(1-u)}\right)^{\gamma}+\left(\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}\right)^{\gamma}\right]^{1/\gamma}}\right)v(W)$$

$$+$$

$$\left(\frac{\left(\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}\right)^{\gamma}}{\left[\left(\frac{pb}{pb+(1-p)(1-u)}\right)^{\gamma}+\left(\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}\right)^{\gamma}\right]^{1/\gamma}}\right)v(-L) > 0$$

and

$$\left(\frac{\left(\frac{p(1-b)}{p(1-b)+(1-p)u}\right)^{\gamma}}{\left[\left(\frac{p(1-b)}{p(1-b)+(1-p)u}\right)^{\gamma}+\left(\frac{(1-p)u}{p(1-b)+(1-p)u}\right)^{\gamma}\right]^{1/\gamma}}\right)v(W)$$

$$+$$

$$\left(\frac{\left(\frac{(1-p)u}{pb+(1-p)(1-u)}\right)^{\gamma}}{\left[\left(\frac{p(1-b)}{p(1-b)+(1-p)u}\right)^{\gamma}+\left(\frac{(1-p)u}{p(1-b)+(1-p)u}\right)^{\gamma}\right]^{1/\gamma}}\right)v(-L) < 0$$

respectively. Eliminating common denominators yields

$$\left(\left(\frac{pb}{pb+(1-p)(1-u)}\right)^{\gamma}\right)v(W)+\left(\left(\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}\right)^{\gamma}\right)v(-L) > 0$$

and

$$\left(\left(\frac{p(1-b)}{p(1-b)+(1-p)u}\right)^{\gamma}\right)v(W)+\left(\left(\frac{(1-p)u}{pb+(1-p)(1-u)}\right)^{\gamma}\right)v(-L) < 0,$$

respectively. Eliminating common denominators once again yields

$$\left((pb)^{\gamma}\right)v(W)+\left(((1-p)(1-u))^{\gamma}\right)v(-L) > 0$$

and

$$\left((p(1-b))^{\gamma}\right)v(W)+\left(((1-p)u)^{\gamma}\right)v(-L) < 0,$$

respectively. These expand to

$$\left(p^{\gamma}b^{\gamma}\right)v(W)+\left((1-p)^{\gamma}(1-u)^{\gamma}\right)v(-L) > 0$$

and

$$\left(p^{\gamma}(1-b)^{\gamma}\right)v(W)+\left((1-p)^{\gamma}u^{\gamma}\right)v(-L) < 0,$$

respectively. Rearranging and factoring yields

$$\frac{p^{\gamma}}{(1-p)^{\gamma}} > \frac{-(1-u)^{\gamma}v(-L)}{b^{\gamma}v(W)}$$

and

$$\frac{p^{\gamma}}{(1-p)^{\gamma}} < \frac{-u^{\gamma}v(-L)}{(1-b)^{\gamma}v(W)},$$

respectively. Thus, we have

$$\frac{-(1-u)^{\gamma}v(-L)}{b^{\gamma}v(W)} < \frac{-u^{\gamma}v(-L)}{(1-b)^{\gamma}v(W)}.$$

Eliminating v(W) and –v(-L) yields

$$\frac{(1-u)^{\gamma}}{b^{\gamma}} < \frac{u^{\gamma}}{(1-b)^{\gamma}}.$$

Taking the γ-root of both sides yields

$$\frac{(1-u)}{b} < \frac{u}{(1-b)},$$

so that we once again have the condition

b+u > 1.

Thus, the tetsable result holds for the decision rules of Pascal, von Neumann-Morganstern, and Kahneman-Tversky when applied to the pragmatic decision problem. That is, if the female responds to the information system with the Bayes strategy <a1, a2>, then b+u > 1. This result is quite robust in that it holds for all risk attitudes, including the risk neutrality of the Pascal model, the simple risk aversion or risk seeking of the von Neumann-Morgenstern model, and the hybrid risk attitude of the Kahneman-Tversky model. Interestingly, this result does <u>not</u> hold for the general specification of the Kahneman-Tversky probability weighting function, nor does it hold for the Prelec (1998) probability weighting function. Thus, we know just how robust the basic result proves to be.

4. THE INDUCTIVE DECISION PROBLEM

If the female faces the inductive decision problem, then she is interested in forming either the belief that the male at hand has or does not have good genes, or she is interested in believing the inductive generalization that all males with beautiful tails have good genes. The former decision problem is modeled via the pure truth-seeker (Marschak 1974, pp. 170-173), and the latter via maximizing expected content (Hintikka 1968/1999, pp. 219-220). The pure truth-seeker's payoff table is

pure truth-seeker's payoff table		
	a1 Bel[good]	a2 Bel[¬good]
s1 good genes	1	0
s2 ¬good genes	0	1

and the expected content maximizer's payoff table is

content payoff table		
	a1 Bel[good]	a2 Bel[¬good]
s1 good genes	cont(s1)	0
s2 ¬good genes	-cont(¬s1)	0

where cont(h) = 1-p(h). The payoff table is based on the view that if h is true, then the gain is the information content, cont(h), of h, whereas if h is false, then the loss is the information content of the negation of h. (See Hintikka 1968/1999, p. 219.)

The pure truth-seeker's decision problem is illustrated on the following page:

pure truth-seeker's decision problem				information system		posterior probabilities	
			P(s)	P(z/s)		P(s/z)	
	a1 Bel[good]	a2 Bel[¬good]		z1 beautiful	z2 ugly	z1 beautiful	z2 ugly
s1 good genes	1	0	p	b	1-b	$\frac{pb}{pb+(1-p)(1-u)}$	$\frac{p(1-b)}{p(1-b)+(1-p)u}$
s2 ¬good genes	0	1	1-p	1-u	u	$\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}$	$\frac{(1-p)u}{p(1-b)+(1-p)u}$

Informed mate selection occurs if and only if the female believes a male who sends the signal 'beautiful' has good genes and believes a male who sends the signal 'ugly' does not have good genes. In decision-theoretic terms, informed mate selection occurs if and only if the Bayes strategy is <Bel[good], Bel[¬good]>, i.e., if and only if

$$\mathrm{E[a1/z1]} = \frac{pb}{pb+(1-p)(1-u)}(1) + \frac{(1-p)(1-u)}{pb+(1-p)(1-u)}(0) >$$

$$\frac{pb}{pb+(1-p)(1-u)}(0) + \frac{(1-p)(1-u)}{pb+(1-p)(1-u)}(1) = \mathrm{E[a2/z1]}$$

and

$$\mathrm{E[a1/z2]} = \frac{p(1-b)}{p(1-b)+(1-p)u}(1) + \frac{(1-p)u}{p(1-b)+(1-p)u}(0) <$$

$$\frac{p(1-b)}{p(1-b)+(1-p)u}(0)+\frac{(1-p)u}{p(1-b)+(1-p)u}(1) = \text{E[a2/z2]}.$$

The foregoing conditions reduce to

$$\frac{\text{p}}{1-\text{p}} > \frac{(1-\text{u})}{\text{b}}$$

and

$$\frac{\text{p}}{1-\text{p}} < \frac{\text{u}}{(1-\text{b})},$$

respectively.

Together, these conditions yield

$$\frac{(1-u)}{b} < \frac{u}{(1-b)},$$

so that once again we have b+u > 1.

The expected content maximizer's decision problem is as follows:

cont maximizer's decision problem				information system		posterior probabilities	
			P(s)	P(z	/s)	P(s	/z)
	a1	a2		z1	z2	z1	z2
	Bel[gen]	¬Bel[gen]		beautiful	ugly	beautiful	ugly
s1 gen	cont(s1)	0	p	b	1-b	$\frac{pb}{pb+(1-p)(1-u)}$	$\frac{p(1-b)}{p(1-b)+(1-p)u}$
s2 ¬gen	- cont(¬s1)	0	1-p	1-u	u	$\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}$	$\frac{(1-p)u}{p(1-b)+(1-p)u}$

where 'gen' is an abbreviation for the generalization 'all beautiful males have good genes'.

The Bayes strategy for the expected content maximizer is <a1,a2> if and only if

$$\text{E[a1/z1]} = \frac{pb}{pb+(1-p)(1-u)}(1-p)+\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}(-p) >$$

$$\frac{pb}{pb+(1-p)(1-u)}(0)+\frac{(1-p)(1-u)}{pb+(1-p)(1-u)}(0) = \text{E[a2/z1]}$$

and

$$\text{E[a1/z2]} = \frac{p(1-b)}{p(1-b)+(1-p)u}(1-p)+\frac{(1-p)u}{p(1-b)+(1-p)u}(-p) <$$

$$\frac{p(1-b)}{p(1-b)+(1-p)u}(0)+\frac{(1-p)u}{p(1-b)+(1-p)u}(0) = \text{E[a2/z2]}.$$

These equations reduce to

$$pb(1-p)+(1-p)(1-u)(-p) > 0$$

and

$$p(1-b)(1-p)+(1-p)u(-p) < 0,$$

respectively. Rearranging terms in each equation yields b+u >1, and the good genes view is again testable.

5. CONCLUSION

Marschak (1974) established that the economic theory of information provides a model of Hintikka's interrogative logic that can be applied in the case of uncertain answers and employed to resolve both pragmatic and inductive decision problems. So applied, Hintikka's logic provides a most interesting result vis-à-vis the good genes view of male ornamentation. Simply put, if females select beautiful males and reject ugly males, then male ornamentation, as an information system, is informative, and the good genes view is testable. Finally, as the foregoing analysis yields, the good genes view is testable whether the female resolves the pragmatic decision problem, via any of the three major decision rules, or the inductive decision problem, via either of the two major decision rules. This suggests that Hintikka's interrogative logic, or "inquiry as inquiry" as he prefers, is remarkably robust.

REFERENCES

Arnauld, A. and P. Nicole. 1662, *La Logique, ou l'Art de Penser*, Paris. Commonly called The Port Royal Logic; available in English as *The Art of Thinking*, 1964, Indianapolis: Bobbs-Merrill, Library of Liberal Arts, and as *Logic or the Art of Thinking*, translated and edited by Jill Vance Buroker, 1996, Cambridge: Cambridge University Press.

Becker, Gordon M., Morris DeGroot, and Jacob Marschak. 1964, "Measuring Utility By a Single-Response Sequential Method," *Behavioral Science* 9:226-232.

Dacey, Raymond. 1985, "A Logic of Detection and Deception," pp. 205-227 in J. Hintikka and F. Vandamme (eds.), *Logic of Discovery and Logic of Discourse*, New York: Plenum Press.

Dacey, Raymond. 1981a, "An Interrogative Account of the Dialectical Inquiring System Based Upon the Economic Theory of Information," *Synthese* 47:43-55.

Dacey, Raymond. 1981b, "Some Implications of 'Theory Absorption' for Economic Theory and the Economics of Information," pp. 111-136 in J. C. Pitt (ed.), *Philosophy in Economics*, Dordrecht, Holland: D. Reidel Publishing Co.

Dacey, Raymond. 1979a, "The Dialectical Inquirer and Decisions Under Uncertainty: With Application to Science Policy in the Peoples' Republic of China," pp. 193-207 in W. Callebaut, M. DeMey, R. Pinxten, and F. Vandamme (eds.), *Theory of Knowledge and Science Policy*, Ghent, Belgium: University of Ghent.

Dacey, Raymond. 1979b, "The Role of Ambiguity in Manipulating Voter Behavior," *Theory and Decision,* 10: 265-279.

Hintikka, Jaakko (ed.). 1999, *Inquiry as Inquiry: A Logic of Scientific Discovery*, Dordrecht: Kluwer. [For the reader's convenience, page references to papers included in this volume are based on the pagination of this volume. Such papers are denoted by combined dates, e.g. Hintikka 1988/1999.]

Hintikka, Jaakko. 1993, "The Concept of Induction in the Light of the Interrogative Approach to Inquiry," pp. 23-43 in John Earman (ed.), *Inference, Explanation, and Other Frustrations: Essays in the Philosophy of Science*, Berkeley: University of California Press.

Hintikka, Jaakko. 1988/1999, "What is the Logic of Experimental Inquiry?," *Synthese* 74: 173-190, and reprinted as pp. 143-160 in J. Hintikka (1999).

Hintikka, Jaakko. 1985, "A Spectrum Logic of Questioning," *Philosophica* 35: 135-150, and reprinted as pp. 127-142 in J. Hintikka (1999).

Hintikka, Jaakko. 1976, *The Semantics of Questions and the Questions of Semantics: Case Studies in the Interrelations of Logic*, Amsterdam: North-Holland.

Hintikka, Jaakko. 1975, "Answers to Questions," Chapter 7 in J. Hintikka, *The Intentions of Intentionality and Other New Models for Modalities*, Dordrecht: D. Reidel.

Hintikka, Jaakko. 1974, "Questions About Questions," pp. 103-158 in M. Munitz and P. Unger (eds.), *Semantics and Philosophy*, New York: New York University Press.

Hintikka, Jaakko. 1973, *Logic, Language-Games, and Information*, Oxford: Clarendon Press.

Hintikka, Jaakko. 1968/1999, "The Varieties of Information in Scientific Explanation," pp. 311-331 in B. van Rootselaar and J. F. Staal Eds.), *Logic, Methodology, and Philosophy of Science III*, Amsterdam: North-Holland (1968), and reprinted as pp. 205-225 in J. Hintikka (1999).

Hintikka, Jaakko. Ilpo Halonen, and Arto Mutanen. Forthcoming/1999, "Interrogative Logic as a General Theory of Reasoning," in R. Johnson and J. Woods (eds.), *Handbook of Applied Logic*, Dordrecht: Kluwer Academic Publishers and reprinted as pp. 47-90 in J. Hintikka (1999).

Hintikka, Jaakko and Merrill B. Hintikka. 1989, *The Logic of Epistemology and the Epistemology of Logic*, Dordrecht: Kluwer Academic Publishers.

Hintikka, Jaakko and Jack Kaulas. 1983, *The Game of Language: Studies in Game-Theoretical Semantics*, Dordrecht: D. Reidel Publishing Co.

Kahneman, Daniel and Amos Tversky. 2001, *Choices, values, and frames*, Cambridge, UK: Cambridge University Press.

Kahneman, Daniel and Amos Tversky. 1979, "Prospect Theory: An Analysis of Decisions Under Risk," *Econometrica* 47: 263-291.

Kolak, Daniel. 2001, *On Hintikka*, Belmont, CA: Wadsworth.

Marschak, Jacob. 1974, "Information, Decision, and the Scientist," pp. 145-178 in C. Cherry (ed.), *Pragmatic Aspects of Human Communication*, Dordrecht, Holland: D. Reidel Publishing Co.

Marschak, Jacob. 1971, “Economics of Information Systems,” *Journal of the American Statistical Association* 66: 192-219.

Prelec, Drazen. 1998, “The Probability Weighting Function,” *Econometrica* 66: 497-527.

Ridley, Mark. 1993, *Evolution*, Boston: Blackwell Scientific Publications.

Ridley, Matt. 1993, *The Red Queen: Sex and the Evolution of Human Nature*, New York: Penguin Books.

Tversky, A., and Kahneman, D. 1992, “Advances in Prospect Theory: Cumulative Representation of Uncertainty,” *Journal of Risk and Uncertainty* 5: 297-323.

von Neumann, John and Oskar Morgenstern, 1947, *Theory of Games and Economic Behavior*, Princeton: Princeton University Press.

A METALOGICAL CRITIQUE OF WITTGENSTEINIAN ‘PHENOMENOLOGY’

William Boos

1. TRACTARIAN MONISM

My principal aim in this brief essay will be to inflect (or coopt) Jaakko and Merrill Hintikka’s ‘phenomenological’ interpretation of Ludwig Wittgenstein’s metaphysical aims (cf. Hintikka and Hintikka 1986 and Hintikka 1996) in directions which may

1 provide tenable if heterodox metalogical interpretations of Wittgensteinian *Sprachspiele*; and

2 partially clarify Wittgenstein’s well-known unremitting hostility to *Mengenlehre*, “*Logik*(*en*) *zweiter Ordnung*” and “*Metalogik*(*en*)” in all their allegedly nefarious forms. (cf., e.g., *PG*, 46*, PB*, 211 *N*114,2 and *N*211, 242)

Along the way, I will also

3 interpret the *Logisch-Philosophische Abhandlung* as an attempt to sketch (but not ‘define’) a ‘*universal*’ *satisfaction-relation* (for ‘*alles, was der Fall ist* [*oder sein könnte*]’), in an *inkonsistenten Inbegriff* one might call *Wittgensteins Paradies*;

4 construe the *LPA*’s ‘*mysticism*’ and ‘*solipisism*’ as a natural concomitant of this relation’s Cantorian *Inkonsistenz*, *self-referential ambiguity* and susceptibility to forms of *semantic paradox* familiar to Wittgenstein as well as his sometime mentor.

If these preliminary assimilations are tenable, they may also offer a straightforward rationale for Wittgenstein’s evident anxiety about the work, and some of his agonistic adjurations of himself and others. Consider, for example, that he was struggling in effect

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 75-99.

5 to provide a discursive sketch in ordinary German of a quasi-Tarskian satisfaction-relation, twenty years *avant l'heure* (cf. also Hintikka 1996, 27-28);

6 to do so without recourse to metatheoretic contexts for the many hypostatic claims he would in the process have to make; and

7 to 'solve' problems of semantic paradox (Berry's as well as Russell's) by recourse to 'ineffable' ascent up Sextus Empiricus' ladder into "*das Mystische*".

The first of these aspirations--which would have been an innovation of the first magnitude--surpassed Wittgenstein's technical abilities. Lest this seem to be *lèse-majesté*, consider that Wittenstein might have included a sketch of such a satisfaction-relation as part of the exposition of his *Bildtheorie*, but contented himself with two pages of graph-theoretic representations of propositional truth-tables.

The *horror metatheoriae* I have attributed to Wittgenstein in 6 above was clearly influenced by his distaste for Russellian type-hierarchies. But it was also forced by his 'phenomenological' aspirations, and his assertion of the shadow-satisfaction-relation's apodeictic '*universality*'.

Whatever his conceptual, temperamental or philosophical motivations may have been, I will argue in the sequel

8 that this commitment eventually became a lifelong imperative for him, and

9 that it had debilitating consequences for his subsequent language-game-theory.

The third aim or "*Ansatz*"--to isolate problems of semantic paradox at the margins of an otherwise 'all'-encompassing 'world'--also accorded with deep currents in Wittgenstein's philosophical and esthetic temperament, however warranted criticisms of it as a copout, *petitio*, *Ausflucht* or *échappatoire*. A certain sensitivity to problems that might arise if one desacralised 'the' ineffable penumbra of 'the' world may also have heightened Wittgenstein's "*impatience*" with questions about such aims (cf. Hintikka 1996, 1-2 and Hintikka 2000, 5).

A natural historical and methodological question arises, assuming I have accurately characterised Wittgenstein's metaphysical 'aims' in 3 through 7 above. Was the solution he achieved *stable*? For assorted reasons unrelated to color-charts or Piero Sraffa's sardonic executions of scurrilous gestures, I believe the answer to this question is 'no'. One of these reasons may related to the Hintikkan notion of 'phenomenology' alluded to in the first paragraph. As Jaakko and Merrill Hintikka distinguished '*phenomenology*' from '*phenomenalism*' (cf., e.g., Hintikka 1996, 210), the former refers to *any* presentations that are '*immediate*', in the metaphysical sense that they require no '*hypotheses*' to ground them, or secure them, or serve as (ancient Greek, twentieth-century metalogical or other) '*axiomata*' for them.

The distinction is intrinsically metaphysical, and as such might serve to clarify just how vaultingly ambitious Wittgenstein's metaphysical aims in 3-7 above (and, I will argue, certain subsequent 'game'-theoretic counterparts).

For such aspirations to '*hypothesis-freedom*' are ancient and venerable. Slight linguistic or methodological variants of them, for example, characterised (e.g.)

10 middle- and late-Platonic '*forms*' (cf., e.g., *Republic*, 531e and 534b, *Sophist*, 238c and 259e, *Parmenides* 133d, 134b-c and 135a);

11 Aristotelian '*first causes*' (cf.1071b3-1073b1), '*first philosophy*' (cf.982a4-983a20 and 1003a20-1008a5) and '*theoretical*' '*thought about thought*' (cf. 1177b1, 1178b7-9, and 1178b22-25);

12 Aquinian and Cartesian '*god*(*s*)' (cf. Aquinas ST 1 q.2 a.3; cf. Descartes, ATVII 40-43, 45-51 and 66-69);

13 Spinozan '*substantia*(*e*)' (cf. Spinoza, 86-159, G46-71);

14 Kantian '*Grenzbegriffe*' (cf. KdrV B310-11), '*Vollständigkeit*(*en*) *der Bedingungen*' (KdrV B443-448, B524, B542-551, B692-693), and the "*reine Selbsttätigkeit*" Kant called "*Freiheit*" (*Grundlegung* 452);

15 Husserlian (*Gebiete der*) '*Voraussetzungslosigkeit*' (cf. Husserl 1913, 19-22), and '*vorurteilslose Seinsmodi*' (cf., Husserl 1977, 37).

To me at least, there seems little doubt that the *LPA*'s grand-syntactical '*Logik*' and equally grand-semantic correlate '*die Welt*' were indeed 'phenomenological' *Inbegriffe* in such 'hypostatic' senses. Historical counterparts of such 'phenomenological' entities have traditionally encountered problems one might draw together under a kind of '*complementarity principle*': the more '*comprehensive*' these entities' 'intended' '*universality*', the less *comprehensible* their '*hypothesis-freedom*' becomes.

More precisely, appeals to a metatheoretic venue in which one could characterise a given class of '*hypotheses*', and assert the 'phenomenological' entities' putative '*freedom*' from them (not to mention their '*existence*') would seem to compromise or relativise their ostensible 'universality' (cf. "*Deuten wir, so machen wir Hypothesen*", *PU* II, 524).

Ancient skeptical as well as twentieth-century metalogical analyses also suggest that one cannot interpret such a metatheory in its own 'phenomenology' without vulnerability to some sort of '*circularity*', *semantic paradox* or susceptibility to Gödelian *diagonalisation* (a well-studied early modern example of such an *aporia* might be found in the '*Cartesian circle*', first observed by Arnauld in 1641; cf. ATVII, 214) In the case of the *LPA*, such problems are more than usually apparent (and well-known). *Either:*

16 '*the*' class of '*hypotheses*' from which Wittgenstein's '*Welt*' is '*free*' is *expressible within* '*it*'.

Or

17 'its' '*existence*'--like that of Anselm's 'god'--is a matter of *metatheoretic* (and therefore '*nonsensical*') *postulation*, in contravention of Wittgenstein's injunction to silence.. (cf. Wittgenstein's rather uncharacteristic remark that "[*d*]*as Überraschende, Paradoxe, ist paradox nur in einer gewissen gleichsam mangelhaften Umgebung. Man muß diese Umgebung so ergänzen, daß, was paradox schien, nicht länger so erscheint*", BGM, 410, VII, 43)

In either case, we would be confronted with semiformal counterparts of the semantic paradoxes of Berry and Russell Wittgenstein claimed to 'resolve' (*oder wenigstens* '*aus der Welt schaffen*'). But all known provisional solutions of the *latter* were known to generate syntactic and semantic hierarchies of the sort he anathematised.

Since Wittgenstein's deeper (and more philosophically defensible) aim was *not* to formulate theological avowals (however 'ethically' regulated his speculations may have been), some '*desacralisation*' of the *Flucht nach vorne* I have just sketched mentioned above may therefore have been inevitable. His readers and his own philosophical conscience, so to speak, enjoined him to provide a '*Sinn*' for his earlier work's 'ethical' '*Unsinngloriole*'. Some of the tensions this 'injunction' occasioned may also have been intensified the apparently widespread expectation, among Wittgenstein's admirers in Cambridge and other readers of the *LPA*, that he would clarify methodological aspects of its conspicuous *Aporien* with the aid of new discoveries in '*logic*'.

Relatively soon after Wittgenstein returned to Cambridge in 1929, for example, it may have become clear to a number of investigators--among them, I suspect, Wittgenstein's clear-headed friend Frank Ramsey--that further clarifications of "logic" might *not* follow the lines of magisterial force Wittgenstein may have thought he had imposed on it.

For between 1922 and 1934,

18 logicians began to realign their representations of Wittgenstein's *Welt* in ways which provisionally but systematically demarcated theories from metatheories; and--

19 mathematical and experimental physicists began provisionally but systematically to distinguish 'observations' of physical 'systems' from the 'states' they measured.

Both these developments called the hypostatic *monism* of Wittgenstein's tractarian ontology into serious question. The logicians Thoralf Skolem and Kurt Gödel, for example, had proved theorems which suggested that

20 'the' integers',

21 'the' real numbers,

22 'the' recursively axiomatisable formal theories, and

23 'the' worlds these axiomatisable theories might 'depict'. might be no more than '*metatheoretically*' '*intended*' notions, metatheoretically interpretable in '*potential*' as well as underdetermined ways, and in vast--in fact continuous--*logischen Räumen* (cf., for comparison, *LPA* 3.4).

Wittgenstein tried for a time to keep up with such developments. After Gödel's results became known, for example, he wrote out attempt after unsuccessful attempt to deconstruct or discredit or explain away various forms of *diagonal arguments* which lay at the heart of Gödel's argument (cf. "*Meine Aufgabe ist es nicht, über den Gödelschen Beweis . . . zu reden;*

sondern an ihm vorbei zu reden", *BGM*, 383,II, 19), and modelled most of these abortive arguments on early jottings in which he had tried to 'prove' that semantic paradoxes were ill-formed (cf., e.g., *LPA*, 3.333).

When these 'refutations' failed, he sought other arguments to justify his deep conviction that the new results 'had to' be wrong, or at least irrelevant (cf. *N*117, 147; *N*117, 152; *N*121, 81v, 82v and 84r; *N*124, 115; and *N*163, 41v). More or less by way of accompaniment, he also drafted from time to time exasperated remarks of the sort cited earlier, to the effect that set-theoretic frameworks for such results were *vom Übel*, and even *des Teufels*.

Why? *Why* did Wittgenstein execrate these new metalogical discoveries *so* fiercely?

One might propose personal as well as philosophical rationales of this hostility (cf., for example, Hintikka 1996, 148-151 and 159-162, and Hintikka 2000, 58). The one I will defend here is that Wittgenstein sought to '*save*--or at least *hold harmless*, at the margins of his nascent *phenomenology* (or *semeiology*) *of* '*language-games*'--an elusive counterpart of the *horror metatheoriae* and *semantic monism* he had defended in the *LPA*.

Somewhat more precisely, the explicit analogy expressed in his remark that "[*w*]*ie es keine Metaphysik gibt, so gibt es keine Metalogik*" (*N*211, 242) seems to me to suggest he hoped to "*save the* [*linguistic*] '*phenomena*'" he had '*uniquely*' interpreted in the *LPA*--even as he disassembled the 'mystical' scaffolding in which he had once enveloped 'them'.

2. 'SPRACHSPIELE' AS 'THEORIEN'

In this section, I will offer a simple but somewhat unorthodox assimilation of Wittgenstein's inconclusive efforts to develop a semeiotic 'phenomenology' of 'language-games' in the *LPA* to well-studied semeiological aspects of the evolving *Metalogik* he so despised.

More precisely, I will

1 assimilate of ('families') of Wittgensteinian *Sprachspiele* to (particular collections of) *first-order theories*, partially (pre)ordered by syntactic interpretation; and

2 argue that significant aspects of his 'game'-theoretic semeiology in the *Untersuchungen* reflected forms of *metalogical theory-relativism*.which naturally arise in the study of such collections.

In (quasi-)'Wittgensteinian' terms, one might formulate the argument as follows:

3 that a 'family resemblance' exists between certain forms of the *Metalogik* Wittgenstein so hated and his *Sprachspieltheorie*; indeed,

4 that several problems, preoccupations and thought-experiments of his 'later' philosophy have archetypes in a conceptual framework (or 'language game') he contemned.

By way of preparation for a partial justification of this assimilation (which differs from more complex generalised-quantifier-interpretations of "language-game" Hintikka and Sandhu have developed; cf., e.g., Hintikka 1996, 162-177 and Hintikka 1998, 1-83), let me return for a moment to 'phenomenological' reconstructions of Wittgenstein's aims and philosophical temperament.

By definition, first, '*phenomenological*' entities would seem in some *prima facie* sense to *relational* in the sense that they appear (*phainontai) to something or another*, *in some fashion or another*. Such '*appearance*(*s*)' might not be '*physical*' or '*phenomenal*' in any terms I (for example) am familiar with. But they *would* seem to compromise the '*hypothesis-freedom*' Hintikka has postulated (cf. 1.16 and the remarks preceding it above).

Leaving such tensions in suspension. and suspending judgment for the moment about the nature of phenomenological '*appearance*', let me

5 call certain formal counterparts of 'language games' *theories*, and assume without much loss of generality (cf. 7-11 below) that the formal counterparts of 'language games' are recursively axiomatisable, first-order theories.

Furthermore, let me

6 call counterparts of certain 'games' which are somewhat 'richer' and more 'complex' that other 'games' (which '*appear*' to them), *metatheories* with respect to (the counterparts of) such 'games'.

What, a philosophical logician might ask, justifies my apparent 'Quinean' stipulation that the theories I will consider be *first-order*?

A preliminary, 'phenomenological' as well as 'Wittgensteinian' answer might be that the 'language games' I want these theories to generalise do *not* (have to) have any *intrinsic ontological priority* over each other. In mathematical-logical jargon, I would first respond to this observation with appeals to now-standard interpretations of *types* as first-order *sorts*--a reconstruction Henkin introduced in print 1950, in a journal Wittgenstein very likely never read (cf. also Enderton, 281-289; Hintikka invoked this interpretation in another context in Hintikka 1996, 23).

A more considered response might invoke a heuristic variant of Ockham's razor.

For the first-order theories I have proposed as generalisations of Wittgensteinian 'language games' are the *simplest* languages of a fairly '*universal*' sort (in a sense to which I will return shortly), about which one can *stipulate*, in other, 'metatheoretic' first-order languages

7 that 'deductive' '*consequence relations*' 'hold' between certain of their inscripts (words and phrases);

8 that their (formal) 'grammars' obey certain rules of 'induction' and 'recursion' (more also about such 'rule-following' later); and

9 that applications of the 'consequence relations' do not lead to certain expressions or utterances call 'absurdities' (in which case we call the games or theories 'consistent').

It is known that the basic patterns sketched in 5-7 can be encoded or represented in myriads of ways, and I will draw on this plurality of syntactical interpretations to argue that first-order theories and their 'inference-rules' are not 'reductive' so much as *ontologically neutral* and *provisionally* '*universal*'. By this I mean

10 that certain first-order *metatheories* of the sort Wittgenstein especially despised, called '*set-theories*', provide *neutral venues for adjudication of semantic questions about every* '*higher-order*' *theory or* '*abstract logic*' which has so far been devised;

and

11 that *processes of transition to* 'richer' first-order metatheories--especially the *set theories* just mentioned--*are indefinitely iterable*: they are subject to no discernible bounds, limits or *Grenzen* anyone has yet been able to discern or anticipate.

Such theories, in short, serve as the *simplest natural metatheoretic venues for other theories which may* '*appear*' *to them* (the motivation for my allusion above to *Ockham's rasor*), in the sense that they can *pose* and *sometimes decide* '*semantic*' questions for them--questions about their '*consistency*'; or their '*existence*'; or their *Anwendung*; or their *Verwendung*; or . . . .

The *interpretations* such first-order metatheories provide--which I have just construed as forms of metalogical '*appearance*'-- also come in various gradations. Certain theories, for example, may be '*interpretable*' in other theories in the (relatively *weak*) sense that the latter have ways to '*understand*' 'what the former are *talking about*', though they may *not* be able to make '*semantic*' metatheoretic decisions of the sort just introduced.

The formal jargon for this (much studied) relation among first-order theories is '*syntactical interpretability*' (cf., e. g, Enderton, 154-163, or Shoenfield, 61-65). One might therefore follow Peirce in calling theories which interpret other theories in this less semantically 'conclusive' way '*interpretants*' (rather than '*metatheories*') for them. Such interpretants, may offer also useful formal counterparts of Wittgensteinian "*Übersicht*" and "*Sehen als*". But they do *not* necessarily confer '*existence*' (*consistency*) on what they '*interpret*' in this way.

Since I also believe such distinctions between stronger and weaker forms of '*interpretation*' track aspects of Kant's elusive demarcation of what is *konstitutiv* from what is *bloß regulativ* (cf., e.g., KdrV B536, B537, B544, B692, B694, B699 and B710-714), let me say that a *metatheory* for a given theory '*constitutes*' that theory (as opposed to an *interpretant*, which merely *interprets* it.

As Kant himself suggested in the *Dialektik* of the first *Kritik* (cf., e.g., B672-B674), venerable forms of Wittgenstein's "*Philosophenunsinn*" may be traceable to tendencies to efface or ignore this distinction. Be that as it may, let me stop and recapitulate.

I have sketched an assimilation of Wittgensteinian '*language games*' to *consistent* first-order theories of the sort just sketched. And I have outlined binary relations between such theories which I have called '*being an interpretant of*', and '*being a metatheory for*'.

How much of the *Problematik* Wittgenstein set forth in the *Untersuchungen* carries over into this formal (or as I have presented it, semiformal) context? Quite a bit.

Consider first an obvious sense in which *metalogical interpretations* of first-order theories (in the strong sense of *metatheoretically complete and quantificationally adequate extensions* of such theories) *formally relativised* as well as *pluralised* the *LPA*'s inchoate prototype of a 'universal' *satisfaction-relation* (its '*Welt*').

Comparable forms of *informal* (and warily admeasured) *relativisation* and *pluralisation* clearly appeared among the *desiderata* to be satisfied by more latitudinarian accounts of language-games and their multiple interpretations that began to apear in Wittgenstein's later writings (though he condemned usages of words such as "*Deutung*" as carriers for them, in accordance with his '*horror metatheoriae*' I attributed to him in section 1).

Consider next the question whether a putative '*game*' could identify '*all*' 'games' (and therefore '*define*', in that sense, *what a* '*game*' *is*). By Gödel's now-classical arguments, there could *be* no such game. For if there *were*, it would be a *consistent* '*game*' *that consituted itself*, and Gödel and Tarski showed in different but closely related ways that this cannot happen.

Thirdly, consider the following metalogical reconstruction of the '*rule-following*' *conundrum*. In a game that provides a metatheory for another game, one might ask which '*rules*' the second, '*object-theoretic*' game '*follows*'.

If we convened that a '*rule*' is given by a metatheoretic *formula* or *predicate* or *property* or *procedure* (as Wittgenstein pointedly declined to do), could we prescribe in the metatheory how far we could '*go on*' in accordance with a given such '*rule*', without '*making a mistake*'?

If we appeal to slightly more complex Gödelian arguments, the curiously 'Wittgensteinian' answer to this question will in general be '*no*.' For a simple application of the parametric diagonal lemma (cf., e.g., Smorynski, 827, Boolos, 49, or Boos 1998, 61) yields that *the metatheory cannot decide which* '*rules*' *it imposes on an interpretation of the object-theory will fail to respect the relevant axioms of* '*mathematical induction*' (A sketch of the proof appears in 6.1 below.)

These observations suggests that my avowedly non-Wittgensteinian assimilations of 'games' to consistent 'ontologically neutral' first-order theories may *not* yield results which are *antithetical* to the *Aporien* Wittgenstein studied. In some cases at least, they *confirm* them.

Returning to the remarks I made at the beginning of the section, let me now reconsider in this quasi-'later-Wittgensteinian' context the extent to which '*language-games*' of the sort I have just studied might form a '*phenomenology*', or '*phenomenological class*'.

By the foregoing remarks, each '*game*' with an induction-scheme adequate to '*engage in linguistic activity*' (parse and interpret its own language) *will* be able to '*define*' a class of theories *it* considers '*games*'. In this sense, it might be said to determine a kind of metalogical *Lebensform* (or Husserlian *Lebenswelt*).

What it *cannot* do is *decide* whether *it itself* (which it can syntactically 'encode', and 'talk about') is a '*game*'; or which '*rules*' have the property that '*games*' it identifies can *apply* them in such a way that they or the metatheory '*knows how to go on*'.

There are, moreover, no undogmatic reasons to '*stop*' the search for wider and wider metatheories, despite Wittgenstein's stern pronouncements (cf., e.g., *PU* 29, *ÜG* 563 and *ÜG* 576) that such recourses to '*secondary*' language-games (which play metatheoretic roles in Hintikka's phenomenological analysis) '*must halt*', and assertions '*we*' cannot engage '*forever*' in such *Philosophenunsinn*. (Indeed. As Keynes' observed, 'in the long run' we're all dead.)

Put more formally, metatheoretic 'games' might provide '*object-theories*' for other 'interpretants' and 'metatheories', which might or might not '*ground*' them (prove their *consistency*, which David Hilbert--anticipating the completeness theorem--construed as a distributive metatheoretic sort of '*existence*').

Such interpretants and metatheories, as I remarked earlier, may continue to do their thing(s), in indefinitely iterable ways. But the very 'Wittgensteinian' absence of any *scharfe Grenze* or *closure* or final '*fixed point*' of such iterations might also suggest a *broader* sort of '*phenomenology*', namely an indeterminately processive 'phenomenological' heirarchy of quasi-Leibnizian '*phaenomena bene fundata*'.

(Cf. "*Ex Hypothesi, quod nihil aliud existat, quam Monades, et quod eae modificentur varie et consentienter, fit ut omnia caetera Entia quae concipimus non sint nisi phaenomena bene fundata*", Leibniz II, 473.)

Such hierarchies would also recapitulate ancient, medieval and post-medieval *hierarchies* of the sort contemporary *skeptics* proffered to their more '*dogmatic*' opponents. It was no accident, for example, that Descartes thought he had to 'bound' or suppress such a metatheoretic hierarchy of '*formal*' and '*objective realities*', in the passages cited in 1.12 above, in order to 'prove' the 'existence' of his [self]-'constituting' '*god*'.

Whatever their skeptical implications, moreover, there seem to be no obvious metalogical reasons why '*dynamical trajectories*' of such '*games*' might not be '*observed*', '*interpreted*' and even '*constituted*' in other '*games*'.

If, for example, one further interpreted '*freedom from hypothesis*' to mean something like '*apparent epistemic sufficiency unto itself*', the class of theories recognisable as (or 'seen as') games in other games might well be adequate to sustain the '*appearance*' (or '*illusion*') of such '*freedom*'.

And if this were the case, once again, such 'appearances' might indeed merit the proud name of (*theory-relative*) '*phenomenonology*'.

And *this*, finally, suggests to me a concluding unscientific thought-experiment, which reposes postulations of '*apparent* epistemic sufficiency' in interrogative form.

If ‘*we*’ were in such a class, would there be any way ‘*we*’ could discern what ‘*we*’ encountered in it from ‘*what there is*’ (or ‘*might be*’)?

I think not, and will defend this view in what follows.

3. . . . MORE ‘GAMES’ THAN ARE DREAMT OF IN ‘OUR’ PHILOSOPHIES. . . .

In the last section, I

1 construed Wittgensteinian ‘*language-games*’ as *consistent first-order theories*, and

2 outlined interpretations for some ‘Wittgensteinian’ predicaments one might encounter in the study of such ‘games’.

What I did *not* do was attempt to argue that every first-order theory is a language-game, in any sense of ‘game’ Wittgenstein himself would have tolerated. For any such ‘reconstruction’ or ‘recuperation’ would be countermanded many times over in Wittgenstein’s *Nachlaß* and published writings.

He explicitly *denied*, for example, that consistently axiomatisable theories which have no counterparts in ‘public’ human speech (so far as ‘we-know) could be language-games, and insisted this was the case whether or not scientists might later conclude that physiological realisations of them function as precursors or prototypes or preconditions of such speech.

He did not deny, of course that scientists might come to such conclusions, or even they might in some sense be ‘correct’. He simply repudiated the view that such languages ‘could’ have any philosophical significance of the sort he attributed (or admitted ‘*we*’ ‘can’ or ‘should’ attribute) to ‘*language-games*’.

Suppose, however, one set his strictures aside once again, and extended usage of the word ‘*language*’ to *any* formal first-order language of the sort considered in the last section, and ‘*game*’ to any *consistent theory* in such a ‘language’? I will argue in this section that such a change might have interesting and significant implications for such languages’ ‘*phenomenology*’.

More precisely, I will argue that the *constraints* and *circumscriptions* Wittgenstein imposed on his *Sprachspiele* are *much* more ‘*reductive*’ than the metalogical assimilations I have introduced in the last section wish to consider further here.

Those constraints make it *very* difficult, for example, to *parse*, much less *credit*, *genuine changes in* '*linguistic*' *as well as semantic perspective*. Wittgenstein struggled to solve this latent problem--without noticeable success, in my view--in his last (and perhaps most readable) work *Über Gewißheit* (cf., e.g., *ÜG* 506-509, *ÜG* 512-513 and *ÜG* 577).

If rigorously enforced, such constraints might even render difficult any serious cognitive analysis of Wittgenstein's own '*Hasenente*'-sequence--one of the reasons, I believe, why it is one of the *PU*'s more charming but less conclusive lines of argument.

A fortiori, they would literally render unintelligible wider-ranging accounts of cognitive loss and recovery in works such as Oliver Sacks's *Man Who Mistook His Wife for a Hat*, a work which seems to me rich in philosophical resonances.

So *stultifying* in fact do Wittgenstein's recusals seem to me that they remind me of David Hume's unwitting rejection of *theory-change* of almost *any* sort in his otherwise reasonable rejection of 'miracles' (cf. *E* I, 114-116); or of the passerby's well-known remark in *ÜG* 467 that "[*d*]*ieser Mensch ist nicht verrückt. Wir philosophieren nur*. . . .").

How did this come to pass? I believe that its antecedents in Wittgenstein's work may be traceable to motives which appear in

3 his elusive but pointed remarks about '*solipsism*' in *LPA* 5.64; and

4 his unusual choice of the word "*intern*" to characterise 'the' *inexpressible* '*structures*' that individuate the tractarian *Welt* (4.122) (semantic relations are commonly thought to be '*external*' to the syntactic predicates they realise)

On my account I sketched above in section 1, Wittgenstein's careful consideration of '*solipsism*' in the *LPA* was an early concomitant or (consequence) of his lifelong *horror metatheoriae*. In what follows, I will argue that his subsequent refusals to broaden the notion of '*game*' to accommodate '*internal*' processes and '*intentional*' faculties reflected an equally persistent conviction that

5 his new, more broadly 'phenomenological' (or 'physicalist', in Hintikka's usage) realm of public language-games must be '*closed*', '*complete*' and *sufficient unto itself*; and that

6 'the' *self* and its '*inner*' properties remain ('publically') *inexpressible*, and therefore marginal to this apparently more extensive and latitudinarian realm of 'linguistic' *phenomena*.

By way of preliminary gloss of 5, I would first ask the reader to consider how often Wittgenstein--surely one of the twentieth century's most relentless critics of ambiguously 'modal' assertions and 'dispositional predicates'--constructed philosophical arguments that eventuated in sententious pronouncements (ostensibly *ad se ipsum*) about what '*can*' and '*cannot*' be *done*, and what '*we*' '*should*' or '*should not*' *say* about it.

To me at least, the abundance of such aphorisms and assertions suggest that *unscharfe Grenzen* of the sort which arise in semantic analyses of '*deontic*' as well as '*alethic*' modalities (cf., e.g., Boos 2003) may have been more or less *invisible* to him, despite the fact that he made acute observations about such *Unschärferelationen* in their '*doxastic*' and '*epistemic*' counterparts..

Be that as it may, notice also that the *LPA*'s *Welt* and the *PU*'s *Lebensform*(*en*) (or *Lebenswelt*(*en*)) shared a common attribute: their '*phenomena*' (unlike those of, say, Kant) are '*appearances*' without any clearly identifiable observer to which they '*appear*' (even, so to speak, "*überhaupt*").

Whatever rationales one might offer for such 'appearances' in other contexts, I believe that Wittgenstein's repudiation of any such '*observability*' or '*intentional inexistence*' reflects once again the '*horror metatheoriae*' I have attributed to him. In the *LPA*, of course, Wittgenstein clearly believed he had found a '*resolution*' of this problem (which seems to me fundamentally methodological rather than epistemic) in his allusive and (literally) recondite analyses of '*solipsism*' and '*mysticism*'.

In his later work, I believe, a modulated form of *horror metatheoriae* underlay the tenacity, even ferocity, of his elenctic assault on whole *galleries* of rhetorical *Strohmänner* and *Schießbudenfiguren* who dared suggest that '*inner*' observations, expectations and other '*intentional*' responses to language-games might merit the accolade of game-theoretic '*existence*', even though they were not (or 'could never be') adequately manifested in *public* language-games' '*An*-' and '*Verwendungen*'.

There are of course no passages in the *PU* comparable to *LPA* 5.64. But there do seem to me lines of argument which converge on a kind of *collective solipsism* (cf. Elizabeth Anscombe's "*linguistic idealism*")--an elusive view in which

5 '*we*' express ourselves entirely in 'public' languages, and these public languages are bounded

6 '*below*' (or '*within*') by 'nonsensical' *vacua*, and

7 '*above*' (or '*without*') by equally 'nonsensical' levels of complexity and metatheoretic ascent ("cf. [*s*]*ie sind mir geschlossen*", *PU* II, 537)

Taken together, such prescriptive and proscriptive limitations apply to '*outer*' limits of 'public' language-games ('*the way up*'), and to '*inner*' ones. which enclose black boxes (or *Käferschachteln*) of intentionality ('*the way down*').They also seem to define new, ever so slightly extended topological boundaries of Wittgenstein's '*Fliegenglas*'. Wittgenstein's closed '*phenomenologies*' of '*public*' language-games *might* indeed have '*left everything as it is*'. But they have also told us very little about *what might be*, and offered few insights into 'the starry heavens above us', much less 'the moral law within us'.

It remains open to Wittgenstein's critics, in any event, to

8 acknowledge that *nonconstructive forms of concept-formation may* lead us into labyrinths of *(Philosophen*)*unsinn*, but

9 hope that they might *also* indicate new heuristic analyses and provisional modes of escape from hiss *Fliegengläsern*,

10 and offer a bit of useful instruction in consequence in *compassion* and *intellectual modesty*.

(*Who are* '*we*', *after all*, *that* '*we*' *are so mindful of* '*us*'?)

4. …WHAT 'WE' (MIGHT) 'SAY' ABOUT WHAT 'WE' (MIGHT) 'USE' TO UNDERSTAND [WHAT 'WE' (MIGHT) 'SAY' ABOUT WHAT 'WE' (MIGHT) 'USE' TO UNDERSTAND] …

I have not made any attempt in this brief essay to trace through any systematic textual correlations or interrelations between Wittgenstein's usage(s) of '*games*', '*grammar*(*s*)', '*meanings*' and '*use*(*s*)'. One rationale (or excuse) for this is that conflicting assertions in the published texts and *Nachlaß* often seem to offer exemplary instances of skeptical *isostheneia*. Instead, I will focus in this section on the *contextuality*, *liminality*, *relationality* and '*dispositionality*' which seem to me characteristic of *all* *four* of these notions, as well as most others to which he gives any serious measure of semantic or semeiotic force.

Many of Wittgenstein's more aggressive assertions about language-games' 'grammar' and dismissive assertions about what '*we*' (can) '*do*'

1 trump dialectical counterarguments which appeal to metatheoretic ascent;

2 block metatheoretic applications of otherwise persuasive elenctic arguments; and

3 permit Wittgenstein's more successful *personae* to retain the (locally) final word.

Collectively, in fact, the elenctic and eristic roles of Wittgenstein's many claims about '(*our*) *grammar*' and what '*we*' (can) '*say*''in the *PU* suggest he thought that

4 such claims might '*constitute*' *Lebensformen* (or *Lebenswelten*), (in the quasi-metatheoretic sense introduced above) much as he had once believed that

5 comparable claims about ('my') *logic*'and what '*I*' can 'show' might'*constitute*' the LPA's '*Welt*'.

To me at least, these analogies also bring to mind another, very different historical comparison: between

6 Wittgenstein rigorous '*use*'-based semantics (or semeiology) of '*meaning*', and

7 David Hume's equally rigorous and dogmatically ambiguous '*Custom*'-based semeiology of '*causation*'.

If the analogy is at all tenable, it suggests that the corresponding 'dogmatic ambiguities' came at a high price in both cases Hume, notoriously, could not even locate '*himself*', much less '*Custom*', in the tines of his 'fork' or the focal field of his 'microscope'. And Wittgenstein--comparably, I believe--could not tell '*us*' who '*we*' are, or provide a '*useful*' or applicable characterisation of '*use*'.

The principal reason for this, I believe, is that ranges of '*use*' are as *dispositional* (open to *indefinite* varieties of alternate *interpretation*) as are ranges of '*purpose*', '*intention*', '*consequence*' and (for that matter) '*consciousness*'.

Put somehat differently, in 'semantic paradoxical' form: if there 'is' no '*game*' of 'all' '*games*', why should there be a '*Verwendung*' of 'all' *Verwendungen*'? Alternatively paraphrased in Wittgenstein's ambiguously 'dialogical' manner: is there a '*use*' for '*us*' to (talk about) '*our*' '*use*'?

Aporien that arise naturally, even inevitably, in straightforward attempts to 'answer' such reflexive questions suggest

8 that *all* the notions italicised in the last paragraph--'*use*' and '*uses of* '*use*'' conspicuously among them--are comparably '*dispositional*': that is,

9 subject to adjudication and readjudication, in the very same sorts of '*interpretants*' and '*metatheoretic*' *thought-experiments* Wittgenstein deployed again and again against his own straw-opponents; and

10 that the only recourses available to Wittgenstein to '*bound*' such hermeneutic and metatheoretic ascents were to arrogate the last word (at least rhetorically), and simply *decree* that they '*must*' '*have an end*'.

(The '*last philosopher*', so to speak, can be anyone who claims the right to turn off the metatheoretical lights and leave the room) Also like Hume--or at least my interpretation of Hume--Wittgenstein

11 *conflated* his 'empirical' *horror metatheoriae* with a shadow-background *monism* (which in my view he never seriously questioned); and

12 '*grounded*' the '*unity*' of his linguistic counterparts of phenomenalists' "impressions" and "ideas" in 'complete' and hypothesis-free '*phenomenology*' for which he could not offer any 'game'-theoretic justification

(Hume, who had a wry sense of humor, once likened the '*Custom*' which grounded *his* '*phenomena*' to a form of "*preestablished harmony*" and "*final cause*"; cf. *E* I, 54-5).

Since I have used the word '*dispositiona*l' more than once, and in senses Wittgenstein roundly rejected (cf., e.g., *PU* II, 501), it may be useful at this point to pause and try to clarify the sense(s) for it (I think) I have in mind.

Consider first

13 the *primary etymology* of its verbal Latin antecedent *dis-ponere*--to place arrange or distribute in *different alternate* locations--and

14 subsequent suggestive *extensions* of this ‘literal’ usage to a kind of *liberum arbitrium*, or intentional ‘*freedom*’ to *choose* among such alternatives

(Both senses linger in odd linguistic corners--the French military expression “*(vous pouvez*) *disposer*”, for example, for the English “dismissed”).

It is presumably evident why a philosopher who rejected attempts to integrate ‘*interior*’ forms of *intentionality* as well as ‘*potentiality*’ into his language games might reject ‘game’-theoretic appeals to14, much as Hume had finessed questions of ‘*interior*’ “Liberty and Necessity” in *E* I 80-103.

Less evident may be the extent to which Wittgenstein was equally wary of conceptual or hermeneutic counterparts of the ‘physical’ *plurality* in 13, and determined to preserve some sort of *monism* or metaphysical *realism* at the (still ineffable) margin of an allegedly ‘complete’ epistemic ‘phenomenology’ of public *Sprachspiele*.

In opposition to these commitments, I conjecture that *hermeneutic plurality* and *game-theoretic incompleteness* (a weak form of ‘*freedom*’) may ‘regulate’ ‘game’-internal *individuation*--make it ‘possible’, in some quasi-Kantian sense.

In physical terms, such *individuation* might be likened to ‘*artificial*’ *isolation and* ‘*localisation*’ *of systems under* ‘*controlled*’ *observation*.

In ‘conceptual’ terms, it may be likened to ‘*intentional*’ *isolation and* ‘*localisation*’ *of* ‘*games*’ *under* ‘*metatheoretic*’ *or* ‘*hermeneutic*’ *observation*.

To me, such analogies also suggest

15 that *‘the world*’ may be *filled* with ‘dispositional’ *‘games*’ and systems open to indefinite varieties of *experimentation and thought-experimentation*;

16 that indefinite semeiotic hierachies of such ‘*experimentation*’ might offer an inherently *in*complete but arguably ‘phenomenological’ account of ‘*experience*’.

Such an account of '*experience*' would clearly be a process-theoretic form of *theoretical relativism*. It might well leave many local aspects of such '*experience*' '*as it is*'. But it would also countenance *indefinite suspension of judgment* about '*global*' counterparts of '*local*' *completeness* (*and* '*individuation*').

Whatever the merits of such an account, I have tried to argue the case that wider interpretations of '*game*' along the lines I have sketched might have permitted a more skeptically and metalogically inclined 'Wittgenstein' to modulate his early semantic monism in the direction of such a view.

Along the way, however, such a counterfactual 'Wittgenstein' would have had to

17 consider '*dispositional*' aspects of game-theoretic accounts of '*meaning*' as well as '*use*';

18 attribute provisionally '*intentional*' roles to certain 'metatheoretic' language games and (potential) *Verwendungen* of them;

19 grant that *some* '*interior*' games *might* (provisionally) 'regulate' other games they '*constituted*' or '*interpreted*' ('the way down', mentioned earlier); and

20 acknowledge that there might be no '*end*' in *any* game to other games' potential (re)interpretations of what that game could be '*about*'.None of this, of course, was acceptable to him. Again and again, he

21 tacitly assumed that certain dispositional or theory-relative notions, such as '*Verwendung*,' were more or less unequivocally applicable, if not well-defined; but

22 denounced other, comparably dispositional or theory-relative notions, such as '*intentionality*', when his thought-experiments made it clear to him that they were *plurally interpretable*, and therefore *Philosophenunsinn*. Wittgenstein, in short could not reconcile his newfound *syntactic pluralism* with

23 the *monist presuppositions* that continued to run in the background of his pluralist thought-experimentation; or with

24 the '*phenomenological*' *demands of completeness* he (wrongly) believed 'must' follow from the game-theoretic homogeneity and

universality he sought. He could only formulate local *elenchoi* which oscillated between the recurrent polarities of these irreconcilable demands.

5. CONCLUSION

Ancient, early modern, Kantian and twentieth-century views of ‘*appearances*’ or ‘*phenomena*’ have clearly had skeptical as well as dogmatic aspects, and I have argued at some length above that ‘*phenomenologies*’ might either be

1 *relativised* to particular classes of ‘*phenomena*’ to be ‘*encompassed*’; or

2 putatively ‘*all-encompassing*’, but indefensibly *begged*.

In this brief final section, I will try to draw together these arguments to conjecture that Wittgenstein equivocations between this dichotomy’s ‘*skeptical*’ and ‘*dogmatic*’ poles was one of the constants of his philosophical career.

‘The *early* Wittgenstein’ in section 1’s reconstruction took a deeply reductive and dogmatic view of the propositional and predicate logic his predecessors Peano, Frege and Russell had devised. But he also characterised an ineffably ‘unique’ *interpretation* of scientific *extensions* of this logic with the aid of an ancient skeptical (or ‘mystical’) image of a ‘*ladder*’ that ‘*sublates*’ itself.

‘The *later* Wittgenstein’ reconfigured this ‘ladder’ as the shifting topological boundaries of a metaphorical *Fliegenglas*--an image in which he superposed a drastically skeptical perspective-shift onto a grotesque miniature of the Platonic ‘cave’. But he also shattered the ‘*mirror*’ of ‘logic’ and its extensions in ‘the world’ into a *kaleidoscope* of ‘language-games’ and ‘*Lebensformen*’.

Throughout all this, both ‘Wittgensteins’ suggested again and again

3 that ‘*we*’ (or ‘one’) could somehow make a (noumenally or collectively) ‘self’-sufficient and perhaps ‘*complete*’ *semeiology* out of ‘*our*’ *Sachverhalte* (respectively, ‘our’ public ‘language-games’ and their ‘*uses*’).

Both these claims--‘early’ and ‘late’--might be assimilated to Hume’s assertion--in his role as the ‘Newton’ of the ‘moral sciences’ (cf. *E* I, 14)--

4 that '*we*' could somehow make an (empirically or collectively) self-sufficient and perhaps '*complete*' *phenomenal psychology* out of '*our*' 'impressions, and the ideas 'we' 'derive' in public and 'customary' ways from them.

Indeed, I believe this analogy with Hume can be extended in informative (and non-Kripkean) directions. In his role as dogmatic 'empirical' phenomenalist, for example, Hume begged the adequacy, sufficiency, closure, completeness and uniformity of what 'we' *perceive and understand*. But he also dismissed '*internal*' interpretations of such perceptions as understandings as irrelevant appeals to 'inner' forms of "animal *nisus*" (cf. *E* I, 77).

In his role as dogmatic 'linguistic' phenomenologist, the later Wittgenstein begged the adequacy, sufficiency, closure, completeness and uniformity of what 'we' *do and say*. But he also dismissed '*internal*' interpretations of such actions and utterances as irrelevant recourses to (pseudo)scientific '*nonsense*'.

In more 'skeptical' modes, finally, *both* Hume and Wittgenstein ('early' as well as 'late')

5 expressed persistent interest in extreme forms of finitism (cf., e.g., Hume, *Tr* II,1);

6 implicily or explicitly denounced forms of 'metaphysics', even as they engaged in it;

7 explicitly renounced efforts to characterise personal identity in their own phenomenalist (or phenomenological) terms; and

8 responded derisively but ambivalently to the prospect of indefinite ranges of mathematical (or physical) thought-experimentation.

To me at least, these analogies and dualities suggest that there were indeed

9 *two philosophical* '*Wittgensteins*'--the *skeptical thought-experimentalist* and the *dogmatic* '*phenomenologist*'--much as there were

10 *two philosophical* '*Humes*'--the '*consequent skeptic*' and the *dogmatic* '*phenomenalist*'.

Not unlike their 'Humean' counterparts, moreover, '*dogmatic' Wittgenstein* and his '*skeptical*' *Doppelgänger*--'early' and 'late'--wanted to have their '*complete*' '*phenomenologies*' and *relativise* them too. The '*early*' '*dogmatic phenomenologist*' Wittgenstein, for example, effectively '*prescinded*' (*from* '*the*' '*world*'*-internal semantics of* '*what is the case*') *whole classes of theories which might be extraordinarily useful to students of anthropology and cognitive science*, as well as others who might to seek to investigate interactions and interrelations between the two. Some (but not I) might call such '*prescission*' '*skeptical*'.

The '*later*' '*dogmatic phenomenologist*' Wittgenstein, by contrast, effectively '*prescinded*' (*from* '*unserer' 'Grammatik*'' *und Lebensform*(*en*)) *whole classes of theories which might be extraordinarily useful to students of anthropology and cognitive science*, as well as others who might to seek to investigate interactions and interrelations between the two Some (but not I) might call *this* '*prescission*' '*skeptical*' as well.

At their very worst, I believe, these internal tensions and dialectical *Gegensätze* in Wittgestein's work did *not* eventuate in '*quietism*', or even in the '*nihilism*' Hintikka observes in the interpretations of the late Burton Dreben, but in aggressive forms of philistine *intolerance*. Little would be achieved by attempts to rationalise such intolerance as '*therapy*', '*social construction*' or '*deconstruction of metaphysics*'.

At his best, however, Wittgenstein was *not* a quietist, or a nihilist, much less a philistine.

'*Quietists*', for example, seldom devote thousands upon thousands of hours to unsuccessful attempts to assimilate the '*centers*' of their *dogmatic presuppositions* to the '*circumferences*' of their *skeptical thought-experiments*. And his agonistic efforts were hardly guided by 'pragmatic' philosophical principles to '*leave everything as it is*'. They were guided by forlorn but tenacious hopes--expressed by the wistfully resigned remarks in his preface to the manuscript of the *Untersuchungen*--to "*bring forth a good book*".

("*Wer immer strebend sich bemüht*, *den können* '*wir*' *erlösen*.")

6. POSTSCRIPT

The purpose of this brief afterword is to sketch (as promised in section 2), the following proposition:

6.1 Proposition

Let M be an a unary predicate of a consistent recursively axiomatisable theory T* which extends Peano arithmetic, and assume (metatheoretically) that

6.1.1 M is a T*-definable structure on the integers N of T* which syntactically interprets a recursively axiomatisable subtheory T of T* in T*; and that

6.1.2 '{...}', an integer-valued function, recursively '*codes*' the syntax and consequence-relations of T and T* (cf., e.g., Smorynski, 835-838).

Then

6.1.3 T* *cannot decide which* '*rules*' R *it imposes on* M *can be* '*followed*'.

More precisely:

6.1.4 *no unary predicate* S *of* T* *has the property that* [S(r) if and only if [r codes a unary predicate R* of T*, and the restriction R of R* to the 'universe' N of M has a least element]] *is provable in* T*.

Proof (**sketch**)

In order to obtain a contradiction, we assume that such a unary predicate S *does* exist in T, and form the following binary predicate L:

[L(u,v) iff

[[S(v) and v codes a unary V* such that the restriction V of V* to N is nonempty but not all of N]

implies that

[u is not the least element in N which is not in V]].

Invoking the *parametric diagonal lemma* (set out in Smorynski, 827, Boolos, 49 or Boos 1998, as I mentioned earlier), we can now adduce the existence of a unary predicate K in the language of T* such that

[K(v) if and only if *not*-L(v,{K($\dot{v}$)}) for every v] is provable in T*.

(Here {...} codes the formal expression '...' in T*, as above, and {K($\dot{v}$)} the result of syntactical *substitution* of v in K; cf. Smorynski, 837.)

Setting r = {K(v)}, T* would therefore prove that

6.1.5 [K(r) if and only if *not*-L(r,{K($\dot{r}$)})].

Since **6.1.5** asserts in T* that r satisfies the nonvacuous predicate K if and only if r is the least element of N which does *not* satisfy K, we can finally infer from our metatheoretic consistency-assumption that no such S can exist in the language of T*.

Among other things, these observations suggest that the *sorites* is essentially a metatheoretic counterpart of the *Berry paradox*, and may therefore be exploited to yield a straightforward alternative derivation of Gödel's results along lines sketched in Boos 1998, 71-72.

REFERENCES

Aquinas, Thomas, Summa Theologiae. Madrid: Biblioteca de Autores Cristianos, 1951

Aristotle, Metaphysics, ed. with an English translation by Hugh Tredennick. Cambridge: Harvard University Press, 1975 and 1977.

Aristotle, Nicomachean Ethics, ed. with an English translation by H. Rackham. Cambridge: Harvard University Press, 1975.

Boolos, George, The Unprovability of Consistency. Cambridge: Cambridge, 1979.

Boos, William, "The *Transzendenz* of Metamathematical 'Experience'", *Synthese* 114 (1998), 49-99

Boos, William, "Virtual Modality", *Synthese*, 136 (2003), 435-491.

Descartes, René, Oeuvres de Descartes, ed. Charles Adam et Paul Tannery, Nouvelle Présentation. Paris: Vrin, 1973-1978 (AT).

Enderton, Herbert, A Mathemtatic Introduction to Logic. New York: Academic, 1972

Henkin, Leon, "Completeness in the Theory of Types", Journal of Symbolic Logic, 15 (1950), 81 - 91.

Hintikka, Jaakko, Ludwig Wittgenstein, Half-Truths and One-and-a-Half Truths (Jaakko Hintikka Selected Papers, Volume I). Dordrecht: Kluwer, 1996.

Hintikka, Jaakko, Language, Truth and Logic in Matheamtics (Jaakko Hintikka Selected Papers, Volume III). Dordrecht: Kluwer, 1998.

Hintikka, Jaakko, Wittgenstein. Belmont: Wadsworth, 2000.

Hintikka, Merrill and Hintikka, Jaakko, Investigating Wittgenstein. Oxford: Blackwell 1986

Hume, Enquiries, ed. P. Nidditch. Oxford: Clarendon, 1975.

Hume, A Treatise of Human Nature, ed. L Selby-Bigge and P. Nidditch. Oxford: Clarendon, 1978.

Husserl, Edmund, Cartesianische Meditationen. Hamburg: Meiner, 1977.

Husserl, Edmund, Logische Untersuchungen II/1. Tübingen: Niemeyer, 1913.

Kant, Immanuel, Kritik der reinen Vernunft, ed. Raymund Schmidt. Hamburg: Meiner, 1956. (KdrV)

Leibniz, Gottfried, Die philosophische Schriften von Gottfried Wilhelm von Leibniz, ed. C. Gerhardt. Hildesheim: Olms, 1978 (GI - GVII)

Plato, Parmenides, ed. with a translation by H. N. Fowler. Cambridge: Harvard, 1977.

Plato, Republic, ed. with a translation by Paul Shorey. Cambridge: Harvard, 1969.

Plato, Theaetetus and Sophist, ed. with a translation by H. N. Fowler. Cambridge: Harvard, 1977.

Sacks, Oliver, The Man Who Mistook His Wife for a Hat. New York: Harper, 1990.

Shoenfield, Joseph, Mathematical Logic. Reading: Addison-Wesley, 1967.

Spinoza, Baruch, Werke II (Tractatus de Intellectus Emendatione and Ethica). Darmstadt: Wissenschaftliche Buchgesellschaft, 1967.

Wittgenstein, Ludwig, Logisch-Philosophische Abhandlung and Philosophische Untersuchungen, in: Werkausgabe in Acht Bänden (Band I). Frankfurt am Main: Suhrkamp, 1984. (*LPA*, *PU*)

Wittgenstein, Ludwig, Philosophische Bemerkungen, in: Werkausgabe in Acht Bänden (Band II). Frankfurt am Main: Suhrkamp, 1984. (*PB*)

Wittgenstein, Ludwig, Philosophische Grammatik, in: Werkausgabe in Acht Bänden (Band IV). Frankfurt am Main: Suhrkamp, 1984. (*PB*)

Wittgenstein, Ludwig, Bemerkungen über die Grundlagen der Mathematik, in: Werkausgabe in Acht Bänden (Band VI). Frankfurt am Main: Suhrkamp, 1984. (*BGM*)

Wittgenstein, Ludwig, Über Gewißheit. New York: Harper and Collins, 1972. (*ÜG*)

Wittgenstein, Ludwig, Nachlaß. Oxford: Oxford (CD Rom Edition), 2000 (*N*)

THEORETICAL COMMENSURABILITY BY CORRESPONDENCE RELATIONS:WHEN EMPIRICAL SUCCESS IMPLIES THEORETICAL REFERENCE

Gerhard Schurz
University of Duesseldorf

1. HINTIKKA ON THE (IN)COMMENSURABILITY OF THEORIES

According to Kuhn (1962), the theoretical terms of two competing theories are incommensurable in the sense that there do not exist logical or conceptual translation relations between them. I call this kind of incommensurability Kuhn-incommensurability. This thesis of Kuhn is supported by the standard analysis of theories in philosophy of science.[39] According to this analysis, the nonlogical terms of a scientific theory T divide into two classes: the so-called non-T-theoretical terms whose meaning is independently given, and the so-called T-theoretical terms, whose meaning is not independently given, but is specified by the theory T itself. For example, time, position and its derivatives w.r.t. (with respect to) time are non-theoretical, while mass and force are theoretical in classical mechanics. It follows from the standard analysis that the meaning of mass and force in classical mechanics is different from their meaning in special relativity theory. The two theories are Kuhn-incommensurable w.r.t. mass and force.

Hintikka (1988) suggests to evaluate the degree of (in)commensurability between two theories in terms of their consequences. I speak here of Hintikka-(in)commensurability (Hintikka speaks of consequential (in)commensurability). Thereby, Hintikka uses the interrogative concept of consequence which is relativized to a certain background information B, where B is the class of answers which are available from an information source. In our context, the information source

[39] Cf. Carnap (1956), Hempel (1951, 1958), Feyerabend (1962), Lewis (1970), Sneed (1971), Papineau (1996) (etc.). Cf. also Schurz (2004).

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 101-126.

B contains the results empirical observations and experiments which one *would* know after a sufficient time of empirical research. According to Hintikka's completeness theorem (1999, 53f) it holds that a statement S is a model-consequence of theory T given the answer set B iff B, T $\Vdash$ S (where "$\Vdash$" stands for ordinary logical consequence). Hintikka's basic idea is to relativize the degree of commensurability between two theories T_1 and T_2 to a set Q of yes-or-no-questions of the form $(?)(S_i \vee \neg S_i)$ $(i \in I)$. Answers to these questions have the form S_i or the form $\neg S_i$. Note that the statements in B can be different from any of statements $(\pm)S_i$ ($i \in I$; "$(\pm)$" for "unnegated or negated"). A theory T, represented as a set of statements, answers a question $(?)(S_i \vee \neg S_i)$ iff $B \cup T \Vdash (\pm)S_i$. Let $Q_1 \subseteq Q$ be the set of questions answered by T_1; so $A_1 = \{S_i: B \cup T_1 \Vdash S; i \in I\}$ is the set of answers given by T_1 to the questions in Q_1. And likewise for Q_2 and A_2. $Q_1 \cap Q_2 \subseteq Q$ are the questions answered by both T_1 and T_2, and $Q_1 \cup Q_2 \subseteq Q$ are the questions answered at least by one of the two theories. Like Hintikka we first assume the case where the two theories are compatible with respect to $Q_1 \cap Q_2$, that is, they *give the same answers* to the questions which they both answer. If and only if this is the case, the union of answers $A_1 \cup A_2$ will be logically consistent, and the intersection $A_1 \cap A_2$ will contain an answer to each of the questions in $Q_1 \cap Q_2$ which is given by both theories. Under these conditions, Hintikka (1988, 234) suggests to measure the degree of commensurability between theories T_1 and T_2 relative to the question set Q, abbreviated as $c(T_1,T_2|Q)$, as follows:

$$\text{(Comm)} \quad c(T_1,T_2|Q) \;=\; \frac{\inf(\bigwedge A_1 \cap A_2)}{\inf(\bigwedge A_1 \cup A_2)}\,.$$

Here "$\bigwedge X$" denotes the conjunction of the members of a finite set of statements X, and "inf S" denotes the degree of information of a statement S, defined as inf S := –log p(S), where p(–) is a given probability measure over the given 1st order language. Hintikka (1988, §6) suggests to use the uniform probability measure over the Hintikka-constituents in a 1st order language. For purpose of this paper, nothing hangs on the choice of a probability measure.

If the set of answers $A_1 \cap A_2$ given by both theories to questions which both can answer is empty but $A_1 \cup A_2$ is non-empty, then the degree of commensurability (relative to Q) is minimal, namely 0. This corresponds to the most radical case of incommensurability, which is given when the two theories do not only have any non-logical concept in common. According to some radical interpretations of Kuhn and Feyerabend, cases of zero-commensurability are possible because observations are generally theory-dependent. But historically speaking, such cases are unrealistic: there is always a shared empirical or pre-theoretical language which two competing theories have in common and in terms of which basic observational data are described. On the other hand, if all questions answered by one theory are also answered by the other theory in the same way, and vice versa, then the degree of commensurability is maximal, namely 1. In this case the two theories share all of their theoretical concepts – at least all those concepts which are relevant in answering questions in Q.

Let L_1 be the language of theory T_1 (the set of formulas constructible in its vocabulary) and L_2 the language of theory T_2. The most interesting case is given when the two theories share a certain empirical or non-theoretical sublanguage $L_e = L_1 \cap L_2$, but have in addition their distinct T_1-theoretical and T_2-theoretical vocabulary. If we assume that the set Q of questions contains all questions which can be expressed in the shared empirical sublanguage L_e, then our assumption that the two theories give the same answers to questions which the can both answer means that the two theories are *empirically compatible*, that is, they do not imply opposite statements of L_e.

Hintikka generalizes his definition of the degree of commensurability to the case of theories whose answers to jointly answerable questions are incompatible in an interesting way which need not be discussed here.[40] If the two theories are empirically incompatible, then it is possible to decide between the two theories by means of empirical investigation. The more challenging case which will be the subject of the following considerations is given when the two theories are empirically compatible – at least w.r.t. a given class of domain-restricted empirical consequences.

Crucial in Hintikka's measure of commensurability is its relativization to a question set Q. First of all, we shall require that Q contains only *non-trivial* questions w.r.t. T_1 and T_2 in the following sense: if $(?)S \vee \neg S \in Q$ and $B \cup T_i \Vdash\!\!\!— S$, then S must not be a completely irrelevant consequence of $B \cup T$ in the sense that a predicate (or function symbol) is uniformly replaceable in S by arbitrary other predicates (of the same arity) salva

[40] Hintikka (1988, 235) suggests to replace the numerator of the definiens in (Comm) by the *average* of the information of the set of answers of T_1 to questions which both T_1 and T_2 can answer, and the information of the set of answers of T_2 to questions which both theories can answer. Hintikka's move is certainly reasonable for the numerator of the measure $c(T_1,T_2|Q)$, but it seems to me that Hintikka's move is not sufficient concerning the denominator of the measure $c(T_1,T_2|Q)$, because the set of answers $A_1 \cup A_2$ will be inconsistent if the two theories give opposite answers to at least one question, and then $c(T_1,T_2|Q)$ will be zero. A plausible suggestion to repair this unwelcome consequence would be to replace the information $\inf(\bigwedge A_1 \cup A_2)$ in the denominator by the sum $\inf(\bigwedge A_1) + \inf(\bigwedge A_1)$.

validitate of $B \cup T_i \Vdash\!\!\!\!\text{—}\, S$ (for details on the concept of relevant consequence cf. Schurz 1991). If we let $Q = Q_e$ be the set of all non-trivial questions expressible in the shared empirical language of T_1 and T_2, then Hintikka's measure gives us a notion $c(T_1,T_2|Q_e)$ of *empirical commensurability*: $c(T_1,T_2|Q_e)$ will be minimal, namely zero, if the two theories have non non-trivial empirical consequences in common, that is, if they speak about entirely different empirical domains, and $c(T_1,T_2|Q_e)$ will be maximal, namely 1, if the two theories agree in all of their empirical consequences, that is, if they are *empirically equivalent*. Interesting cases which lie in between these two extreme cases are those where empirical consequences classes of the two theories have a significant overlap, but either (i) T_1 has additional empirical consequences which T_2 does not have, but not vice versa, or (ii) both theories have empirical consequences which the other one does not have. If all the empirical consequences are confirmed, then in case (i) T_1 is empirically more successful than T_2, and in case (ii) T_1's and T_2's empirical success complement each other.

Let us now consider the set $Q = Q_t$ of non-trivial questions which are expressible in the distinct theoretical language of the theories. So Q_t is the set of questions expressible in the language $L_1 \Delta L_2 := (L_1 - L_2) \cup (L_2 - L_1)$, which is the so-called symmetric difference between the two languages and contains all formulas expressible in *only* one of the two languages. Let us call the resulting measure of commensurability $c(T_1,T_2|Q_t)$ the measure of *theoretical commensurability*. What we then obtain is the following:

Incommensurability Theorem: Provided that the background information B is expressed in the shared (non-theoretical) language of two theories T_1 and T_2, and the question set Q contains only non-trivial questions (w.r.t. T_1 and T_2), then the measure of the theoretical commensurability $c(T_1,T_2|Q_t)$ between two theories is zero.

Proof: Every question $(?)(S \vee \neg S)$ in $L_1 \Delta L_2$ S will either belong to the language $L_1 - L_2$ or to the language $L_2 - L_1$; let us assume the first case (the second is similar). We proof that $B \cup T_2 \Vdash\!\!\!\!\text{—}\, S$ *cannot* hold for a non-trivial question $(?)(S \vee \neg S)$ (w.r.t. T_1, T_2). Assume $B \cup T_2 \Vdash\!\!\!\!\text{—}\, S$. S contains at least one T_1-theoretical predicate which is not the language of $B \cup T_2$, and so this predicate must be uniformly replaceable in S salva validitate of $B \cup T_2 \Vdash\!\!\!\!\text{—}\, S$, which means that $(?)(S \vee \neg S)$ is a trivial question. It follows that no non-trivial question in Q_t is answerable by both theories.

What we have obtained from Hintikka's commensurability measure is a result which has been widely accepted in post-Kuhnian philosophy of science: two rivalizing theories may be commensurable with respect to their common empirical vocabulary, but with respect to their theoretical vocabulary, they are incommensurable. Thereby, I have taken it for granted that there is no way of isolating analytic meaning postulates which connect the theoretical terms of the two theories and which are contained in the theory-independent background information B. This follows from the assumption that has been explained in the beginning of this section, namely that the meaning of the theoretical concepts of a theory is specified by the theory itself. Although Hintikka does not seem to share this view (1988, 227), I have argued that this view is well established, or at least I assume that it is.

What I conclude from this result is that Hintikka's logical measure of commensurability is *adequate* because it does what it should do: it clearly reveals the crucial problem which has beset philosophers of science since Kuhn and does not sweep it under the carpet. Concerning the relation between Kuhn-(in)commensurability and Hintikka-(in)commensurability, our result can also be expressed by saying: Kuhn-incommensurability implies Hintikka-incommensurability. For if T_1 and T_2 are theoretically incommensurable in Kuhn's sense, then there are no conceptual relations between the distinct theoretical terms of two theories T_1 and T_2, and this implies that T_1 and T_2 are theoretically incommensurable in Hintikka's sense. This result supports Hintikka's assertion that his "consequential definition of commensurability does do justice to the idea of incommensurability as untranslatability" (1988, 239).

2. INCOMMENSURABLE THEORIES WITH EQUAL EMPIRICAL SUCCESS: THE ARGUMENT AGAINST SCIENTIFIC REALISM

The systematic possibility of competing theories which have equal empirical success in a given domain of applications, but which are theoretically incommensurable (in the Kuhn- or Hintikka-sense) is also the main argument for scientific anti-realism, or instrumentalism. Quine (e.g., 1960, 141ff; 1975; 1992, §41) has repeatedly demonstrated that given an empirically successful theory T, one can usually construct an empirically equivalent theory T* which has a completely different theoretical superstructure. Quine has called this phenomenon the *empirical underdetermination* of theories, and what follows from it is that the inference from the empirical success of a theory to the real existence of its theoretically postulated entities and to truth of its theoretical claims is unjustified. Quine's argument of empirical underdetermination is also the main couterargument to Putnam's miracle argument (e.g. 1975, 73), which goes as follows: if a scientific theory T is empirically successful for a long period of time, then it would be as improbable as a miracle if this theory would not also be approximately true in the realistic sense. Quine has shown that for every empirically successful theory T it is almost always possible to construct an equally empirically successful theory T* with a completely different theoretical superstructure. So Putnam's miracle argument *cannot* be right, *under the assumption that the theoretical terms of the two incommensurable theories cannot both have reference altogether.* This assumption that joint reference is impossible is usually made, and it seems to be supported by the artificial examples of Quine as well as by historical examples. For example, it is said that if Lavoisier's oxygen theory of combustion is true, then Stahl's concept of a "phlogiston" cannot have reference. It should be observed, however, that this assumption does *not* follow from the logical analysis of commensurability: even if the two theories are theoretically incommensurable in Hintikka's sense, it is logically possible that they both are realistically true. This possibility which is revealed by Hintikka's analysis will be of crucial importance for the main result of this paper: the correspondence theorem.

In his deep-going analysis of Putnam's miracle argument Carrier (2003) imposes an important qualification upon the notion of empirical success. It is always possible to construct a theory *post factum* as to fit certain given empirical data which were already known *beforehand*. What really counts for the empirical success of a theory are episodes where the theory correctly predicts a *novel* empirical phenomenon which at the time of the theory construction was neither known nor expected to be true by means of empirical induction. Carrier calls this kind of success *strong empirical success* (2003, §4), and many philosophers of science agree that what really counts in the history of science is this strong form of empirical success – the prediction of novel phenomena. Newtonian mechanics predicted a variety of novel phenomena. For example, the very prediction of a gravitational force which acts between all kinds of physical bodies was completely novel and unexpected: most of Newton's contemporaries did not believe that gravitational forces would also exist between normal-sized objects such as this pencil and this blackboard. Only much later, in 1798, when Henry Cavendish gave an experimental demonstration of the effect of gravitation between normal-sized objects with a torsion scales, the majority of empirical scientists became convinced. So let us accept Carrier's qualification of Putnam's miracle argument, because it is very well confirmed by the history of science.

Carrier suggests to explicate Putnam's miracle argument in the following *improved* version: if a theory is strongly empirically successful (i.e., has correctly predicted novel phenomena), then it would as improbable as a miracle if the theory's theoretical terms would not refer, and the theory would not be at least approximately true. Vis à vis this weakened version of the miracle argument, almost all of Quine's counterexample loose their bite, because almost all of them are *post hoc* constructions. What Quine shows is that if a certain amount of empirical data E and a certain theory T confirmed by that data in E are *given*, then by some logico-mathematical tricks it is usually possible to construct a theory T* which differs from T significantly in its theoretical superstructure but explains the data in E equally well. Of course, none of these artificially constructed alternative theories T* has ever predicted a novel phenomenon. This is even true for the example of Poincaré's alternative cosmology which is frequently quoted by Quine (1975, 322; 1992, §41). Carrier argues that what one would really need to refute Putnam's improved miracle argument would be real examples of scientific theories which were strongly empirically successful for a certain period of time, but which have been supplanted afterwards by equally or more successful theories with a significantly different theoretical superstructure, so that from that time on nobody believed in the reference of the theoretical terms of the old theory any more, despite of its strong empirical success. What makes Carrier's paper so interesting is that he in fact gives two historical examples of this sort: the phlogiston theory and the theory of caloric (2003, §7). We now turn to these examples.

3. CARRIER ON THE STRONG EMPIRICAL SUCCESS OF THE PHLOGISTON- AND THE CALORIC-THEORY

According to the phlogiston theory of combustion which goes back to Georg Stahl around 1730 and was significantly developed by Henry Cavendish in 1766 and Joseph Priestley in 1782, every material which is capable of being burned or calcinated (roasted) contains a simple substance which is immaterial or at least different from ordinary matter, and which was called phlogiston.[41] Phlogiston was though to be the bearer of combustibility. When the combustion or calcination takes place, the burned or calcinated substance delivers its phlogiston, usually in form of a hot flame or an evaporating inflammable gas, and a dephlogistonated substance-specific residual remains. In the 1780s, Lavoisier introduced his alternative oxygen theory according to which every combustion and calcination consists in the

[41] For the following cf. Carrier (2003, §7); Thagard (1992, §3.1-3.3); Ströker (1967, 115-144).

oxidation of the substance being burned or calcinated, that is, in the formation of a chemical bond of its molecules with oxygen; the assumption of the existence of a special bearer of combustibility became superfluous in Lavoisier's theory. In modern chemistry, Lavoisier's theory is still accepted a generalized form (see below) and nobody believes in the existence of phlogiston any more.

An important domain of application of the phlogiston theory was the calcination of metals. In terms of the modern oxygen theory, calcination of metals corresponds qualitatively the following chemical reaction:[42]

Metal + Oxygen $\rightarrow$ Metalcalx (= Metaloxide) (e.g.: $3Fe + 2O_2 \rightarrow Fe_3O_4$)

but in terms of the phlogiston-theory this reaction is described as follows:

Metal $\rightarrow$ Metalcalx + Phlogiston$\uparrow$

The phlogiston theory was carried on by Henry Cavendish who dissolved various metals in acids (such as hydrochloric acid HCl, or sulfuric acid H_2SO_4; Cl is the element Chlorium and S is Sulfur). In general, the molecular structure of an acid has the form H_nX; where the n hydrogenium atoms give off their electrons to their negatively charged n-valenced partner X. While in terms of modern chemistry, the qualitative reaction is the following

Metal +Acid (H-X) $\rightarrow$ Metal-X + Hydrogenium$\uparrow$ (e.g. $Fe + 2HCl \rightarrow FeCl_2 + H_2\uparrow$),

while in terms of phlogiston theory the reaction was

Metal + Acid $\rightarrow$ Metalcalx-Acid-Solution + Phlogiston (inflammable gas) $\uparrow$.

Cavendish thought that the evaporating inflammable gas was identical with phlogiston. It should be emphasized, however, that the identification of phlogiston with hydrogenium gas was not fully acceptable because it did not

[42] *Chemical Notation:* the substances mentioned left from the arrow are input substances and those right from the arrow are output substances of the chemical reaction. "↑" at a substance means that this substance is an evaporating gas. Fe stands for Ferrum (iron), O for oxygen, H for Hydrogenium; the lower indices denote the number of atoms in the molecule

work in other domains of the phlogiston theory. For example, the end product of the combustion of coal – Carbondioxide gas (CO_2; C for Carbonium) – was also identified with phlogiston or with phlogistonated air; but it was (of course) never possible to gain hydrogenium gas from carbonium dioxide. Another domain of phlogiston theory was animalic respiration. Priestley had identified oxygen gas with completely dephlogistonated air; and respiration was identified with inhaling dephlogistonated air and exhaling phlogiston; but what one exhales is again not Hydrogenium but Carbondioxide. Nevertheless phlogiston theory need not be given up by these problems, for one could think phlogiston is always released in combination with some other gaseous component.

The empirically important point of the phlogiston theory was that combustion, calcination, and saltification (solution in acid) was modeled as chemical process in the which substance under consideration *delivers* something – let us speak here of *dephlogistonation*. It was known from many other chemical reactions that these reactions can be *inverted*. In 1782 Joseph Priestley set up the bold conjecture that it should be possible to invert the process of dephlogistonation by adding phlogiston in the form of inflammable air (hydrogenium) to a metal calx. Priestley heated several metal calxes in inflammable air and observed that the inflammable air was almost completely absorbed and that the calxes were slowly reconverted into the metals. Priestley had also recorded the emergence of water droplets in this reaction, but he assumed that the water was contained in the inflammable air from the beginning. In modern terms, Priestley has synthesized the following reaction of *reduction*, which is the inversion of a process of oxidation

> Metaloxide + Hydrogenium $\rightarrow$ Metal + Water (e.g., $FeO + H_2 \rightarrow Fe + H_2O$),

while in terms of the phlogiston theory he had performed the following reaction:

> Metalcalx + Phlogiston (+ Water) $\rightarrow$ Metal (+ Water)

In this way, the phlogiston theory has predicted a completely novel phenomenon, the transformation of metalcalxes into pure metals by the addition of phlogiston, which was celebrated as a great success of the phlogiston theory. Phlogiston theory was strongly successful in Carrier's sense, although phlogiston does not even exist.

There were several empirical reasons why Lavoisier later concluded that his oxygen theory was true and that the postulate of phlogiston was superfluous. For example, the combustion of certain substances such as phosphorus or sulfur produced a residue substance which paradoxically increased in its weight after the combustion process was over and the phlogiston has left the substance. Moreover, while phlogiston theorists were unable to isolate phlogiston in a way which could explain all cases, Lavoisier was able to isolate oxygen in a way which could explain all cases of combustion in a uniform way. Nevertheless, if we restrict the phlogiston theory to a certain range of applications, such as the oxidation and saltification of metals and the retransformation of metalcalxes into pure metals, then the phlogiston theory was strongly empirically successful with respect to these domains.

Now let us turn to Carrier's second example. At the time when Lavoisier has successfully developed his oxygen theory, chemists still believed in the theory of caloric as an explanation of heat. Every material substance contains some amount of caloric, and this amount is responsible for the heat or temperature of the substance − the more caloric it contains, the hotter it will be. Thereby caloric was assumed to be a substance consisting of weightless particles. While the particles of material substances *attract* each other in a substance-specific way, which is demonstrated by the forces of cohesion, the caloric particles *repel* each other, which is confirmed by thermal expansion, i.e. by the fact (almost) all substances expand in their volume when their temperature and hence the amount of caloric which they contain increases. In the solid state of substances, the attractive forces among the material particles dominate the repulsive forces among the caloric particles, and this holds the solid substance together. In the fluid state, the repulsive forces between the caloric particles becomes stronger but not yet dominant. Finally in the gaseous state these repulsive forces become completely dominant, so that the attractive forces between the material particles are negligible. These principles of the theory of caloric imply that the thermal expansion of gases, that is, the dependence of their volume on their temperature, should be entirely caused by the thermal expansion of caloric and, hence, should be *the same* in all gases, independent from their material nature. This was a prediction of a novel phenomenon, which was independently confirmed by John Dalton and Joseph-Louis Gay-Lussac in 1802 in their phenomenological gas law which asserts that under constant pressure, the volume of any (sufficiently ideal) gas is proportional to its absolute temperature.

The modern theory of thermal expansion is based on the kinetic theory of gases, which has been developed by Ludwig Boltzmann and Clerk Maxwell at the end of the 19th century. According to this theory, a special immaterial substance such as caloric does not exist and need not be postulated in order to explain heat and thermal expansion. Heat is nothing but the mean kinetic energy of the molecules. In the gaseous state, the distance between the molecules of the gas is so large that the volumes of its molecules and the attractive forces between its molecules are negligible. Therefore gases of equal temperature under the same pressure will have the same volume.

Again, the theory of caloric had several disadvantages, for example, it was never possible to isolate caloric, and it was strange to assume that caloric had no mass. But if the caloric theory is restricted to certain domain of applications, namely to the thermal expansion of gases, this theory was strongly successful. So we have a second example of a strongly empirically successful theory whose central theoretical term – caloric –does not refer to anything real.

4. HIDDEN CORRESPONDENCE RELATIONS

In this section I want to argue that Putnam's improved miracle argument is still right in a certain sense, *in spite* of these historical examples. A *closer look* on these and other historical examples shall confirm this claim. Also Carrier gives some arguments which seem to point into the direction of this claim. Carrier argues that if an outdated theory T was strongly successful in a certain domain, then it would be a miracle if there would not be *something* in T's superstructure which T has got right (2003, §8). But the question is in what this 'something' consists – and in this respect, Carrier's analysis seems to me not sufficient. Carrier argues that what the phlogiston theory and the caloric theory have got right is their ordering or *classification* of the phenomena (§8). But this description is not sufficiently clear – which phenomena does Carrier mean? For example, both the phlogiston theory and the caloric theory have got their classification of *substances* wrong, because according to presently accepted chemical theory, neither phlogiston nor caloric is an existing substance. In some passages, Carrier defends a more narrow claim, namely: what phlogiston and caloric theory have got right is the grouping of phenomena as being the result of a common underlying process. What is right in phlogiston theory according to Carrier is that the chemical process of dephlogistonation (oxidation) and phlogistonation (reduction) are identified as inverse reactions; and what is right in caloric theory is that thermal expansion is a common phenomenon of all gases. This is true, but it seems to be too weak. That chemical processes can be inverted was known independently from various chemical reactions, and that the thermal expansion of all (sufficiently ideal) gases is the same is a purely empirical consequence of caloric theory whose content does not depend on any theoretical assumptions at all. What we really are after is something in the *theoretical superstructure* of the phlogiston- and of the caloric theory which is right in the light of modern chemistry. We are looking for something X which from the view of presently accepted chemistry really exists and which corresponds to something in the theoretical superstructure of the phlogiston- or of the caloric-theory.

There is nothing which directly corresponds to "phlogiston" from the viewpoint of modern chemistry. But this is no wonder, because as we have already explained, phlogiston theory itself was not able to provide a general criterion of how phlogiston can be empirically identified. So the theoretical term "phlogiston" was empirically underdetermined in phlogiston theory. The theoretical expressions of phlogiston theory which did all the empirically relevant work and which were *not* empirically underdetermined were the expressions of *phlogistonation = assimilation of phlogiston*, and of *dephlogistonation = release of phlogiston*. For these two expression there is a perfect correspondence in modern chemistry which goes much farer than then identification of phlogistonation with oxidation in Lavoisier's sense. To explain this correspondence we need a bit more of modern chemistry.

Every substance consists of molecules, and molecules consist of atomic elements bound together by chemical bonds. The *electropositivity* of an element measures the tendency of an element to contribute electrons to its neighboring atoms in electrically polarized or ionic bonds.[43] Conversely, the electronegativity measures the tendency of an element to attract electrons from the neighboring atom in polarized or ionic bonds. The lower (higher) the column of an element in the periodic table, and the lower (higher) its row, the more electropositive (electronegative) is the element – with the exception of the elements of row 8, the inert gases. Metals (as well as hydrogenium) are typically electropositive, becoming more inert if their column and row becomes larger. Elements in column 4 and rows 2 and 3 (Carbonium and Silicium) are in the middle of the spectrum. Non-metals such as oxygen are typically electronegative, with the extremes being the elements of column 7, the halogens. Oxidation of an elementary substance X (a metal, coal, sulfur, etc.) in the *generalized sense* consists in the formation of a polarized or ionic bond of X, in which the atomic elements of X are electropositive and thus donate electrons to their electronegative neighbors in the bond. Every process of combustion, calcination or saltification consists of such an oxidation process. The inversion of the process of

[43] For the following cf., e.g., Oxtoby et al. (1999), ch. 3, ch. 6.3.

oxidation is called the process of reduction: here the polarized or ionic bond between an electropositive X-ion and its electronegative neighbor is broken, X regains its missing electrons and reappears in its pure elementary form. Therefore we have the following correspondence relations between phlogiston theory and modern chemistry:

Correspondence relations between phlogiston theory and modern chemistry:

Dephlogistonation of X corresponds to (and hence implicitly refers to) the donation of electrons of X-atoms to the bonding partner in the formation of a polarized or ionic chemical bond.

Phlogistonation of X corresponds to (and hence implicitly refers) to acceptance of electrons from the bonding partner by positively charged X-ions in the breaking of a polarized or ionic chemical bond.

What was wrong in phlogiston theory is that phlogiston was thought of as a special substance which is *emitted* during an dephlogistonation process. The electrons do not *leave* the chemical substance but just move a little bit to the electronegative neighbors in the molecule. What really is emitted as the end product of an oxidation process (besides the oxidized material) depends on the oxydans, that is, the input substance which causes the oxidation and which spends the electronegative partner. If the oxidans is an acid, then what is emitted is hydrogenium, whence in these cases phlogiston could be identified with hydrogenium. If the oxydans is pure oxygen and the oxidized material is coal, then carbon dioxide is emitted. In the combustion of phosphorus and sulfur nothing is emitted, and therefore the weight increases after dephlogistonation – these most problematic cases for the phlogiston theory. But apart from these cases, the phlogiston theory was strongly successful, and this strong empirical success is explained by the above correspondence relations.

In the case of caloric theory we enter again the situation that the term "caloric" was empirically underdetermined – an empirical identification of caloric was impossible. The crucial theoretical expression of caloric theory which did all the empirical work were *the amount of caloric particles* in a substance and *the repulsion forces between the caloric particles*. According to the presently accepted theory, what corresponds to the amount of caloric particles contained in a substance is the *mean kinetic energy* of its molecules. If this mean kinetic energy increases, the molecules start to rotate and oscillate and bang around, and if these movements get too strong to be compensated by the cohesion forces, then the substances converts into the gaseous state. In this state, the cohesion forces between the gas molecules are negligible as compared to the expansion forces, which are measured macroscopically in the *pressure* of the gas, and which corresponds in caloric theory to the repulsion forces between the caloric particles. So we have the following correspondence relation:

Correspondence relations between the caloric theory and modern physical chemistry:

The amount of caloric particles in a substance X = the mean kinetic energy of X's molecules

The repulsion force between the caloric particles in X = the expansion forces of X's molecules which in the gaseous state correspond to the pressure of X.

These correspondence relations explain the strong success of the caloric theory which correctly predicted that in the gaseous state the thermal expansion is the same for all gases.

Generally speaking, with a correspondence relation between theories T_1 and T_2 I mean a *synthetic* (non-analytic) statement which connects theoretical terms of the two theories in a *non-trivial* way, and which is 'somehow' obtained from considering the two theories together, although the way it is obtained is usually notoriously unclear. We have found correspondence relations for even such outdated theories as phlogiston theory and caloric theory. We can find similar correspondence relations in seemingly all other areas of strongly empirically successful theories. For example, the expression of absolute temperature in the phenomenological gas laws corresponds to the mean kinetic energy in the kinetic gas theory. Masses and velocities in Newtonian mechanics approximate the masses and velocities in special relativity theory when the velocity of the given physical particles become small relative to the velocity of light. The expectation values of the dynamic variables in quantum mechanics (position and momentum) obey the laws of classical mechanics (the theorem of Ehrenfels). And so on. For these reason I want to conjecture the following (bold) thesis:

Thesis: If a theory T has been strongly successful in a domain of applications A, but was superseded later on by an empirically superior theory T (which was likewise successful in domain A), then the empirically essential theoretical terms or expressions of T correspond implicitly to certain theoretical expressions of T*, and given T* is true, they refer implicitly to the entities denoted by these expressions of T*.*

If this thesis were true, then all Kuhn- and Hintikka-incommensurable theories which have equally strong empirical success in a certain domain A would be connected by these correspondence relations and, hence, would be theoretical commensurable in a sense which is different from Kuhn's or Hintikka's sense. Moreover, if this thesis were true, then Putnam's improved miracle argument would be valid. My historical analysis support this thesis, but opponents may consider this support as a rather weak one and my correspondence relations as rather accidental ones. Is there a way to establish the validity of my thesis in a systematic or even logical way? In the next section I will try to confirm my thesis by a general logical theorem.

5. THE CORRESPONDENCE THEOREM

I have argued in §1 that correspondence relations between two Kuhn- or Hintikka-incommensurable theories cannot be provided by logic alone, and they cannot be provided by theory-independent conceptual principles, because such principles do not exist. Also, non-trivial correspondence relations between T_1 and T_2 cannot be provided by one of the two theories alone, because the language of T_1 does not contain the theoretical terms of T_2, and vice versa. More precisely, for every statement S such that $B \cup T_1 \Vdash S$, all occurrences of T_2-theoretical predicates in S are uniformly replaceable in S by arbitrary other predicates, salva validitate of $B \cup T_1 \Vdash S$, which means that S is trivial correspondence relation.[44] So the only remaining possibility is that such correspondence relations can may obtained as consequences of the two theories joined together in their unified language. But how should it possible to unify the conceptual frameworks of two theories which are theoretically incommensurable?

At this point, Hintikka gives as an decisive hint: "there is in principle nothing that prevents a philosopher of science from pooling together the resources of the two conceptual frameworks" (1988, 229). This remark is important as it makes clear that from the logical viewpoint, two theories can be compared even if they do not possess any non-logical concept in common

[44] Lewis (1970, §VI) has argued that correspondence relations (he calls them bridge principles) can be deduced from one theory alone, but a closer look to his arguments shows that Lewis includes the definite description 'definitions' of the theoretical terms of the other theory among the premises. Lewis only asserts that correspondence relations may *possibly* be derived, but he does not show *how* and under *which conditions* they can be derived. The correspondence theorem of this section gives an answer this question.

(we assume that both languages share the same logical concepts). So let L_{12} be the united language of both theories (with languages L_1 and L_2); note that L_{12} is not just the union $L_1 \cup L_2$, but it includes in addition all *mixed* formula containing both concepts of L_1 and of L_2. For example, if T_1 is the phlogiston theory of combustion, and T_2 the oxygen theory, then L_{12} contains statements such as "phlogiston is different from oxygen", or "phlogiston is delivered iff electron are donated", etc.

Hintikka's point is absolutely crucial as a defense against the idea of global incommensurability which has become popular after Kuhn and Feyerabend. Not only is there nothing which prevents us from analyzing statements connecting the concepts of two Kuhn-incommensurable theories – analyzing such statements is even necessary if we are interested in finding correspondence relations between the two Kuhn-incommensurable theories. Let us first explain how both theories are needed in the process of deducing a correspondence relation at hand of a real example, the correspondence between absolute temperature and mean kinetic energy in statistical thermodynamics. One may object against this example that the phenomenological gas law is not really a theory because all of its terms are empirically measurable. This objection is not important because nothing in our example hangs on the empirical measurability of the absolute temperature (apart from that, the notion of absolute temperature involves theoretical assumptions concerning the fixation of the absolute zero point of temperature). We use the example only as a demonstration of how the derivation process of important correspondence relations in science makes use of *both* theories.

In 1811, when Dalton's atomic hypothesis was already well confirmed, Avogadro stated his famous gas law according to which the same volumina of a gas contain the same number of molecules. By one *mole* of an elementary substance one understands its molecular weight in gram, and one mole always contains the same number of molecules, namely $N_A = 6.023.10^{+23}$ molecules (Avogadro's number).[45] The volume of one mole of a sufficiently ideal gas at 1 atm pressure and 0^o Celsius gas is always 22,414 liter. In the light of these findings, the phenomenological gas law has been expressed around the middle of the 19^{th} century in the following form:

(1) $p \cdot V = n \cdot R \cdot T$

where p, V, T are the pressure, volume and absolute temperature of a gas, respectively, R is a constant (the Rydberg constant), and n is the number of moles in the gas, that is, the gas contains $n \cdot N_A$ molecules.

[45] For the following cf. Barrow (1966), Vol. 1, chs-1-2.

When the kinetic gas theory was developed by Maxwell and Boltzmann, the idea was to reduce this phenomenological regularity to the mechanical behaviour of the gas viewed as an swarm of myriads of tiny molecules which fly around and bounce against each other and against the walls of the gas container in a random way. Mole number and volume retain their meaning when passing from phenomenological gas law to the framework of mechanics, and also the pressure of a gas has a sharply defined mechanical meaning, although this meaning was a little bit harder to explicate. The pressure which is exerted orthogonally on an area is defined as the quotient of a force F and the area A, p = F/A. So the pressure which a gas exerts on the walls of its container is the sum of the forces which are exerted by its molecules in each time unit. By averaging over the squares of the velocities of the randomly moving molecules one obtains after certain calculations the following formula for the pressure of the gas (see Barrow 1966, ch. 2.2):

$$\text{(2)}\quad p = n \cdot \frac{2 \cdot N_A}{3 \cdot V} \cdot \frac{m \cdot \overline{v^2}}{2}\text{, hence}$$

$$p \cdot V = n \cdot \frac{2 \cdot N_A}{3} \cdot \frac{m \cdot \overline{v^2}}{2}$$

where $\overline{v^2}$ = the mean velocity of the gas molecules, m is the mass of one gas molecule; so

$$m \cdot \frac{m \cdot \overline{v^2}}{2}$$

is the mean kinetic energy of one gas molecules.

The crucial problem was to find a mechanical interpretation of the concept of (absolute) temperature. Phenomenologically, temperature corresponds to a special sensory quality, which is qualitatively different from any mechanical sense experience. So for temperature no mechanical account can be given independently from the gas law: the phenomenological concept of temperature is *incommensurable* with mechanical concepts, both in the Kuhn- and in the Hintikka-sense. At this point, the *uniting* of the two theories, the phenomenological gas law (1) and the consequence (2) of kinetic gas theory, becomes essential: from (1) and (2) together the following correspondence relation follows:

$$(3)\ T = \frac{2 \cdot N_A}{3 \cdot R} \cdot \frac{m \cdot \overline{v^2}}{2}$$

in words: the absolute temperature of a gas is directly proportional to the mean kinetic energy of the gas molecules.

This fundamental correspondence relation between two seemingly incommensurable concepts was established by drawing a consequence from the union of both theories.

I want to show now by general logical means that there is a systematic reason behind this possibility: under certain conditions it is always possible to derive a correspondence relation between two theories which have equally strong empirical success in a common domain of applications A. I assume that the theoretical term φ is empirically characterized within T by means of Carnapian bilateral reduction sentences of the form

$$\text{(BR)}\quad A_i \rightarrow (\varphi[x] \leftrightarrow R_i) \qquad \text{where } V_f(x) \subseteq V_f(R_i),\ V_f(x) \subseteq V_f(A_i)$$

in words: under empirical circumstances A_i, the presence of φ is indicated or measured by an empirical phenomenon or process R_i.

The formulas A_i and R_i are possibly complex empirical (non-T-theoretical) formulas. They contain variables which are understood to be universally quantified and which obey the condition at the right side ($V_f(x)$ = the set of variables occurring free in x). With x I mean a vector of variables describing the given individual or system (for example, time and number of particles, etc.). So φ abbreviates an atomic formula containing the T-theoretical term φ and the variables in x (it may also contain variables r over real numbers). I will frequently omit the variables and mention them only when I want to speak explicitly about the individual or system x to which certain claims apply.[46]

[46] Certain subtleties are involved concerning the variable t for time. Since I do not want to distract the reader, I mention them only in the footnote. If φ is an intrinsic property of x which does not depend on time, then the bilateral reduction sentences has the explicit form $\forall x,t$: $A_ixt \rightarrow (\varphi(x) \leftrightarrow R_ixt)$. This sentence is empirically creative insofar it implies $\forall x$: $\exists t(A_ixt \wedge R_ixt) \rightarrow \forall t(A_ixt \rightarrow R_ixt)$. This is *okay*: if a certain behaviour R_ixt under circumstances A_ixt (such as a certain chemical reaction) empirically indicates an intrinsic

In a modern understanding (different from Carnap) bilateral reduction sentences are conditions of empirical identification or measurement of the theoretical entity or process φ under special conditions A_i. They are usually not part of T's axiomatization but are obtained as logical consequences of T and the background knowledge B. For example, if φ stands for "x delivers phlogiston", and A_i stands for "x is a metal which is put into hydrochloric acid", then R_i stands for "x dissolves in the acid and inflammable air evaporates". I wish to emphasize that the characterization of theoretical terms by bilateral reduction sentences in this general sense covers almost all important kinds of statements in which a theoretical term is empirically identified or measured. In particular, it covers all quantitative measurement laws for theoretical terms. For example, if φ stands for the formula "mass(x) = n gram" and A_i for the circumstance "x is put on a balanced beam scales", then R_i stands for the formula "the number of one gram units on the other side of the balanced beam is n". The formula $A_i(x) \rightarrow \forall r(m(x)=n \leftrightarrow bal(x)=n)$ is logically equivalent with the conditional identity statement $A_i(x) \rightarrow m(x)=bal(x)$, which is the standard form of a theoretical measurement law.

It is a crucial feature of an empirically successful theory − as opposed to a simple law of disposition − that its theoretical terms are characterized by *several different* bilateral reduction statements of the above form. Therefore I assume that the domain A of the theory T consists of several special kinds of applications, described by empirical circumstances $A_1,\ldots,A_n$, in which φ is characterized by special bilateral reduction sentences. For example, A_1 may be the exposition of a metal to air and water, A_2 may be the exposition of a metal to hydrochloric acid, etc. From these bilateral reduction sentences, empirical consequences or the form $A_i \wedge R_i \rightarrow (A_j \rightarrow R_j)$ are deducible by which one can infer something from what has happened in one domain of application about what will happen in another domain of application.

If the empirical success of T should count as a *strong* empirical success, then the subdomains A_i must be *qualitatively different* from each other: from empirical descriptions of what goes on in a system x in domain A_i nothing should be follow by *empirical induction alone* about what goes on in system x in domain A_j; such an inference should only be possible by passing to the theoretical description of the system x under consideration. To reflect this qualitative difference of the subdomains, I shall assume that

time-independent property of x, then it must be expected that x will exhibit this behavior every time when it is put into circumstances A_i. This subtlety causes further subtleties discussed in the next footnotes.

whenever a theory T implies a conditional which asserts that whenever the system x has behaved in a certain way Q_i in domain A_i, then it would behave in another way Q_j in the different domain A_j, then this implication can only hold because it is possible in T to conclude from $A_ix \wedge Q_ix$ a certain theoretical description $\tau(x)$ of x about the intrinsic properties of x from which again it can be concluded in T that $(A_jx \rightarrow Q_jx)$ will hold. More formally, whenever T $\Vdash$ $(A_ix \wedge Q_ix \rightarrow (A_jx \rightarrow Q_jx)$, then there will exist a theoretical 'interpoland' $\tau(x)$ such that T $\Vdash$ $(A_ix \wedge Q_ix \rightarrow \tau(x))$ and T $\Vdash$ $(\tau(x) \rightarrow (A_jx \rightarrow Q_jx))$.[47] If a theory T has this property w.r.t. a partition $\{A_i{:}1 \leq i \leq n\}$ of a domain A into subdomains, then I say that the unification of A by T is *essentially dependent on T*. This condition is rather natural: for example, by observing reactions of metals in hydrochloric acid nothing can be concluded by empirically-inductive means about the behaviour of metals in oxygen or water, or about the inversion of the process by passing hydrogenium into the metal-salt-solution. Likewise, by observing planets moving around the horizon nothing can be concluded by empirically-inductive means about the movement of projectiles on the earth. All connections of this sort are provided by the theory.

A third condition which I impose on the theories is that they are *causally normal* in the following sense. The non-theoretical (empirical) predicates or parameters of the theory divide into a set of causally independent parameters which describe the circumstances A_i of the special subdomains, and a set of causally dependent parameters. The behaviour of the system x w.r.t. its dependent parameters under given independent parameters can be deduced from the theory. But it is impossible to derive from a purely theoretical description $\tau(x)$ of a system x any empirical assertion about the status of the independent parameters of x. This is again a very natural condition. For example, nothing can be concluding from the theoretical nature of a certain substance about what humans do with it, about whether they expose it to hydrochloric acid or to heat or whatever. Nothing can be concluded from a purely mechanical description of a physical body about its initial conditions, about whether the body is thrown into the air, put

47 If $\varphi(x)$ is a time-independent property, then these empirical consequences have the exact form: $\forall x{:}\ \exists t(A_ixt \wedge Q_ixt) \rightarrow \forall t(A_jxt \rightarrow Q_jxt)$, saying that whenever x has behaved at *some* time in subdomain A_i in the way R_i, then every time x is put into subdomain A_j it will behave in the way Q_j. The interpolation statements following from T then have the form T $\Vdash$ $\forall x{:}\ (\exists t(A_ixt \wedge Q_ixt) \rightarrow \tau(x))$ and T $\Vdash$ $\forall x{:}(\tau(x) \rightarrow \forall t(A_jxt \rightarrow Q_jxt))$. But these statements are logically equivalent with $\forall x,t{:}\ A_ixt \wedge Q_ixt \rightarrow \tau(x)$, and $\forall x,t{:}\ \tau(x) \rightarrow \forall t(A_jxt \rightarrow Q_jxt)$. Therefore it is possible to represent these statements without quantifiers, which means that they are universally quantified.

under pressure or split into pieces. I summarize my three conditions in the following definition.

Definition:

1) A T-theoretical term or expression φ of a theory T with partitioned domain $A = A_1 \cup \ldots \cup A_n$ domain A is *empirically successful* iff $n \geq 2$ and for every subdomain A_i the theory implies a bilateral reduction sentence of the form $A_i \rightarrow (\varphi[x] \leftrightarrow R_i)$, which says in words*:* if a system x is exposed to the circumstances of subdomain A_i, then x has theoretical property φ iff it x's behaviour satisfies the empirical description R_i.

2) The unification of the partitioned domain $A = A_1 \cup \ldots \cup A_n$ by a theory T is essentially T-dependent iff for every formula of the form $(A_i \wedge Q_i \rightarrow (A_j \rightarrow Q_j)$ following from T, there exists a theoretical description $\tau[x]$ of the underlying system x such that $(A_i \wedge Q_i \rightarrow \tau[x])$ and $\tau[x] \rightarrow (A_j \rightarrow Q_j)$ follow from T (cf. fn. 9).

3) A theory T is *causally normal* w.r.t. its partitioned domain $A = A_1 \cup \ldots \cup A_n$ iff (i) the non-T-theoretical (empirical) vocabulary of T divides into a set of independent parameters (predicates or function terms) ID and a set of dependent parameters D, (ii) the descriptions "$A_i x$" of the subdomains A_i are formulated solely by means of the independent parameters in ID (plus logico-mathematical symbols), and (iii) no non-trivial claim about the state of the independent parameters of a system x can be derived in T from a purely T-theoretical and T-consistent description of x.

Correspondence theorem: Let T be a theory with partitioned domain $A = A_1 \cup \ldots \cup A_n$ which is causally normal w.r.t. this partitioned domain and which contains a T-theoretical term φ which is empirically successful w.r.t. this partitioned domain. Let T* be a successor theory of T which is likewise causally normal w.r.t. this partitioned domain and which has at least the same empirical success in domain A as T has, such that the unification of the partitioned domain obtained by T* is essentially T*-dependent. (T's theoretical terms need not stand in any logical or conceptual relation to T*s theoretical terms.) Then: T and T* together imply a correspondence relation of the form

$A \rightarrow (\varphi\,[x] \leftrightarrow \tau^*[x])$

in words: whenever a system x is exposed to the circumstances in one of the sub- domains of A, then x has the T-theoretical property φ iff x satisfies the (possibly complex) T*-theoretical description τ^*

which implies that $\varphi(x)$ implicitly refers to the theoretical state of affairs described by described by $\tau^*(x)$, provided T* is true.

Remark*:* The correspondence theorem applies, of course, to all empirically successful T-theoretical terms of T. The condition of empirical success of φ must be imposed on T, which means that the T-theoretical term or expression φ of T − such as "addition of "phlogiston" or "amount of caloric" − can be empirically identified in various empirical circumstances. That T* is empirically as successful as T in domain A means simply that every empirical consequence of T about what goes on in one of T's subdomains A_i must also be an empirical consequence of T* Both theories must be causally normal. The condition of theory-dependent unification must be imposed on the theory T*. However, the condition of theory-dependent unification for T* implies that the subdomains $A_1,\ldots,A_n$ are qualitatively different so that nothing can be transferred from one subdomain to the other by empirically-inductive means. This implies in turn that T's empirical success w.r.t. the subdomains $A_1,\ldots,A_n$ must be a *strong* empirical success in Carrier's sense.

Proof*:* We will explain the following proof in an illustrative way, so that the reader gets an impression of what goes on and how the conditions on T and T* come into play.

(1.) φ in T is empirically successful w.r.t. a partition $A_1,\ldots,A_n$ of A ($n\geq2$), and T is causally normal w.r.t. this partition; so it must hold that

(Red) $\forall i\in\{1,\ldots,n\}$: $T \Vdash\!\!\!- A_i \rightarrow (\varphi[x] \leftrightarrow R_i)$

where the A_i's are expressed in terms of independent parameters and the R_i's in terms of dependent parameters of T.

(2.) The empirical consequences of T following from these bilateral reduction sentences in (1) must also be consequences of T*, because T* is equally empirically successful in domain A as T, and so the following must hold (but cf. fn. 9):

$\forall i\neq j \in \{1,\ldots,n\}$: $T^* \Vdash\!\!\!- A_i\wedge R_i \rightarrow (A_j\rightarrow R_j)$, and

$T^* \Vdash\!\!\!- A_i\wedge\neg R_i \rightarrow (A_j\rightarrow\neg R_j)$.

These are $2\cdot n\cdot(n-1)$ empirical consequences which T* must imply.

(3.) Because the unification of A yielded by T* is essentially T*-dependent, (2.) gives us the following (cf. again fn. 9: the interpolating formulas can be understood as universally quantified; so the proof remains essentially propositional):

For every i there must exist theoretical T*-descriptions $\tau_i(x)$ and $\mu_i(x)$ such that

$T^* \Vdash\!\!\!- A_i\wedge R_i \rightarrow \tau_i{}^*[x]$, and $\forall j\neq i\in\{1,\ldots,n\}$: $T^* \Vdash\!\!\!- \tau_i{}^*([x] \rightarrow (A_j \rightarrow R_j)$, and

$T^* \Vdash A_i \wedge \neg R_i \rightarrow \mu_i^*[x]$, and $\forall j \neq i \in \{1,\ldots,n\}$: $T^* \Vdash \mu_i^*[x] \rightarrow (A_j \rightarrow \neg R_j)$.

It is sufficient for our purpose to choose just one fixed i. So we put $\tau_i^* = \tau^*]$ and $\mu_i^* = \mu^*$. We will show below that the τ_i^*'s and μ_i^*'s must be equivalent for all $1 \leq i \leq n$.

We can see why the condition of theory-dependent unification is crucial. Without this condition it would be impossible to say *how* the empirical consequences in (2.) are obtained from T*. Each of this $2 \cdot n \cdot (n-1)$ consequences could be obtained in a completely different way. With this condition we can 'interpolate' two T*-theoretical descriptions of x, one interpolates the empirical consequences involving positive R_i's and R_j's , and the other one interpolates the empirical consequences involving negative R_i's and R_j's.

(4.) From (3.) it follows by propositional logic that (for fixed $i \in \{1,\ldots,n\}$:

$T^* \Vdash (A_i \rightarrow (R_i \rightarrow \tau^*[x])$ $T^* \Vdash A_j \rightarrow (\tau^*[x] \rightarrow R_j)$ $(\forall j \neq i \in \{1,\ldots,n\}$

$T^* \Vdash (A_i \rightarrow (\neg R_i \rightarrow \mu^*[x])$ $T^* \Vdash A_j \rightarrow (\mu^*[x] \rightarrow \neg R_j)$ $(\forall j \neq i \in \{1,\ldots,n\}$

What we want is, of course, the derivation of bilateral reduction sentences for T*-theoretical descriptions which have the same form as the bilateral reduction sentences for φ. From this we could derive, by joining the two theories together, the intended correspondence relation between φ and the T*-theoretical descriptions (similar as in the derivation of the correspondence between absolute temperature and mean kinetic energy, but in a logically general way). The problem is that so far we have *two different* T*-theoretical descriptions, τ and μ. Now the crucial condition (iii) of causal normality comes into play.

(5.) We prove $T^* \Vdash (\tau^*[x] \rightarrow \neg\mu^*[x])$, thereby using condition (iii) of the causal normality of T*. It follows from the right side of (4.) by propositional logic that

$T^* \Vdash (\tau^*[x] \wedge \mu^*[x]) \rightarrow \neg A_j (\forall j \neq i \in \{1,\ldots,n\}$

But since T* is causally normal w.r.t. the given partition of A, A_j is an assertion described in terms of independent parameters, and so it follows that $(\tau^*[x] \wedge \mu^*[x])$ must be T*-inconsistent, for otherwise T* could not be causally normal. Therefore we have

$T^* \Vdash (\tau^*[x] \rightarrow \neg\mu^*[x])$.

This gives us together with (4.) the following:

$T^* \Vdash (A_i \rightarrow (R_i \leftrightarrow \tau^*[x])$ $T^* \Vdash A_j \rightarrow (\tau^*[x] \leftrightarrow R_j)$ $(\forall j \neq i \in \{1,\ldots,n\}$

Therefore we have

(Red*) $\forall i \in \{1,\ldots,n\}$: $T^* \Vdash A_i \rightarrow (\tau^*[x] \leftrightarrow R_i)$

(6.) Now we have reached the situation which we have wanted. For if we join the consequence set of T in (Red) and the consequence set of T* in (Red*), we can derive from both by propositional logic the intended correspondence relation:

$\forall i \in \{1,\ldots,n\}$: $T \cup T^* \Vdash A_i \rightarrow (\varphi[x] \leftrightarrow \tau^*[x])$

and hence

(Corr) $T \cup T^* \Vdash A \rightarrow (\varphi[x] \leftrightarrow \tau^*[x])$.

(7.) We also emphasize that in the derivation of the correspondence relation we have used only the bilateral reduction sentences Red(T) following from (T). We can sharpen our assertion as follows:

(Corr*) $\mathrm{Red}(T) \cup T^* \Vdash A \rightarrow (\varphi[x] \leftrightarrow \tau^*[x])$.

This means that T may contain some assertions in its theoretical superstructure which are even inconsistent with T*. What we only need to assume is that Red(T) is logically consistent with T*.

(8.) Concerning the uniqueness of the T*-theoretical expression $\tau^*[x]$ module equivalence: Let us again put $\tau^*[x] = \tau_j^*[x]$. (Red*) can be derived for all j ($1 \leq j \leq n$). So it holds:

$\forall j \in [1,\ldots,n\}$, and $\forall i \in \{1,\ldots,n\}$: $T^* \Vdash A_i \rightarrow (\tau_j^*[x] \leftrightarrow R_i)$.

Hence by propositional logic:

$\forall j \neq k \in [1,\ldots,n\}$, and $\forall i \in \{1,\ldots,n\}$: $T^* \Vdash A_i \rightarrow (\tau_j^*[x] \leftrightarrow \tau_k^*[x])$

Therefore by propositional logic:

$\forall j \neq k \in [1,\ldots,n\}$: $T^* \Vdash A \rightarrow (\tau_j^*[x] \leftrightarrow \tau_k^*[x])$

which means that T* entails that all the theoretical descriptions $\tau_i^*[x]$ ($1 \leq i \leq n$) are equivalent in the domain A. So it is natural to assume that T* identifies all of them.

End of the proof.

Let me apply this correspondence theorem to our two historical examples. The caloric theory implies bilateral reduction sentences of the form

If x is a gas with mole number n under pressure p, then the amount of caloric in it is directly proportional to its volume.

Given that kinetic gas theory explains all the (confirmed) empirical consequences which follow from these bilateral reduction sentences for various gases, mole numbers, and pressures, which constitute the different subdomains, then our correspondence theorem tells us that both theories together must imply a correspondence relation which says:

If x is a gas, then the amount of caloric in it is proportional to a certain functional expression in terms of the kinetic gas theory

– and in fact, we have found such a function al expression: the mean kinetic energy of the gas particles.

The phlogiston theory implies bilateral reduction sentences of the form

If x is a substance of kind X_i (e.g., a metal), and x is exposed to the influence of certain chemical input substances Y_i (e.g., solution in hydrochloric acid)$_i$, then x gets dephlogistonated iff the chemical reaction outputs Z_i (e.g., iff it dissolves and inflammable air evaporates).

Given that modern chemical oxidation and reduction theory implies all the (confirmed) empirical consequences of these bilateral reduction sentences then again, the correspondence theorem says that the two theories together must imply a correspondence relation saying:

For all substances x in chemical reactions y of the domain of applications: x is dephlogistonated during reaction y iff during reaction y, x satisfies a theoretical description in terms of modern chemical oxidation and reduction theory

– and we have found such a description: oxidation in the generalized sense. The same must hold for phlogistonation and reduction.

Let me summarize what I hope to have achieved in the preceding investigation:

(1.) I have shown that if a theory T was empirically successful in a domain A, and this empirical success has involved essentially a theoretical term φ, and the theory T was superseded later on by a theory T* which is Kuhn- and Hintikka-incommensurable with T and at least equally successful in domain A, and if both theories satisfy some plausible conditions, then from both theories together a correspondence relation will be derivable which asserts that in domain A, the presence of φ in a system x corresponds to the satisfaction of a certain T*-theoretical description $\tau*$ of the system x. The conditions on the theory T* imply that the domain A must split into qualitatively different subdomains, which means that the empirical success of T must have been a strong success of T.

(2.) Thereby I have shown that under the given conditions, the equally strong empirical success of two theories T and T* in a domain A implies a certain commensurability between the two theories, in form of correspondence relations between T's theoretical terms and T*-theoretical descriptions, *even when* the two theories are Kuhn- and Hintikka-incommensurable.

(3.) I have motivated substantiated the correspondence theorem in (1.) by various historical examples.

(4.) If the theory T* is true, (1.) implies that the T-theoretical terms must have a certain implicit reference – implicit because the nature of this reference is usually not grasped by the proponents of T. But this implicit reference of T's theoretical terms explains the strong empirical success of T. Of course, a theory of the kind T* may not always be known. But one assume that it always *exists*, in form of the "true" theory about the physical nature of the world. Therefore, the improved version of Putnam's miracle argument can be strengthened as follows: *that a theory is strongly empirically successful although its theoretical terms do not refer, even not implicitly, is not only very improbable: under certain plausible conditions it is even impossible*.

REFERENCES

Barrow, G. M. (1966): *Physical Chemistry*, McGraw-Hill, New York.

Carnap, R. (1956): "The Methodological Character of Theoretical Concepts", in: Feigl, H./ Scriven, M. (Hg.), *Minnesota Studies in the Philosophy of Science Vol. I*, Univ. of Minnesota Press, Minneapolis, 38-76.

Carrier, M. (2003): "Experimental Success and the Revelation of Reality: The Miracle Argument for Scientific realism", in P. Blanchard et al. (eds.), *Science, Society and Reality*, Springer, Heidelberg.

Feyerabend, P. (1962). "Explanation, Reduction, and Empiricism, in H. Feigl, G. Maxwell (eds.), *Minnesota Studies in the Philosophy of Science Vol. 3*, Univ. of Minnesota Press, Minneapolis.

Hempel, C. G. (1951): "The Concept of Cognitive Significance: A Reconsideration", reprinted in Hempel 1965, ch. II.1.

Hempel, C.G. (1958): "The Theoretican's Dilemma. Studies in the Logic of Theory Construction", reprinted in: Hempel 1965, 173-228.

Hempel, C.G. (1965): *Aspects of Scientific Explanation and Other Essays in the Philosophy of Science*, New York-London.

Hintikka, J. (1988): "On the Incommensurability of Theories", *Philosophy of Science* 55, reprinted in Hintikka (1999), 227-240 (quoted from there).

Hintikka, J. (1999): *Inquiry as Inquiry: A Logic of Scientific Discovery*, Kluwer, Dordrecht.

Kuhn, T.S. (1962): *The Structure of Scientific Revolutions*, Univ. of Chicago Press, Chicago.

Lewis, D. (1970): "How to define Theoretical Terms", *Journal of Philosophy* 67, 427-446.

Oxtoby, D. et al. (1999): *Principles of Modern Chemistry*, Saunders College Publishing, Orlando.

Papineau, D. (1996): "Theory-dependent Terms", *Philosophy of Science* 63, 1-20.

Putnam, H. (1975): "What is Mathematical Truth?", in H. Putnam, *Mathematics, Matter and Method*, Cambridge Univ. Press, Cambridge, 60-78.

Quine, W.v. O. (1960): *Word and Object*, MIT Press, Cambridge/Mass.

Quine, W.v.O. (1992): *Pursuit to Truth*, Harvard Univ. Press, Cambridge/Mass.

Schurz, G. (1991): "Relevant Deduction", *Erkenntnis* 35, 391 - 437.

Schurz, G. (2004): "Semantic Holism and (Non)-Compositionality in Scientific Theories", to appear in: M. Werning et al. (eds.), *Proceedings of the Cococo-Conference*.

Sneed, J.D. (1971): The Logical Structure of Mathematical Physics, Reidel, Dordrecht.

WHAT IS ABDUCTION? AN ASSESSMENT OF JAAKKO HINTIKKA'S CONCEPTION

James H. Fetzer
Department of Philosophy
University of Minnesota

Few nations of the world have a more distinguished philosophical tradition than Finland, and few philosophers have attained the distinction of Jaakko Hintikka. His research displays a breadth of interest and a depth of analysis that brings a sense of admiration tinged with envy to most members of the profession. It would be asking too much of anyone to attempt to encompass the complete content of his voluminous publications. Indeed, I shall make no such effort here but shall focus instead on one of his books, *Inquiry as Inquiry: A Logic of Scientific Discovery* (1999), and from it one of his articles, "What is Abduction?" (1998), and its primary theme, namely: that abduction cannot be adequately understood as inference to the best explanation, but that, when taken as a method for providing answers to an inquirer's questions, it can fit comfortably within Hintikka's own interrogative model of knowledge acquisition.

There appear to be at least two crucial issues here. One is what Charles S. Peirce had in mind when he introduced "abduction" as a mode of inference comparable to"deduction" and "induction". This is an historical question. The second is whether "abduction" can be adequately explicated as "inference to the best explanation". This is a philosophical question. While there are passages in Peirce that tend to support Hintikka's position and Peirce makes some misleading claims, Hintikka's contention that abduction in Peirce's sense appears to be no more than "a mysterious power of guessing right" cannot be sustained. The notion of "abduction", however, turns out to be ambiguous. "Abduction" in its narrow sense entails "a capacity for guessing" that is an indispensable element of "abduction" in its broader sense, which defines a procedure for "guessing right" by means of inference to the best explanation and raises some intriguing questions about Hintikka's interrogative theory of knowledge.

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 127-155.

1. THE HISTORICAL QUESTION.

As Ilkka Niiniluoto, another distinguished Finnish philosopher, has noted, Peirce drew distinctions between three broad classes of reasoning in 1865, proposing that there is "a large class of reasonings" that are neither deductive nor inductive but which involve inferences from effects to causes, which Peirce called "hypothesis" (Niiniluoto 1999). In "Deduction, Induction, and Hypothesis" (1878a, CP 2.623), he illustrates their differences in relation to a syllogism of the form Barbara, where *deduction* is an inference from a rule and a case to a result, *induction* from a case and a result to a rule, and *hypothesis* from the rule and the result to the case, thus:

(I) DEDUCTION:	(P1) All the beans in this bag are white (rule)
	(P2) These beans are from this bag (case)
	(C1) These beans are white (result)
(II) INDUCTION:	(P3) These beans are from this bag (case)
	(P4) These beans are white (result)
	(C2) All the beans in this bag are white (rule)
(III) ABDUCTION:	(P5) All the beans in this bag are white (rule)
	(P6) These beans are white (result)
	(C3) These beans are from this bag (case)

Peirce maintains that, while all inference may be reduced to Barbara, it is not necessarily the most appropriate form in which to represent the distinctive properties of each, where the following schematic diagrams may be helpful in drawing out some the important features distinguishing each from the others.

Consider, for example, the following schematic characterization of deduction,

(IV) DEDUCTION: (P1) All As are Bs (rule)

(P2) This is an A (case)

―――――――――――――――

(C1) This is a B (result)

where the single line between premises and conclusion represents (the claim that) this inference is deductive (in the semantical sense), which will be the case when the conclusion cannot be false if the premises are true and (in the syntactical sense) when the conclusion follows from the premises in accordance with acceptable rules.

And consider additionally the following schematic characterization of induction,

(V) INDUCTION: (P3) This is an A (case)

(P4) This is a B (result)

═══════════════

(C2) All As are Bs (rule)

where the double line between premises and conclusion indicates (the claim that) this inference is inductive (in the semantical sense), which will be the case when, although its conclusion has more content than its premises and can be false even when those premises are true, the conclusion receives some degree of evidential support from those premises (in the syntactical sense) on the basis of acceptable rules of inductive inference, none of which is a truth-preserving deductive rule.

This generates the fundamental problem of inductive logic, namely: which of the infinity of non-truth-preserving rules qualify as acceptable rules of inductive inference? The pattern Peirce offers to exemplify abduction, after all, appears to be a candidate, since it is an obvious example of affirming the consequent, which, were it embraced as an acceptable rule, would exemplify the witticism that, in texts on logic, the logical fallacies are explained in their first part (on deduction) and are committed in their second (on induction). Here then is the analogous schematization,

(VI) ABDUCTION: (P5) All As are Bs (rule)

(P6) This is a B (result)
\- - - - - - - - - - - - - - - - -
(C3) This is an A (case)

where the single broken line means that the character of the inference is neither deductive nor inductive. Without begging the question by assuming the meaning of abduction, it should at least be apparent that Peirce has a formal justification in the three distinct arrangements of rules, cases, and results, which provides an objective foundation for distinguishing abduction from induction and deduction.

2. CHARACTERIZING ABDUCTION.

The question becomes whether there are any objective rules that govern the form of inference known as "abduction". Hintikka expresses profound skepticism as to whether abduction should even be considered to be a form of "inference". He advances a rather more complex version of schema (VI), which he offers as follows:

(VII) ABDUCTION:
(P7) All As which are B are C (rule)

(P8) This A is a C (result)

(C4) Therefore, this A is a B (case)

And, indeed, in Peirce's own characterizations in this article, he uses the "therefore" sign in designating conclusions for "deduction", "induction", and "hypothesis", alike.

This not only justifies Hintikka's schema (VII) and suggests that schemata (I) to (VI) are incomplete by virtue of the omission of "therefore" but offers prima facie motivation for a variety of critical remarks that revolve about its usage as follows:

> But in such a scheme an abductive inference does not necessarily yield even probabilistic support for its conclusion. Hence it is extremely puzzling why Peirce should have claimed that abduction in his sense was an inference… This puzzle about what abduction really is deepened by

> Peirce's explicit acknowledgment of the presence of a conjectural element in abduction. It even seems that our abductive hypothesis-forming power is nothing but a mysterious power of guessing right. (Hintikka 1998, p. 92)

Hintikka's objection, it appears, is that the conclusion of an inference ought to be detachable and independently assertable, which may or may not be the case for inductions but is certainly not the case for abduction, thus understood. Indeed, from this point of view, Peirce's use of the "therefore" sign is puzzling in itself.

My introduction of the single line, the double line, and the broken line, are interpretive and subject to argument. No one, I suppose, would be inclined to deny that the difference between inductive and deductive inferences is such that the use of a double line and of a single line, respectively, can be justified on several grounds, including that inductive arguments are non-demonstrative, ampliative, and non-additive, whereas deductive arguments are demonstrative, non-ampliative, and additive (Fetzer 1993a, pp. 106-110). It is a very good idea, but not common practice, to call inductive arguments with the right form *proper* and deductive arguments with the right form *valid*. Proper arguments with true premises are then called *correct* and valid arguments with true premises *sound*.

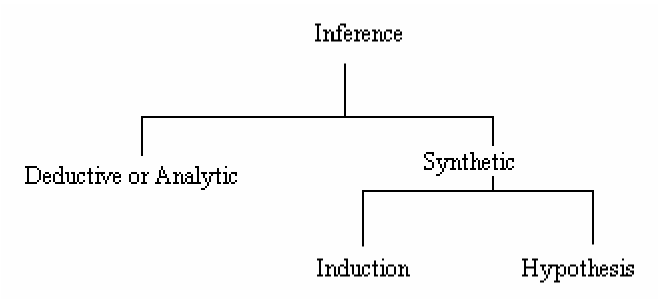
3. INDUCTION VS. HYPOTHESIS.

Similarly, the use of the terms "argument" and "inference" are not completely standard. Some prefer to restrict the term "argument" exclusively to special kinds of linguistic entities consisting of sets of sentences divided into two parts, premises and conclusions, where premise indicators (such as "given", "assuming", and such) and conclusion indicators (such as "therefore" and "consequently", in the deductive case; "probably" and "likely", in the inductive case) indicate how those arguments are to be taken, namely: the intended kind of support that the given premises are supposed to provide to their conclusions. And the term "inference" can be used to refer to the psychological process of drawing conclusions from premises, no matter whether they are proper or valid, correct or sound. Since the use of the phrase, "rules of inference", is perfectly common, here I shall not draw such a distinction.

Peirce distinguishes between "deductive or analytic" inferences, on the one hand, and two forms of synthetic inference, "induction and hypothesis", on the other.

He offers an explicit diagram in support, which presents his classification as follows:

(VIII)



remarking, by way of partial explanation, that induction includes inferences from samples to populations and that hypothesis includes reasoning by analogy, namely: Induction is where we generalize from a number of cases of which something is true, and infer that the same thing is true of a whole class. Or, when we find a certain thing to be true of a certain proportion of cases and infer that it is true of the same proportion of the whole class. Hypothesis is where we find some very curious circumstance, which would be explained by the supposition that it was a case of a certain general rule, and thereupon adopt that supposition. Or, where we find that, in certain respects two objects have a strong resemblance, and infer that they resemble one another strongly in other respects.(CP 2.624) Peirce commits a blunder here, alas, which supports Hintikka's objection that the method of abduction is no more than some "mysterious power of guessing right".

In particular, in asserting that, in hypothesis, having discovered a supposition which, if true, would explain some curious circumstance, we "thereupon" adopt it, Peirce implies that every hypothesis that might explain a puzzling phenomenon is, on that basis alone, an acceptable hypothesis that can be detached and adopted as true. This, of course, cannot possibly be correct, since there are many inconsistent potential explanations of a person's death, for example, where the discovery of a body is a curious circumstance, where that she has been shot (stabbed, strangled, and so on) might explain it does not warrant detachment and acceptance. Usually Peirce is more circumspect, especially about its application to various conjectures to ascertain which of them provides a possible explanation for the phenomenon.

4. MODAL DIFFERENCES.

Returning to Hintikka's concern about the use of the term "inference" in the sense in which an inference yields a detachable conclusion, there appear to be important modal distinctions between deductions, inductions, and abductions, which sheds some light on this puzzle and might help to clarify the situation. The conclusions of valid deductions, for example, are "certain" in relation to their premises. It thus appears to be very plausible to schematize deductions (of the very basic forms under consideration within this context) as follows:

(IX) DEDUCTION: (P1) All As are Bs (rule)

(P2) This is an A (case)
____________________ [certainly]

(C1) This is a B (result)

where the use of the word "certainly" implies that the conclusion must be the case, provided, of course, that the specified premises themselves are both true. And it appears equally plausible to schematize inductions of his form as follows:

(X) INDUCTION: (P3) This is an A (case)

(P4) This is a B (result)
==================== [probably]

(C2) All As are Bs (rule)

assuming the premises are both true, but where the strength of the support could range anywhere from zero to one on a hypothetical quantitative scale. And it appears very plausible to schematize abductions in this context thus:

(XI) ABDUCTION: (P5) All As are Bs (rule)

(P6) This is a B (result)
---------------------------- [possibly]
(C3) This is an A (case)

since the most that should be said is that we thereupon adopt it as possibly true. But there is an alternative schematization that addresses Hintikka's concerns more directly, where "possibly" characterizes the conclusion itself rather than the strength of the relationship between the premises and the conclusion as follows:

(XII) ABDUCTION: (P5) All As are Bs (rule)

(P6) This is a B (result)

(C3) Possibly, this is an A (case)

On this interpretation, the conclusion (C3) might be regarded as detachable and assertable, which would tend to satisfy Hintikka's desiderata for an "inference".

So, if (XII) is a defensible construction, perhaps his objection can be overcome. Others will observe that the difference between (XI) and (XII) is very much like one addressed by Carl G. Hempel (1965), who argued, relative to deductive arguments of form (IX) and inductive arguments of form (X), that "certainly" and "probably" qualify the strength of the support for the conclusion which is provided by the premises without making assertions about the modal status of the conclusion itself, where, "Certainly, this is a B", for example, would be true only if "This is a B" were a necessary truth as an analytic sentence. If we take "certainty", "probably", and "possibly" as standing for the strength of support provided the conclusion by the given premises, therefore, and not as qualifying the conclusion itself as a logical possibility, for example, then the problem that Hintikka has raised arises again. These conclusions do not appear detachable.

5. PROBABILITIES AND POSSIBILITIES.

Notice, in particular, this problem clearly arises with inductive arguments of form (X) whose probabilities are low. Peirce's example illustrates this point, since ordinarily a single case of an A that happens to be a B would not provide much support for the inference that all As are Bs, unless the sample exhausted the population because it had only a single member. The question becomes one of how many members must belong to a sample for it to offer sufficient support to justify detaching the conclusion. Presumably, Hintikka is not questioning the logical status of induction in Peirce's sense because he would acknowledge there are acceptable rules of inference for assigning specific degrees of probability to specific conclusions within the framework of inductive logic. What he doubts is the prospect of acceptable rules of inference for assigning comparable degrees of possibility to specific conclusions within the framework of an abductive logic.

A common rule of inference within inductive reasoning that has been taken as fundamental by Hans Reichenbach (1949), Wesley Salmon (1967), and others is known as *the straight rule* (SR), whose application can be formulated as follows:

(XIII) (P1) If m/n observed As are Bs, then infer (inductively) that m/n As are Bs;

(P2) m/n observed As are Bs;

═══════════════════════════

(C3) m/n As are Bs

provided that a large number of As are observed over a wide variety of conditions. Observe that even if every A observed under a wide variety of conditions were a B, no matter how numerous, that would not make the inference deductive. It would be a special case of an inductive inference in which the number m happens to equal n.

That an inference involves probabilities does not imply that it must be inductive, however. If the probability that a bean in a bag is white is 2/3, for example, then the probability that a bean in a bag is non-white must be 1/3, as a simple deductive inference. Peirce (1878a) goes even further, however, arguing that, "If, from a bag of beans of which we know that 2/3 are white, we take one at random, it is a deductive inference that this bean is probably white, the probability being 2/3" (CP 2.623). I am inclined to schematize the inference involved as possessing the following form:

(XIV) DEDUCTION:

(P1) 2/3 of the beans in this bag are white (rule)

(P2) This is a (random) bean from this bag (case)

(C3) The probability this is white is 2/3 (result)

Peirce offers a long-run justification for this argument, according to which this is a "random" draw just in case, over the long run, draws made this way have a relative frequency equal to the relative frequency of white beans in the bag. It is supposed to be analytic that the long-run probability that the bean is white must equal 2/3.

Leaving the long-run aspect aside for the moment, consider that no bean is 2/3 white: every bean is either white or non-white. (Curiously, Peirce even interprets his first premise as, "The beans in this bag are 2/3 white", which is rather striking.) When we talk about specific beans, presumably we want to know the strength of the evidential support for the conclusion that they are white. Thus, in accord with Hempel's analysis, (XIV) should be recast as an inductive argument having the form,

(XV) INDUCTION:

(P1) 2/3 of the beans in this bag are white (rule)

(P2) This is a (random) bean from this bag (case)
════════════════════════════════[2/3]

(C3) This bean is white (result)

where the bracketed number represents the strength of the evidential support for the conclusion provided by its premises, which we can call its "logical probability". Peirce even provides a justification for regarding its value as the long-run truth frequency with which such conclusions are true, given the truth of such premises.

6. INDUCTION AND EXPLANATION.

Since schema (XV) is an inductive argument of the form, rule-case-result, in that order, it should be a deductive argument on Peirce's system of classification, which suggests that something must be wrong. As Niiniluoto has observed, in his article, "A Theory of Probable Inference" (1883), Peirce offers several modes of probable "deduction" from statistical premises, including what he called "simple probable deduction", whose structure is the same as that of (XV). As Niiniluoto remarks, Peirce even emphasizes that its conclusion is "This is a B" and not "The probability that this is a B is 2/3" (Niiniluoto 1993, pp. 197-198). I do not mean to be picking nits in observing that "simple probable deduction", thus understood, is a form of inductive inference, since all such arguments are non-demonstrative, ampliative, and non-additive. But it is still of the Peircean form, rule-case-result.

Hintikka could argue that even if the straight rule were an acceptable rule of inductive inference,that would not demonstrate that abduction is an acceptable inferential practice. That would require demonstrating that there are acceptable rules of abductive inference as well. What may be most interesting about schema (XV) from this point of view, therefore, is that is represents the logical structure of inductive-statistical explanations on one or another of Hempel's accounts, provided that the statistical generalization that subsumes the specific case is lawlike (Hempel 1965 and Hempel 2001). This raises the intriguing prospect that the straight rule might qualify as an inductive form of inference to the best explanation, which is a possibility that has been explored in several papers by Gilbert Harmon (1965, 1967, 1968). Perhaps Harmon's approach can shed light on these fundamental problems.

Observe that, if the straight rule (or enumerative induction) were to provide an acceptable form of inference to the best explanation, that would establish that there is at least one acceptable mode of inference to the best explanation, but it would not necessarily establish that that mode is a mode of abduction. An inference from cases and results to rules, after all, is not the same as an inference from rules and cases to results, much less from rules and results to cases. Strictly speaking, it would appear that an argument of this kind conflates two kinds of arguments, one inductive (the straight rule inference to a generalization), the other deductive (subsuming specific cases by means of that generalization). It seems implausible that the conjunction of an inductive inference with a deductive inference would yield an abductive inference, but let us consider whether Harmon has identified the pieces that unlock this puzzle.

7. INDUCTION BY ENUMERATION.

We know from Hempel that, apart from theoretical explanations of specific laws, adequate scientific explanations for specific events entail the subsumption of those events by means of covering laws of one or another kind, as Harmon acknowledges:

> Enumerative induction argues from an observed correlation either to a generalization of that correlation or to correlations in the next instance. If the relevant generalization is a law of nature, it is relatively easy to see why the inference must be a special case of inference to the best

explanation, since the relevant law will explain the observed correlation. (Harmon 1968, p. 531).

The catch, of course, is the straight rule provides no basis for distinguishing laws of nature from mere correlations. Which means that (SR) affords no way to ascertain when Harmon's hypothetical antecedent--"the relevant generalization is a law of nature"--happens to be satisfied. Enumerative induction is not inference to laws.

Consider the simplest case. Even if 100% of Ferraris world-wide were red, so that a large number have been observed under a wide variety of conditions, it would not follow that this generalization is a law of nature. As Karl R. Popper sought to explain, laws have *the force of prohibitions*: they cannot be violated, the cannot be changed, and they require no enforcement. Insofar as there are processes or procedures by means of which the color of a Ferrari could change, that generalization, even if it were true, would be no law. Correlations exist for random properties, but that does not entail the existence of corresponding laws. Absent principles that distinguish genuine laws from accidental generalizations--colloquially, causation from correlation--what Harmon has to say is either true but trivial or significant but false, since the straight rule does not separate them.

If enumerative induction does not sustain inference to laws, then it does not sustain inference to the best explanation. Does anyone think that, even if 100% of Ferraris were red, that would explain why? The explanation in this case, as in the case of other accidental generalizations, would be distinct for each vehicle in the class: the original paint was red, their owners liked it, and so forth. But the only evidence that counts for the existence of a law is evidence that they cannot be violated or changed, say, by repainting. While successful attempts to falsify lawlike hypotheses tend to refute them, unsuccessful attempts to falsify the same hypotheses tend to support them, but fallibilistically in the mode of corroboration (Popper 1965; Fetzer 1980, 1993a). Yet the discovery of correlations can serve the heuristic function of suggesting the possible existence of corresponding laws to test.

8. VALIDATION, VINDICATION, EXONERATION.

Reichenbach (1949), Salmon (1967), and others have been profoundly affected by the Humean conception of universal laws as constant conjunctions and statistical laws as relative frequencies, which harmonizes well with accounts of probabilities as limiting frequencies. The identification of natural laws with limiting frequencies and of the straight rule as the foundation of scientific inference, moreover, can be given a "pragmatic vindication" in the form of the argument that, when the goal of science is taken to be the discovery of natural laws and those laws are identified with limiting frequencies, reliance upon the straight rule is guaranteed to lead to their discovery,.which is not the case for any alternative rule. It can be laid out as a decision matrix, where if those limits exist and we employ (SR), then they are sure to be discovered, where no method can succeed if those limits do not exist (Fetzer 1993a, pp. 114-117).

Speaking generally, demonstrating that specific arguments satisfy acceptable rules of induction or deduction thereby "validates" them, while demonstrating that specific rules of induction or deduction can fulfill their intended function thereby "vindicates" them. Just as arguments that satisfy various deductive forms, such as *modus ponens* and *modus tollens*, can be shown to be valid thereby as instances of acceptable rules, those rules in turn can be shown to be acceptable in relation to the desideratum that the conclusions they validate can never be false when their premises are true, which is their vindication. Similarly, specific inductive rules can be shown to be acceptable in relation to the objective that they are intended to fulfill, which Reichenbach and Salmon take to be the discovery of limiting frequencies across empirical sequences.

While the truth-preserving desideratum that underlies the vindication of rules of deductive inference appears to be impeccable and capable of its own exoneration, the identification of limits with laws is not. Even if infinite extensions of finite sequences were features of the physical world, sentences describing them would be descriptive and non-explanatory. They would describe properties of one history of the physical world rather than what would have to be true of every lawful history. The existence of stable relative frequencies over finite sequences, however, can serve as evidence for the existence of natural laws under a far more adequate conception as logically contingent subjunctive conditionals that attribute permanent attributes to objects instantiating reference properties, where those attributes in turn are dispositions of universal or of probabilistic strength. The discovery of laws proceeds by a process of conjectures and attempted refutations, as Popper proposed (Fetzer 1980, 1993a).

9. IS "ABDUCTION" AMBIGUOUS?

These distinctions contribute to understanding why enumerative induction fails to separate correlations from causation and cannot qualify as inference to the best explanation. It appears to be an acceptable mode of inductive inference when the existence of limits and infinite sequences is not in doubt, which may be the case in many abstract contexts. For reasoning within the context of pure mathematics, (SR) appears unobjectionable. For reasoning within the context of applied mathematics, however, where an ontology of permanent attributes and dispositional properties appears to be required, it cannot sustain inference to the existence of natural laws. Hintikka's skepticism over the prospect of establishing acceptable rules of inference for assigning degrees of possibility to specific conclusions in the framework of an abductive logic has yet to be overcome. Is abduction no more than guessing right?

A familiar distinction between the contexts of discovery and of justification would appear to imply that an inference to the existence of a *possible* explanation has to be separated from an inference to the acceptance of an *actual* explanation. Surprisingly, Peirce does not consistently preserve this distinction. Some of his examples display inferences to possible explanations, while others imply their acceptance as adequate. He states, "As a general rule, hypothesis is a weak kind of argument. It often inclinesour judgment so slightly toward its conclusion that we cannot say that we believe the latter to be true; we only surmise that it might be so" (CP 2.624). This supportsthe conception of abduction as a process of hitting upon an hypothesis as a possibleexplanation, where there is no implication that the possible explanation must be true. Other passages, however, might be interpreted as supporting a far stronger position.

A few paragraphs earlier, for example, Peirce offers the following example of the application of the method of hypothesis in a typical situation of abductive inference:

> Suppose I enter a room and there find a number of bags, containing different kinds of beans. On the table there is a handful of white beans; and, after some searching, I find one of the beans contains white beans only. I at once infer as a probability, or as a fair guess, that this handful was taken out of that bag. This sort of inference is called *making an hypothesis*. It is an inference of a *case* from a *rule* and a *result*. (CP 6.623)

Here Peirce seems to suggest that the inference drawn is not merely a possibility but a probability, where the identified hypothesis is not merely preferable but may even be acceptable. And this suggests in turn that Peirce's conception might be ambiguous.

10. THE COMPREHENSION THESIS.

In the earliest stages of his discussion, Hintikka cites four theses that have been advanced by Tomis Kapitan (1997, pp. 477-478), which Hintikka believes indicate inconsistency in Peirce and therefore imply problems for understanding abduction:

> (T1) *The Inferential Thesis:* Abduction is, or induced, an inferential process or set of processes (CP 5.188-189, 7.202);
>
> (T2) *The Thesis of Purpose:* The purpose of "scientific" abduction is both (i) to generate new hypotheses and (ii) to select hypotheses for further examination (CP 6.525);
>
> (T3) *The Comprehension Thesis*: Scientific abduction includes all the operations whereby theories are engendered (CP 5.590); and,
>
> (T4) *The Autonomy Thesis:* Abduction is, or embodies, reasoning that is distinct from, and irreducible to, either deduction or induction (CP 5.146).

The inferential thesis (T1), the thesis of purpose (T2), and the autonomy thesis (T4) appear to be consistent with the conception of abduction as inference to the existence of a possible explanation in accordance with the pattern rule-result-case, as schemata (III), (VI), and (VII), reflect. But it would be difficult to suppose that such a pattern would reasonably encompass "all the operations whereby theories are engendered".

Hintikka observes that, "The identification of the problem of Peircean abduction with the nature of ampliative inference is largely justified by the Comprehension Thesis" (Hintikka 1998, p. 93). He suggests that it is a merit of an hypothesis or a theory to explain new, previously unknown facts. But if they are genuinely new-- if these facts were unknown at the time of the abduction--they cannot be among their premises. More often scientific reasoning looks less like an inference to some hypothesis that explains the facts than the extension of an hypothesis to new data it turns out to explain. Insofar as abductive reasoners might not possess adequate explanations even for *known* data, "the abductive inference cannot be a step [from] the known data to a hypothesis or theory that explains them" (Hintikka 1998, p. 95).

Hintikka thereby poses a dilemma for the construction of abduction as inference to the best explanation. In one version, it is *epistemic,* since (he suggests) it cannot be an inference to genuinely new, unknown facts, because genuinely unknown facts cannot figure as the premises in any inference. Here the use of the terms "known" and "unknown" may be misleading, since an abductive inference (presumably) can yield a new hypothesis not previously considered as a conjecture whose truth value is epistemically indeterminate. In another, however, it is *logical,* since an inference of the pattern rule-result-case presupposes the availability of the rule to serve as a premise in the inference to a possible case as an explanation of the result. Insofar as hypotheses and theories, especially concerning natural laws, are said to qualify as "rules", the basic abductive pattern should instead be from cases and results to rules.

11. IS ABDUCTION INCOHERENT?

The logical version of Hintikka's dilemma appears to contradict Peirce's scheme of classification, where Deduction proceeds from rules and cases to results, Induction from cases and results to rules, and Abduction from rules and results to cases. So if inference to the best explanation can sustain inferences to hypotheses and theories of broad scope and systematic power--which incorporate general principles having the character of natural laws for the purposes of explanation and prediction--then some forms of Abduction proceed from cases and results to rules, which means that, given his classification scheme, they should be instances of Induction rather than instances of Abduction. But inferences to hypotheses and theories incorporating principles that have the character of general laws presumably ought to be exemplars of Abduction!

Largely on the basis of his epistemic dilemma, Hintikka concludes that "the first and crucial step to a scientific hypothesis or theory that abduction is supposed to be cannot be thought of as an inference to the best explanation" (Hintikka 1998, p. 96).While I believe that Hintikka has identified a serious problem in Peirce's exposition, I do not believe that its ramifications are quite as drastic as he suggests. No doubt,because of his preoccupation with epistemic relations, he commits a subtle fallacy in not appreciating the provisional and tentative character of conjectural explanations of the evidence to be explained. Coming up with hypotheses and theories as possible explanations for the phenomena entails creative acts of imagination and conjecture that are proper paradigms of abductive inference, especially if they were previously unknown. Indeed, they might have never before entered the mind of any thinker!

What I want to suggest, therefore, is that the accent in abductive reasoning ought to fall on the nature of explanations rather than the rule-result-case formulation. An explanation of either universal-deductive or inductive-statistical form will have both a rule and a case among the premises by means of which its result is subsumed. That means an inference to the best explanation might proceed from the result backward to the rule, when the case is given, or backward to the rule, when the result and the case are given. Peirce may have been preoccupied with the pattern rule-result-case because that is typical of explanations for singular events, especially when the laws relating cases and results of that kind are already known. Indeed, inferences from correlations to laws would appear to be yet another variation, where such inferences are justified only when repeated attempts to violate those correlations have failed.

12. ABDUCTION AND EXPLANATION.

Consider, for example, the discovery of a woman's body. Suppose that there are signs of strangulation. Then if would be an abductive inference from the result and the case to the rule if one were to infer that she may have died from strangulation, insofar as strangulation induces death. Suppose those signs of strangulation were subtle and had been overlooked. Then it would be an abductive inference to infer that, given strangulation induces death, she might have died from strangulation. In this instance, the abductive inference would be from the result and the rule to the case. Both inferences would be to possible explanations. Both appeal to laws. And if the relationship between strangulation and death were not yet known, it is not a stretch to suppose that, confronted with a series of deaths with signs of strangulation, it might be inferred from repeated cases that perhaps strangulation induces death.

All of these are examples of abductive inference involving inference to possibleexplanations. None of them would be considered to be conclusive or to require no further investigation. In the case of a body with no obvious signs of the cause of death, it might be appropriate to look once again for signs of strangulation. In the case of a body with signs of strangulation, it would be appropriate to conduct an autopsy to confirm the cause of death to insure she had not died from some other cause, such as an overlooked gunshot wound, instead. And when confronted with recurrent phenomena associating one possible cause of death with deaths over a number of trials, it might be appropriate to examine more and more trials across widely varied conditions to insure that these correlations are effects of causation.

If we pursue this avenue of approach, then the key to abduction turns out to be the theory of explanation. Hintikka remarks that, "the nature of explanation is scarcely any clearer than the nature of abduction" (Hintikka 1998, p. 94). Recent work on explanation, however, supports the conception that adequate scientific explanations for the occurrence of singular events must satisfy four requirements.

> (CA-1) the explanans must stand in a suitable (inductive or deductive) logical relationship to the explanandum by virtue of subsuming that event's description;
>
> (CA-2) the subsumption relation must obtain by virtue of the presence of one or more lawlike sentences as general premises that would be laws if they were true;
>
> (CA-3) the antecedent of the subsuming general premise must not include the description of any properties whose presence or absence is not nomically relevant;
>
> (CA-4) the sentences that constitute the explanation -- the explanandum as well as the explanans--must be true (Fetzer 1980, Ch. 5; Niiniluoto 1993; Fetzer 1993a).

Conditions of adequacy (CA-1) and (CA-2) implement the cover law conception of explanation by subsumption, while (CA-3) excludes factors that, although present, were not nomically relevant to the occurrence of the explanandum event. (CA-4) imposes the truth condition on both the explanans and the explanandum, since the truth of premises of inductive arguments do not guarantee those of their conclusions. Adequate explanations then cite all and only those factors that brought them about.

13. POSSIBLE EXPLANATIONS.

Abductions would then differ from Inductions at least insofar as Abductions are potentially explanatory, while Inductions need not be. In later work, Peirce offered a characterization of "abduction" as an inferential step for engendering hypotheses:

(XVI) The surprising fact C is observed;

> But if A were true, C would be a matter of course;
>
> Hence, there is reason to suspect that A is true. (CP 5.189)

This approach could be generalized, as Niiniluoto (2001) has done, by offering in its place a schema for specific classes of possible causes rather than one specific cause:

(XVII) The surprising fact C is observed;

But if something like A were true, C would be a matter of course.

Hence, there is reason to suspect that something like A is true.

Thinking a problem through, a student of science, such as Kepler, might have thought of a type of non-circular hypotheses to account for the orbits of the planets around the Sun as a plausible approach before adopting an elliptical hypothesis specifically. And similarly for Newton's contemplation of various laws for universal gravitation (Niiniluoto 2001, pp. 239-240). And this approach has a broad scope of application.

The conditions for plausible abductions of this kind, however, are going to be the same as those for potential explanations, which are the same as those for adequacy minus truth. In other words, a possible explanation for the occurrence of a puzzling event must subsume that event by means of a covering law and exclude conditions irrelevant to the event as explanans stands to explanandum. And while producing a material conditional of the form "If A then C" for any value of "A" and any value of "C" may not be challenging--since you can always simply assert "If A then C"—that the relationship between them must present a possible law as a logically contingent subjunctive conditional attributing a permanent property to every member of some reference class is another matter entirely, not to mention the synthetic character of the nomic relations between properties in the world that can stand as cause to effect

Hintikka's concern with respect to whether abduction is no more and no less than "some mysterious power of guessing right" appears less well-founded in light of the consideration that abduction as inference to a possible explanation has to satisfy the conditions for an explanans to serve as a possible explanation. This tends to clarify why differences in background knowledge, intelligence and ingenuity might make a difference here. Kepler and Newton were capable of impressive feats of abduction in part because they understood the problem situation better than their peers. But even if this makes abduction as inference to a possible explanation plausible, it does not establish the existence of corresponding rules of inference for assigning comparable degrees of possibility to specific conclusions in the framework of an abductive logic.

14. LIKELIHOODS AND PROBABILITIES.

A common characterization of the conception of science known as "Inductivism" suggests that science proceeds by a process of four stages consisting of Observation, Classification, Generalization, and Explanation, where the straight rule is the basic rule of inference. Since the straight rule is incapable of distinguishing correlation from causation, it cannot sustain inference to law and cannot satisfy the conditions required for explanation. The strongest result that Inductivism can in fact provide, therefore, is Prediction. A counterpart conception of science called "Abductivism", by contrast, holds that science proceeds by a process of four alternative stages of Puzzlement, Speculation, Adaptation, and Explanation (Fetzer 1980, 1993a, 2002). Thus understood, Abductivism incorporates the inference to possible explanations as the foundation for its second stage, Speculation, which depends upon the exercise of imagination and conjecture in attempting to specify a complete set of alternatives.

The third stage, Adaptation, incorporates both deductive and inductive forms of reasoning, since any alternatives that are incompatible with the available evidence are tentatively excluded from further consideration. Just as the logical structure of inductive arguments on Peirce's account display the properties of *probabilities*, the logical structure of abductive arguments on this extension of Peirce's account display the properties of *likelihoods*, where the likelihood of hypothesis h, given evidence e, is defined as the probability of evidence e, given hypothesis h. Because more than one hypothesis can confer a high probability on an outcome, the sum of two or more likelihoods may be greater than 1. The appraisal of alternative possible explanations can then proceed on the basis of comparisons between their respective likelihoods, where a rule of maximum likelihood can be employed (Fisher 1958, Michalos 1969).

According to the rule of maximum likelihood, the hypothesis that has the highest likelihood on the available evidence among the set considered ought to be accepted. By adopting likelihood measures of evidential support, it becomes possible to utilize Hacking's *laws of likelihood*, according to which an hypothesis h1 is better supported than alternative h2 when the likelihood of h1, given e, is greater than the likelihood of h2, given e, or when the likelihood ratio of h1 given e to h2 given e is greater than 1 (Hacking 1965). In order to render these principles applicable in the framework of inference to the best explanation, the explanandum must be taken to be the evidence and the explanans as the hypothesis (in variations that are described above), where the general rule that supports the likelihood assignment must be lawlike. Inference to the best explanation thus depends upon nomic probabilities and nomic likelihoods.

15. ABDUCTION VS. ABDUCTION

Although the principles of probability and of likelihood can be applied within abstract domains and to inferences about correlations (between samples and their parent populations), for example, their applicability within the context of inference to the best explanation requires the availability of lawlike premises concerning not frequencies but propensities as causal probabilities in determining the counterpart nomic likelihoods (Fetzer 1980, 1993a, 2002). Advancing explanations for singular events presumes access to relevant covering laws. If, under specifiable conditions, strangulations (gunshots, poisoning, etc.) induce death, that Peggy was strangled can explain her death, but not if she was shot or poisoned, etc. Thus, explanations from universal laws have the Peircean logical structure of deductions of rule-case-result,

(XVIII) DEDUCTION: (P1) Strangulations induce death (rule)

(P2) Peggy was strangled (case)
______________________________ [certainly]

(C1) Peggy is dead (result)

which qualifies as an adequate explanation for her death provided that it satisfies (CA-1) through (CA-4), including the exclusion of explanatorily irrelevant factors. Notice that, in cases of this kind, strangulation stands to death as cause to effect.

Or, given that Peggy is dead and that strangulations induce death, it would be an appropriate inference to a possible explanation to infer that she was strangled, even before undertaking a systematic evaluation of the alternative explanations,

(XIX) ABDUCTION: (P1) Strangulations induce death (rule)

(P2) Peggy is dead (result)
-- [possibly]

(C1) Peggy was strangled (case)

but only at the stage of Speculation. There would be no justification for inferring that she was strangled as an acceptable conclusion unless alternative hypotheses (that she was shot, that she was poisoned, etc.) had been eliminated, where, say, that she was strangled might eventually be established beyond a reasonable doubt by establishing that no alternative explanation is reasonable, which might be done by establishing that there are indications of strangulation but not of gunshots, etc.

Thus, when Abduction in the sense of inference to a possible explanation is taken as the object of discussion, Hintikka's concern that it does not support an inference to an acceptable conclusion is completely defensible and the apparent tension between Abduction as a process of speculation and the thesis that scientific abduction includes all the operations whereby theories are engendered would seem impossible to deny. So there is an historical basis for objections of this kind. But when we acknowledge that possible explanations must satisfy specific conditions of adequacy that are not arbitrary or capricious and that Abduction in this narrow sense is only one stage in the process of Abduction in the broader sense as "Inference to the Best Explanation",it also becomes apparent that, granting the role of imagination and conjecture, neither of them can justifiably be characterized as "some mysterious power of guessing right".

16. INFERENCE TO THE BEST EXPLANATION.

As an historical question, therefore, the evidence appears to be equivocal, since some passages from Peirce support the interpretation of Abduction as a method for inference to possible explanations, but others support the interpretation of Abduction as inference to the best explanation. The suggestion that I am advancing is that the evidence is equivocal because there was a profound ambiguity in Peirce's conception between Abduction as a method for inferring to possible explanations and Abduction as a method for inferring to the best explanation, where Speculative Abduction may most appropriately be envisioned as a stage of any Inference to the Best Explanation.And that the comprehension thesis only becomes plausible in relation to the totality of Peirce's work when this distinction is drawn and the proper object of evaluation is taken to be Inference to the Best Explanation and not simply Speculative Abduction.

Perhaps the most fundamental logical difference between them is that Speculative Abduction applies to hypotheses individually while Inference to the Best Explanation applies to them collectively. Since, under specific conditions, gunshots and poisoning also induce death, the situation supports additional Speculative Abductions as follows:

(XX) ABDUCTION: (P1) Gunshots induce death (rule)

(P2) Peggy is dead (result)
------------------------------------ [possibly]

(C1) Peggy was shot (case)

(XXI) ABDUCTION: (P1) Poisoning induces death (rule)

(P2) Peggy is dead (result)
------------------------------------- [possibly]

(C1) Peggy was poisoned (case)
…

The mere possibility that Peggy was strangled (shot, poisoned, etc.), as we well know, doesn't make it so, which makes it all the more apparent why, construed in this way, the very idea that Speculative Abduction includes all the operations whereby theories are engendered is not merely implausible but at least faintly ridiculous, if not absurd.

17. ABDUCTION VS. INDUCTION.

It should be all the more apparent that Inference to the Best Explanation has to proceed through the successive stages of Puzzlement, Speculation, Adaptation, and Explanation, where the process of Adaptation requires (1) rejecting alternatives that are incompatible with the available evidence and (2) comparing the alternatives that are compatible in relation to their respective degrees of evidential support. Using a likelihood measure, therefore, and employing the laws of likelihood, the hypothesis with the highest likelihood is the preferable hypothesis among the alternatives under consideration. But even though that means it is the hypothesis that is most worthy of acceptance, that does not mean it is therefore acceptable. The preferable hypothesis only become acceptable when sufficient evidence becomes available or settles down.

When we consider drawing inferences from correlations to corresponding laws, it becomes apparent that distinguishing between them requires undertaking repeated tests in an attempt to violate them, because lawful generalizations cannot be changed and require no enforcement. When relative frequencies over repeated runs of trials have resisted our best attempts to alter them, then the existence of those stable short run frequencies affords appropriate evidence for the existence of corresponding laws, provided, of course, that acceptance is understood as tentative and fallible. Although Harmon (1967) appeals to maximum likelihood in defense of enumerative induction as inference to the best explanation, that is not enough for inference to covering laws.And while (SR) cannot support inference to covering laws, it can support predictions.

Suppose, for example, that 2/3 of the Ferraris in a sample have been observed to be painted red. If a large number of Ferrari's had been observed over a wide variety of conditions (in different countries, seasons, etc.), (SR) would justify inference (XXI),

(XXII) (P1) If m/n observed As are Bs, then infer (inductively) that m/n As are Bs;

(P2) 2/3 of observed Ferraris are reds;

═══════════════════════

(C3) 2/3 of Ferraris are reds

which, in turn, could be used to make a prediction about the next Ferrari observed,

(XXIII) (P1) 2/3 of Ferraris are red (rule)

(P2) This is another Ferrari (case)

═══════════════════════[2/3]

(C1) This Ferrari is red (result)

where, unlike schema (XVIII) relating strangulation to death, schema (XXII) is not explanatory, because being red is not a lawful effect of being a Ferrari as its cause. Although this argument satisfies condition (CA-1) of derivability, it does not satisfy condition (CA-2) of lawlikeness nor condition (CA-3) that excludes irrelevant factors.

18. SIMPLE PROBABLE DEDUCTION

While predictions can be based upon merely correlations, explanations require inference from covering laws. As Niiniluoto (1993) observes, Peirce's pattern of "Simple Probable Deduction", which has the same logical form as (XXII), namely:

(XXIV) (P1) m/n As are Bs (rule)

(P2) This is an A (case)

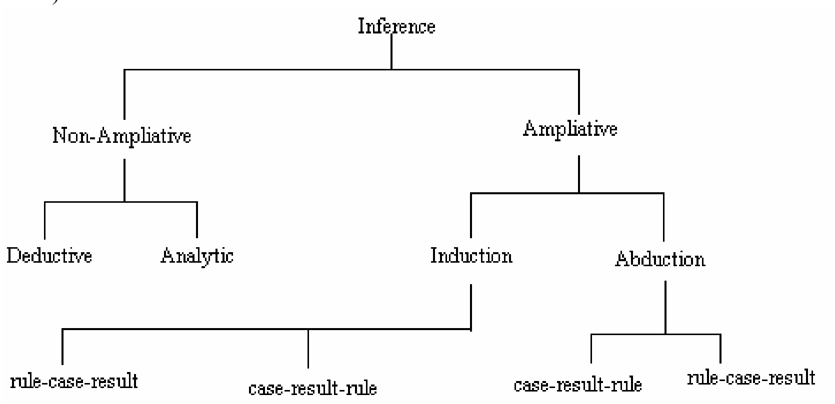
════════════════════[m/n]

(C1) This is a B (result)

would provide a potential explanation if its rule were lawlike, which would be the case if it were subjunctively conditional and attributed a dispositional property of probabilistic strength to every instance of its reference property. Peirce endorses a long run propensity conception that moves in this direction (CP 2.664), but which can be improved upon by its single case counterpart (Fetzer 1993b), for which short and long runs of trials are properly envisioned as short and long runs of single cases.

Under this interpretation, the logical probability [m/n] that specifies the degree of support that the premises confer upon their conclusion can be understood not only as the value of the frequency with which conclusions of that form will tend to be true when the premises are true but as a degree of entailment with which the conclusion will be true in any single case. Since the conclusion can still be false even when the premises are true, these arguments are inductive with respect to their conclusions, even though they display the rule-case-result pattern Peirce classifies as "Deduction". Since Induction includes inferences from samples to populations and from rules and cases to results, while Abduction includes inferences to the existence of laws and to cases from rules and results, an alternative to Peirce's diagram might work better where the ampliative/non-ampliative distinction takes precedence over logical form. (See (XXV) on the following page)

(XXV)



19. QUESTIONS AND ANSWERS.

Hintikka suggests that Peirce's distinction between Deduction, Induction, and Abduction does not reflect the difference between (what he calls) "definitory" and "strategic" rules, where definitory rules are truth-preserving rules for which the truth of their premises guarantees the truth of their conclusions. Strategic rules, by contrast, promote inquiry by yielding truth in the long run (Hintikka 1999, p. 100). Peirce, of course, promoted a pragmatic conception of truth as a property of those beliefs the community of inquirers is ultimately destined to accept or to adopt over the long run as a result of applying scientific methods to answerable questions forever. This conception appears difficult to defend when we consider that certain properties and relations might have no instances, too few instances, or enough instances that are unrepresentative "by chance" (Fetzer 1983, p. 31).

Hintikka interprets Peirce as implying that "all inferences have to be judged strategically" for their contribution to the discovery of truth in the long run as an effect of their propensity to yield truth. He regards the distinction between "ampliative" and "non-ampliative" arguments as fundamental, which suggests that he would be sympathetic to the relationships reflected by schema (XXV). His solution to "the problem of abduction", which he generalizes as the problem of ampliative inference in general, situates abductive inferences as answers to to an inquirer's questions in relation to some specific source of answers. And he also endorses the conception of the theory of explanation as the study of the logic of why-questions, which in turn suggests the conception of abductive inferences as potential answers to why-questions (Hintikka 1999, pp. 102-103).

Hintikka's emphasis upon inquirers and the pragmatic dimensions of inquiry suggest the prospect of a logic of inquiry that might potentially be formalized as a sextuple of *premises, conclusions, rules, languages, strategies,* and *proponents.* This reconstruction has the standing of a conjecture--it might even be said to be an abductive inference regarding Hintikka's position--but to the extent to which premises provide answers to conclusions as questions in accordance with rules for providing those answers (which are expressed in languages and pursued on the basis of specific epistemic strategies by inquirers as their proponents), it may contribute a schematic framework for appraising his position. The question that it appears to raise is whether explanations as answers to questions in accordance with rules *require* relativization to proponents and strategies as well as languages.

20. WHERE THINGS STAND.

It may be worth summarizing the situation as I see it at this point in time. While there are passages in Peirce that tend to support Hintikka's position and Peirce does make some misleading claims, Hintikka's contention that abduction in Peirce's sense appears to be no more than "a mysterious power of guessing right" ultimately cannot be sustained. The notion of "abduction" itself turns out to be ambiguous, where in its narrow sense, "abduction" entails "a capacity for guessing" that presumes the exercise of imagination and conjecture. It is not therefore "mysterious" but appears to bound by the conditions that must be satisfied by possible explanations. This is the process that I refer to here as "Speculative Abduction", which does not warrant acceptance. "Abduction" in this sense appears to be an indispensable element of "abduction" in a broader sense, which can be properly explicated as Inference to the Best Explanation.

Peirce's diagram for representing different kinds of inference based upon their logical form as defined by combinations of rule/case/result does not turn out to be the most promising depiction of the underlying relationships involved here, since both Induction and Abduction have case-result-rule varieties, just as Deduction and Induction likewise have rule-case-result varieties. More could be said about the nature of the bracketed values of the logical probabilities that relate premises to conclusions, including, for example, that ordinarily predictions will only be made when the value of [m/n] is equal to or greater than .5. But no specific value must be satisfied by explanations to be adequate nor, for that matter, by inferences to be acceptable, provided only that they are the best supported by virtue of being preferable hypotheses when the available evidence has suitably "settled down".

Hintikka ultimately endorses the conception of abductive inference as offering answers to an inquirer's questions in relation to some specific source of answers. When those questions are construed as why-question in the sense appropriate to explanation, then there would appear to be a broad convergence on the meaning of "abduction" that bears comparison with Peirce's conception, which, as we have seen, requires disambiguation. The existence of objective standards of adequacy for explanations, however, poses questions for Hintikka's interrogative model of inquiry, especially since those conditions for adequacy appear to be independent of inquirers and their strategies. I know all too well that I have not done justice to his article much less to the tapestry of his work. But by raising questions about a few of its sentential threads, I invite answers that should illuminate the whole.

REFERENCES

Fetzer, J. H. (1980), *Scientific Knowledge* (Dordrecht, Holland: D. Reidel).

Fetzer, J. H. (1983), "Transcendent Laws and Empirical Procedures", in N. Rescher, ed., *The Limits of Lawfulness* (Pittsburgh, PA: University of Pittsburgh Press), pp. 25-32).

Fetzer, J. H. (1993a), *Philosophy of Science* (New York, NY: Paragon House).

Fetzer, J. H. (1993b), "Peirce and Propensities", in E. C. Moore, ed., *Charles S. Peirce and the Philosophy of Science* (Tuscaloosa, AL: University of Alabama Press), pp. 60-71.

Fetzer, J. H. (2002), "Propensities and Frequencies: Inference to the Best Explanation", *Synthese* 132 (2002), pp. 27-61.

Fisher, R. A. (1958), *Statistical Methods for Research Workers* (London, UK: Oliver & Boyd, Ltd.)

Hacking, I. (1965), *Logic of Statistical Inference* (Cambridge, UK: Cambridge University Press).

Harman, G. (1965), "The Inference to the Best Explanation", *The Philosophical Review* 74/1 (January 1965), pp. 88-95.

Harman, G. (1967), "Detachment, Probability, and Maximum Likelihood", *Nous* 1/4 (December 1967), pp. 301-411.

Harman, G. (1968), "Enumerative Induction as Inference to the Best Explanation",*Journal of Philosophy* 65/18 (September 19, 1968), pp. 529-533.

Hartshorne, C. and P. Weiss, eds. (1974), *Collected Papers of Charles Sanders Peirce*, Vols. I and II (Cambridge, MA: Harvard University Press).

Hempel, C. G. (1965), *Aspects of Scientific Explanation* (New York, NY: The Free Press.

Hempel, C. G. (2001), *The Philosophy of Carl G. Hempel*, J. H. Fetzer, ed. (New York, NY: Oxford University Press).

Hintikka, J. (1998), "What is Abduction? The Fundamental Problem of Contemporary Epistemology", *Transactions of the Charles S. Peirce Society* 34 (1998), pp. 503-533.

Hintikka, J. (1999), *Inquiry as Inquiry: A Logic of Scientific Discovery* (Dordrecht, The Netherlands: Kluwer Academic Publishers).

Kapitan, T. (1997), "Peirce and the Structure of Abductive Inference", in N. Hauser et al., eds., *Studies in the Logic of Charles Sanders Peirce* (Bloomington, IN: Indiana University Press), pp. 477-496.

Michalos, A. (1969), *Principles of Logic* (Englewood Cliffs, NJ: Prentice-Hall, Inc.).

Niiniluoto, I. (1993), "Peirce's Theory of Statistical Explanation", in E. C. Moore, ed., *Charles S. Peirce and the Philosophy of Science* (Tuscaloosa, AL: University of Alabama Press), pp. 186-207.

Niiniluoto, I. (1999), "Defending Abduction", *Philosophy of Science* 66 (1999), pp. S436-S451.

Peirce, C. S. (1867), "On the Natural Classification of Arguments", in C. Harshorne and P. Weiss, eds.,

Peirce, C. S. (1878a), "Deduction, Induction, and Hypothesis", in C. Hartshorne and P. Weiss, eds., *Collected Papers* (1974), CP 2.696.

Peirce, C. S. (1878b), "A Theory of Probable Inference", in C. Hartshorne and P. Weiss, eds., *Collected Papers* (1974), CP 2.731.

Peirce, C. S. (1878c), "The Doctrine of Chances", in C. Hartshorne and P. Weiss, eds., *Collected Papers* (1974), CP 2.646.

Peirce, C. S. (1883), "A Theory of Probable Inference", in C. Hartshorne and P. Weiss, eds., *Collected Papers* (1974), CP 2.716.

Popper, K. R. (1965), *Conjectures and Refutations* (New York, NY: Harper & Row).

Reichenbach, H. (1949), *The Theory of Probability* (Berkeley, CA: University of California Press).

Salmon, W. C. (1967), *The Foundations of Scientific Inference* (Pittsburgh, PA: University of Pittsburgh Press).

THE DIALOGIC OF JUST BEING DIFFERENT
HINTIKKA'S NEW APPROACH TO THE NOTION OF *EPISTEME* AND ITS IMPACT ON "SECOND GENERATION" DIALOGICS

Shahid Rahman
Université Lille 3

1. HINTIKKA AND THE LOSS AND RECOVERY OF THE NOTION OF *EPISTEME*

The relation between the philosophy of knowledge and science became difficult after the anti-psychologist arguments of Frege and the further attacks of the logical positivists which virtually destroyed traditional gnoseology. However gnoseology had accomplished an important role in building the bridge between philosophy and science, and its disappearance also made the re-establishment of that relation difficult.

In the 1960s Jaakko Hintikka launched what he now calls the “first generation” of epistemic logic and what some others call “explicit” epistemic logic. The point of explicit epistemic logic was that the inference relation - that is the relation between a subject and a proposition - was made propositional and brought into the object language with the help of modal operators. Knowledge thus became an operator with which well formed formulae could be built. This way of making explicit the inference relation should have made it possible to study the logic of knowledge as the study of the logic of a certain kind of propositions. Moreover, explicit epistemic logic should have filled the gap left by the vanishing of traditional gnoseology, which understood the inference relation by means of the doctrine of "judgement", and it promised a new start for epistemology. Unfortunately these first attempts failed to have a real impact on epistemology for diverse reasons. Let me mention here one general point concerning the reception of epistemic logic and which might be seen as having prevented the success of explicit epistemic logic.

The origins of epistemic logic were attached to game semantics which understood knowledge as a passage of information. Now, the passage of information is, in game semantics, action-driven and thus has a predominantly dynamic character. More precisely: knowledge is associated

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 157-187.

to a certain type of choice functions allowing the passage of information. The problem is that the dynamic character of the early game-theoretical approach was not yet ready for the challenges of epistemology and the main stream re-translated epistemic logic into standard modal logic where the language of the possible world semantics somehow converted the dynamic approach into a static one: the semantics of standard modal logic seems to be object-driven rather than action-driven.

More recently, following once more the proposals of Hintikka and after further development by Hintikka and Gabriel Sandu, the game-theoretical approach to information became the formulation which took seriously the dynamic approach of the very beginnings of game semantics. The point of the new proposal is to abandon the idea of perfect information which pervaded the earlier approaches to epistemic logic. Moves are sometimes made by a player in ignorance of what happened in certain specified earlier moves. Such moves are said to be independent of the earlier ones. The new element rendering IF-Logic (independence-friendly logic) is a notation for informational independence. In IF the flow of information can be *propositionalised* by means of a slash which scores the different dependencies and independencies during such a game, e.g. the slash in $K(\forall x)(\exists y/K)S(x, y)$ signalises that the defender must made the choice of a value for the variable y in ignorance of the challengers choice of scenarios (worlds) corresponding to K. The conceptual impact is impressive and changed even the main-stream approaches to epistemic logic. One of the questions opened by this new approach to epistemology concerns the logic of enquiry rather than the logic of justification. In this new conception the relation of epistemic logic to epistemology and more generally to science concerns the informational process of acquiring knowledge rather than the characterisation of what knowledge is. Whatever it is, the thesis nontheless remains that knowledge has an explicitly propositional character:

The aim of this paper is to explore some of the consequences of the new proposals of the "second generation" of epistemic logic from the point of view of an earlier sister of game-theoretical semantics which appeared around 1958: the dialogic of Paul Lorenzen and Kuno Lorenz. Dialogic shared from the very beginnings the dynamic approach to semantics. In the earlier times a family struggle produced some misunderstandings: while game-theoretical semantics was understood as concerning reasoning under conditions dialogics was thought to concern only the concept of logical proof. The time is ripe to pool both approaches together and indeed new researches in computer sciences (linear logic), artificial intelligence (legal reasoning) and philosophy follow this path. Let me here briefly present some new results concerning what we could call "second generation dialogic". More precisely, the aim of the paper concerns the dialogical interpretation of impossible worlds in the context of non-normal modal logic and will suggest some IF-explorations beyond the concept of non-normality. In this interpretation non-normality will not necessarily apply to a world where impossible logical truths hold but to a world where the logic is just different. The discussion of these topics will be connected to the development of two issues typical of Hintikka's new approach, namely:

games of inquiry concerning the search for frame conditions for modal formulae while studying counterlogicals and

the difficulties involved in the application of the so-called *Hintikka strategy* and *hybrid languages* while constructing tableau systems for non-normal modal logics where explicit knowledge concerning frame conditions should be implemented.

2. NON-NORMAL DIALOGICS WITHOUT IMPOSSIBLE WORLDS AND WITHOUT IMPOSSIBLE LOGICAL TRUTHS

2.1 The epistemic role of counterlogicals: would the real logic please stand up?

Convincitur ergo etiam insipiens esse vel in intellectu ...
Anselm of Canterbury,
Proslogion, capitulum II, Ps 13, 1, 52, 1

(Thus, even he who knows nothing will be convinced that
at least it is in the intellect…)

Around the 1970s non-normal logics were associated with the problem of omniscience in the epistemic interpretation of modal logic, especially in the work of Jaakko Hintikka and Veikko Rantala, where impossible worlds with impossible logical truths were postulated.[48] Nowadays, the study of non-

[48] Cf. Hintikka [1975] and Rantala [1975]. See too Cresswell [1972] and Girle [1973].

normal logics has been connected too to the study of counterlogicals and its applications in epistemology[49].

In fact, conceiving situations in which not every mathematical or logical truth holds is a usual argumentation practice within formal sciences. However, to formulate the precise conditions which could render an adequate theory of logical arguments with counterpossibles in formal sciences is a challenging issue. Hartry Field has felt the need to tackle this challenge in the context of mathematics. Field writes:

> It is doubtless true that nothing sensible can be said about how things would be different if there were no number 17; that is largely because the antecedent of this counterfactual gives us no hints as to what alternative mathematics is to be regarded as true in the counterfactual situation in question. If one changes the example to 'nothing sensible can be said about how things would be different if the axiom of choice were false', it seems wrong …: if the axiom of choice were false, the cardinals wouldn't be linearly ordered, the Banach-Tarski theorem would fail and so forth (Field [1989]; pp; 237)

These lines actually express the central motivation for a theory of counterpossibles in formal sciences. Namely, the construction of an alternative system where e.g. the inter-dependence of some axioms of a given formal system could be studied. If we were able to conceive not only a counterpossible situation where some axioms fail to be true but also even an alternative system without the axioms in question, then a lot of information could be won concerning the original "real" system. By the study of the logical properties of the alternative system we could e.g. learn which theorems of our "real system" are dependent on axioms missing in the alternative one.[50] Moreover, I would like to add that a brief survey of the history of mathematics would testify that this usage of counterpossibles seems to be a common practice in formal sciences.

The case of the study of counterpossibles in logic called *counterlogicals* is an exact analogue of the case of mathematics and motivates the study of alternative systems in the very same way. We learned a lot of intuitionistic logics, even the *insipiens* classical logical monist learned about his system while discussing with the antirealist. This seems to be a generally accepted

[49] Especially in the work of such people as Graham Priest, Stephen Read, Greg Restall and Richard Routley-Sylvan.

[50] See also Read [1994], 90-91 and Priest [1998], 482.

fact, but why should we stop there? From free logics we learned about the ontological commitment of quantifiers, from paraconsistent logic ways of distinguishing between triviality and inconsistency;[51] from connexive logics the possibility of expressing in the object language that a given atomic proposition is contingently true; from relevance logics that it is not always wise to distinguish between metalogical and logical "if, then"; from IF and epistemic dynamic logic we learned about arguments where various types of flow of information are at stake, for linear how to reason with limited resources, and so forth.

Are these alternative logics "real" or even the "true" logic? Well actually to motivate its study the mere mental construction of them is enough, the mere being *in intellectu*, provided such a construction is fruitful. I would even be prepared to defend that as a start it is enough if they teach us something about the logic we take to be the "real" one. The construction of alternative logics, which in the latter case is conceived as resulting from changes in the original "real" logic, can be thought of as following a substructural strategy: changes of logic are structural changes concerning logical consequence.

In the next section I will offer a dialogical interpretation of non-normal modal logics where non-normality will not necessarily apply to a world where impossible logical truths hold but to a world where the logic is just different.

Actually, in this interpretation the pair *standard-non-standard* will be added to the pair "normal"-"non-normal". Furthermore, the adjectives *standard* and *non-standard* will qualify the noun *logic* rather than *world*, e.g. I will write "the standard logic Lk in the argumentative context m ". *Normal* will qualify those contexts, which do not allow the choice of a logic other than the standard one. Non-normal contexts do allow the choice of a new logic underlying the modalities of the chosen context. Before we go into the details let us distinguish between the following different kinds of counterlogical arguments:

Assume an intuitionist logician who puts forward the following conditional:

If tertium non-datur were valid in my logic, then the two sides of de Morgan Laws would hold (in my logic) too.

We take here once more our intuitionist

[51] Already Aristoteles used counterlogical arguments while studying the principle of non-contradiction, which he saw as the principal axiom of logic.

If tertium non-datur were valid in the non-standard logic Lk, then the two sides of de Morgan Laws would hold in Lk too.

In the first case the alternative logic – here classical logic – might be thought of as a conservative extension of the standard one here intuitionistic logic– i.e. any valid formula of the standard logic will be valid too in the non-standard logic. In the second case this seems to be less plausible: Lk could be a logic which is a combination of classical logic with some other properties very different from the intuitionistic ones. The situation is similar in the following cases where it is assumed that the standard logic is a classical one and the alternative logic can be a restriction:

If tertium non-datur were not valid in my logic, then one side of de Morgan Laws would fail (in my logic).

If tertium non-datur were not valid in the non standard logic Lj, then one side of de Morgan Laws would fail (in Lj).

Because of this fact it seems reasonable to implement the change of logics by means of a substructural strategy (akin to the concept of dialogics) - i.e. a strategy where the change of logics involves a change of the structural properties.[52]

Now in these examples the precise delimitation of a logic is assumed as a local condition. However; the conditional involved in the counterlogical seems to follow another logic which would work as a kind of a metalogic that tracks the changes of the local assumption of a given logic while building arguments with such conditionals. The point here is that in this type of study classical logic has no privileged status. Classical logic might be "the metalogic" in many cases but certainly not here.

[52] This strategy, as developed in Rahman/Keiff [2003], could be implemented either implicitly or explicitly. The implicit formulation presupposes that the structural rules are expressed at a different level than the level of the rules for the logical constants which are part of the object language. The explicit formulation renders a propositionalisation of the structural rules using either the language of the linear logicians or hybrid languages in the way of Blackburn [2001].

2.2 Non-normal dialogics or on how to be just different

2.2.1 Motivation

Let us call *non-standard* such argumentation contexts (or "scenarios" or "worlds") where a different logic holds relative to the logic defined as *standard*. Thus, in this interpretation of non-normal modal logic the fact that the law of necessitation does not hold is understood as implementing the idea that no logically valid argument could be proven in such systems to be unconditionally necessary (or true in any context and logic). Logicians have invented several logics capable of handling logically arguments that are aware of such a situation. The main idea of their strategy is simple: logical validity is about standard logics and not about the imagined construction of non-standard ones; we only have to restrict our arguments to the notion of validity involved in the standard logic. Actually there is a less conservative strategy: namely, one in which a formula is said to be valid if it is true in all contexts whether they are ruled by a standard or a non-standard logic. The result is notoriously pluralistic: no logical argument could be proven in such systems to be unconditionally necessary.

Anyway if we have a set of contexts, how are we to recognise those underlying a standard logic? The answer is clear in modal dialogics if we assume that the players can not only choose contexts but also the (non-modal) logic which is assumed to underlie the chosen context. In this interpretation the Proponent fixes the standards, i.e. determines which is the (non-modal) standard logic underlying the modalities of a given context. However under given circumstances the Opponent might choose a context where he assumes that a (non-modal) logic different from the standard one is at work. Now, there are some natural restrictions on the Opponent choices. Assume that in a given context **O** has *explicitly conceded* that **P** fixes the standards. In other words, the Opponent concedes that the corresponding formulae are assumed to hold under those structural conditions which define the standard logic chosen by the Proponent: we call these contexts *normal*. Thus, **O** has conceded that the context is normal – or rather, that the conditions in the context are normal. In this case **O** cannot choose the logic: it is **P** who decides which logic should be used to evaluate the formulae in question, and as already mentioned, **P** will always choose the logic he has fixed as the standard one. That is what the concession means: **P** has the choice.

Notice once more that "standard" logic does not really simply stand for "normal": normality, in the usual understanding of non-normal modal logic, is reconstructed here as a condition which when a context **m** is being chosen restricts the choice of the logic underlying the modalities of **m**.

2.2.2 Dialogics for S.05, S.2 and S3

The major issue here is to determine dynamically – i.e., during the process of a dialogue – in which of the contexts may the Opponent not have to conceded that it is a non-normal one and allowing him thus to choose a non-modal propositional logic different of the standard one. This must be a part of the dialogue's structural rules (unless we are not dealing with dialogues where the dialogical contexts with their respective underlying propositional logic are supposed to have been given and classified from the start). I will first discuss the informal implicit version of the corresponding structural rules and in the following section we will show how to build tableaux which implement these rules while formulating the notion of validity for the non-normal dialogics. Let us formulate a general rule implementing the required dynamics but some definitions first:

Definitions:

Normality as condition: We will say that t a given context m is *normal* iff it does not allow to choose a (propositional) logic underlying the modalities of m other than the standard one. Dually a context is *non-normal* iff it does allow the choice of a new logic

Standard logic: **P** fixes the standards, i.e. **P** fixes the (propositional) logic which should be considered as the standard logic underlying modalities and relative to which alternatives might be chosen.

Closing dialogues: No dialogue can be closed with the moves (P)*a* and (O)*a* if these moves correspond to games with different logics

Particle rules for non-normal dialogics:
The players may choose not only contexts they may also choose the propositional logic underlying the modalities in the chosen contexts:

□, ◊	*Attack*	*Defence*
□*A* m (□*A* has been stated at context m underlying a logic *Lk*)	**? □n$_{Lj}$ m (at the context m the challenger attacks by choosing an accessible context n and logic *Lj*)**	**A_{Lj} n**
◊*A* m (◊*A* m has been stated at context m underlying a logic *Lk*)	**?◊ m**	**A_{Lj} n (the defender chooses the accessible context n and the logic *Lj*)**

Or in the more formal notation of state of game (see appendix):

□-*particle rule*: From □*A* follows <R, σ, *A*, λ**A* $_{Lj/}$ **n** >, responding to the attack ?□$_{/Lj}$ **n** stated by the challenger at **m** (underlying the logic ***Lk***) and where λ**A* $_{L/}$**n** is the assignation of context **n** (with logic ***Lj***) to the formula *A*, and **n and *Lj*** are chosen by the challenger.

◊-*particle rule*: From ◊*A* follows <R, σ, *A*, λ**A* $_{Lj/}$ **n** >, responding to the attack ? ◊ **n** stated by the challenger at **m** (underlying the logic ***Lk***) and where λ**A* $_{L/}$ **n** is the assignation of context **n** (with logic ***Lj***) to the formula *A*, and **n** and ***Lj*** are chosen by the defender.

The accessibility relation is defined by appropriate structural rules fixing the global semantics (see appendix). To produce non-normal modal dialogic we proceed by adding the following (structural) rule:

(SR-ST10.O5) (SO5-rule):

O may choose a non-standard logic underlying the modalities while choosing a (new) context **n** with an attack on a Proponent's formula of the form $\Box A$ or with a defence of a formula of the form $\Diamond A$ stated in **m** if and only if **m** is non-normal.

P chooses when the context is normal and he will always choose the standard logic but he may not change the logic of a given context (generated by the Opponent).

The logic underlying the modalities of the initial context is assumed to be the standard logic.

Three further assumptions will complete this rule:

SO5 assumptions

(i) The dialogue's initial context has been assumed to be normal.

(ii) The standard logic chosen by **P** is classical logic ***Lc***.

(iii) No other context than the initial one will be considered as been normal.

The dialogic resulting from these rules – combined with the rules for T - is a dialogical reconstruction of a logic known in the literature as S.O5. In this logic validity is defined relative to the standard logic being classical and has the constraint that any newly introduced context could be used by **O** to change the standards. Certainly $\Box(a\vee\neg a)$ will be valid. Indeed, the newly generated context, which has been introduced by the challenger while attacking the thesis, has been generated from the normal starting context and thus will underlie the classical structural rule SR-ST2C (see appendix). The formula $\Box\Box(a\vee\neg a)$ on the contrary will not be valid. **P** will lose if **O** chooses in the second context, e.g., the intuitionistic structural rule SR-ST2I:

contexts	**O**			**P**			*contexts*
					$\Box\Box(a\vee\neg a)$	0	$1_{\{Lc\}}$
$1_{\{Lc\}}$	1	$<?\Box_{/1.1}>$	0		$\Box(a\vee\neg a)$	2	$1.1_{\{Lc\}}$
$1.1_{\{Li\}}$	3	$<?\Box_{/1.1.1\{Li\}}>$	2		$a\vee\neg a$	4	$1.1.1_{\{Li\}}$
$1.1.1_{\{Li\}}$	5	$<?\vee>$	4		$\neg a$	6	$1.1.1_{\{Li\}}$
$1.1.1_{\{Li\}}$		a	6		—		$1.1.1_{\{Li\}}$
					The Proponent loses playing with intuitionistic rules		

O wins by choosing in 3 the structural rule, which changes the standard logic into an intuitionistic logic.

Let us produce a dialogical reconstruction of another logic, known as S2, where we assume not only that the logic of the first context is normal and in general SR-ST10.O5, but also:

(SR-ST10.2) (S2-rule):

> If **O** has stated in a context $\mathtt{m}$ a formula of the form $\Box A$ (or if **P** has stated in $\mathtt{m}$ a formula of the form $\Diamond A$), then the context $\mathtt{m}$ *can* be assumed to be normal**.** Let us call (**O**) $\Box A$ and(**P**) $\Diamond A$ *normality formulae.*
>
> **P** will not change the logic of a given context but he might induce **O** to withdraw a choice of a non-standard logic by forcing him to concede that the context at stake is a normal one.
>
> A normal context can only be generated from a(nother) normal context.

The first two points establish that a formula like $\Box B$ could be stated by **P** under the condition that another formula, say, $\Box A$ holds. In this case **O** will be forced to concede that the context is normal and this normality will justify the proof of B within the standard logic. The third point of the rule should prevent that this process of justification from becoming trivial: formulae such as (**P**) $\Box\Diamond A$ $\mathtt{m}$, or (**O**) $\Diamond\Box A$ $\mathtt{m}$ should not yield normality if $\mathtt{m}$ is no normal themselves: the normality of $\mathtt{m}$ should come from "outside" the scope of (**P**) $\Box\ldots$ $\mathtt{m}$ and (**O**) $\Diamond\ldots$ $\mathtt{m}$**.**

This is, for our purposes, a more appealing logic than S.05 because it makes of the status of the contexts at stake a question to be answered within the dynamics of the dialogue. One can even obtain certain iterations such as $\Box(\Box(a\rightarrow b)\rightarrow(\Box a\rightarrow\Box b))$ which is not valid in S.05, but is in S2: the first context underlies the standard classical logic by the second S.05 assumption, the second context too because **O** will concede $\Box a$ there. Now, because the second context has been Ls-conceded by **O**, he cannot choose a logic different of the classical one, and **P** will thus win. Adding transitivity to S2 renders S3.

2.2.3 Dialogics for E.05, E2 and E3

The point of the logics presented in the section before was not to ignore the non-standard logics, but only to take into consideration the standard one while deciding about the validity of a given argument. We will motivate here a less conservative concept, namely, one in which a formula is said to be valid if it is true in all contexts whether they are ruled by a standard or a non-standard logic. These logics are known as E. In no E system will $\Box A$ be valid for any formula A.

Suppose one modifies S.05 in such a way that no context is assumed to be normal and thus every modality will induce a change of logic. This logic, called E.05, is unfortunately not of great interest: a formula will be valid in E iff it is valid in non-modal logics (think of $\Box(a\rightarrow b)\rightarrow(\Box a\rightarrow\Box b)$, which in this logic cannot be proven to be valid). Modality seems not be of interest there, and this logic can be thought of as a kind of a modal lower limit.

Now the elimination of the assumption that the first context is normal in S2 – that is, take SR-ST10.O5 and SR-ST10.2 but drop the first and third S0.5 assumptions – yields an interesting dialogic for our purposes. $\Box(a\rightarrow b)\rightarrow(\Box a\rightarrow\Box b)$ is valid there, signalising a more minimal structural condition for the validity of this formula than K (for it does not even assume, as K does, that validity concerns only contexts with the same kind of logic). Similarly one could produce D versions, etc. Indeed E2 seems to be the appropriate language where the logical pluralist might explore the way to formulate statements of logical validity which do not assume a universal scope

In fact, up to this point; this interpretation only offers a way to explore the scope of the validity of some arguments when confronted with counterlogical situations, where no middle term is to be conceived between what is to be considered standard and what not. Moreover, that a central aim of this dialogic is to explore fruitful counterlogicals seems not to have been implemented yet. In the next section I would like to suggest some further possible distinctions in order to perform this implementation.

2.2.4 Beyond non-normality

2.2.4.1 Non-normality and the IF-slash

One way of looking at non-normality is epistemically. In this reading non-normality arises because **P** might have to move without knowing if the context he is in is normal. That is, **P** will have to move in ignorance of the logic which applies to a given context.Notice that in this case he might not be able to induce **O** to withdraw from a choice of a non-standard logic. The difference between **S.05** and **S2** can be understood as a question of scope in the very sense of Hintikka. In the following examples, where the slash is introduced to signalise independence, the fourth necessity operator will be outside the scope of the first three:

$$\Box_1(\Box_2(a\rightarrow b)\rightarrow(\Box_3 a\rightarrow\Box_4 b))$$
$$\Box_1(\Box_2(a\rightarrow b)\rightarrow(\Box_3 a\rightarrow\Box_4(\Box_4/\Box_{i=1,2,3})b))$$

In our first example the formula will be valid in **S2** but not in the second example. In the non IF-formulation the failure of the formula is understood as a failure in the context of **S.05**. However, the slash formulation allows us to see the difference between the two logics as a case of imperfect flow of information: **P** will lose because he cannot induce **O** to withdraw from the choice at the second context of a non-standard logic simply because he does not know that **O** placed a necessity at this context.

What is interesting is the fact that in non-normal logics the slash can score the independence between two necessity operators and not only between a diamond and a box. The reason should be clear: the semantics of the necessity operator in non-normal logics requires restrictions on the choices of the challenger to be considered.

Let us fix the local semantics of the slash in non-normal modal dialogic with the help of the arbitrary modal operators Δi.

Δi, Δj	***Attack***	***Defence***
ΔiΔj (Δi/Δj) A m	? Δi m (at the context m the challenger attacks the modal operator with an attack adequate for Δi)	A_{Lj} n (the defender responds with a defence adequate to Δi and performed independently of the choice concerning Δj

Even more generality could be achieved if we allow the slash not only between modal operators but between the logic underlying a given context and modal operators. This is to take seriously the epistemic point mentioned above that **P** will have to move in ignorance of the logic which applies to a given context. However, I will leave the details for a future research

2.2.4.2 Dialogical games of inquiry: the seek of frame conditions

Let us take once more the following example, where the standard logic is classical logic:

> *If tertium non-datur were not valid in my logic, then one sense of double negation would fail (in my logic).*

One possible formalisation consists of translating not-valid by "non-necessary". Now the problem with this example is; that, if **P** does not change the logic; he can win the (negative) conditional in, say, **S2** in a trivial way.

Indeed, **O** will attack the conditional conceding the protasis, **P** will answer with the apodosis and after the mutual attacks on the negation **P** will win defending *tertium non-datur* in classical logic. But then the argument seems not to be terribly interesting. This follows from the fact that in the interpretation displayed above **P** may not change the standard logic once it has been fixed. In general this is sensible because validity should be defined relative to one standard and we cannot leave it just open to just any change. Moreover, though there is some irrelevance there this irrelevance concerns only the formula conceded at the object language: in our case double negation. But what is relevant and is used is the concession that the standard logic is the one where the classical structural rule applies. Finally why should **P** change the logic if he can easily win in the one he defined as standard?

However, in order to implement the dialogic of counterlogicals, one could leave some degree of freedom while changing the logical standard without too much complexity and inducing a more overall relevant approach: a given standard logic may change into a restriction of this logic. In other words, the standard logic may be changed to a weaker logic where any of its valid formulae are also valid in the stronger one **P** first defined as standard. True, the problem remains that it does not seem plausible that **P** will do it on principle: on principle he wishes to win, and if the proof is trivial all the better for him. There are two possibilities:

One is to build a dialogue under conditions determining from the start which contexts are played under the standard logic and which are the ones where the restriction of the standard logic hold (fix a model).

The other is to leave **O** to choose a conservative restriction of the logic **P** first defined as standard.

(SR-ST10.2) *:

> If **O** has stated in a context $\mathbf{m}$ a formula of the form $\Box A$ (or if **P** has stated in $\mathbf{m}$ a formula of the form $\Diamond A$), then the context $\mathbf{m}$ *can* be assumed to be normal. In these cases **O** might choose once a restriction of the standard logic and **P** must follow in his choices the restrictions on the standard logic produced by **O**.
>
> A normal context can only be generated from a(nother) normal context.

In our example **O** will choose intuitionistic logic and there **P** will need the concession of double negation if he wants to prove *tertium non-datur*. One way to see this point is that **O** actually tests if in the substructural rules defining the standard logic there are not some redundancies. Perhaps a sublogic might be enough.

For the example of this section this seems enough but one could even allow such restrictions in the case of the initial context in S.05. Moreover one could even drop the second S.05 assumption and let **P** choose an arbitrary standard logic. Take for example the case:

If transitivity were not holding in my logic, then $\Box a \rightarrow \Box\Box a$ *would fail too (in my logic).*

Suppose the standard logic is S4. We should use a notation to differentiate the modality which defines the standard logic and which is normal from the modalities which are used within the corresponding non-normal logic. Let us use "Δ" (or "∇") for necessity (or possibility) in the standard logic. Furthermore let us use Blackburn's hybrid language to "propositionalise" the properties of the accessibility relation. We could thus write

$\neg\Box(\nabla\nabla\nu_i \rightarrow \nabla\nu)_i$ *(transitivity) (in my S4 logic)* $\rightarrow \neg\Box(\Delta a \rightarrow \Delta\Delta a)$ *(in my S4 logic).*

If SR-ST10.2* applies then the Opponent will choose, say, the logic K and the Proponent will win. In these types of dialogue the Opponent functions more constructively than in the sole role of a destructive challenger. In fact, the Opponent is engaged in finding the minimal conditions to render the counterlogical conditional. Actually there has already been some work done concerning the dialogic adequate for seeking the minimal structural conditions for modal logic. The dialogues have been called *structure seeking dialogues* (SSD) and have been formulated in Rahman/Keiff [2003]. In these dialogues the "constructive" role of the Opponent is put into work explicitly.[53]

Here is another kind of example:

If the principle of non-contradiction were not valid in my logic, then one sense of double negation would fail (in my logic).

[53] In the context of the SSD with the thesis; say, *A*, the Proponent's claims that he assumes that a determined element δ_i (of a given set Δ of structural rules) is the minimal structural condition for the validity of *A*. Informally, the idea is that structural statements can be attacked by the challenger in two distinct ways. *First*, by conceding the condition δ_i, claimed by the player X to be minimal, and asking X to prove the thesis. *Second*, by (counter)claiming that the thesis could be won with a (subset of) condition(s) of lesser rank in Δ. In that case, the game proceeds in a subdialogue, started by the challenger who now will claim that the formula in question can be won under the hypothesis δ_j, where δ_j is different from δ_i and has a lesser rank as δ_i. Since the challenger (Y) starts the subdialogue *he now has to play formally*. See details in Rahman/Keiff [2003].

One other way to formalise this would be to put the negation inside the scope of the necessity operator:

> *If it were necessary that the principle of non-contradiction does not hold, then it would be necessary that one sense of double negation will fai.*

If we assume here too that SR-ST10.2* applies then the Opponent will choose some sort of paraconsistent logic (such as Sette's P1). Certainly, the Opponent will lose, anyway but other choices would lead to a trivial winning strategy of the Proponent.

If, instead of using SR-ST10.2*, we leave the choice of the standard logic open, **P** might choose any logic as standard and then it would seem that almost anything goes. It is perhaps not the duty of the logician to prevent this but the application of SR-ST10.2*and the corresponding SSD can help there, leaving the Opponent to search for the "right" structural conditions under which the formula should be tested.

The point may be put in a different way. In the dialogues of the preceding section's the role of the Opponent is to test if the thesis assumes surreptitiously that its validity holds beyond the limits of the standard logic. In this role the Opponent may choose any arbitrary logic without any constraints. Let us now assume, that the Opponent, still in the role already mentioned, comes to the conclusion that the thesis of the Proponent holds as it is. The Opponent can then play a slightly different role and explore the possibilities of another strategy: he might try to check if the standard logic chosen is not too strong concerning the thesis at stake. The latter is the aim of the structure seeking dialogues.

The preceding considerations hardly settle the matter of the ways the change of logics can be studied dialogically. There are many other possible variations – one could for example think that the SSD would be activated when some problematic assumption of the standard logic arises which might not actually concern the thesis. This will do for the present though.

3. TABLEAUX

The aim of this section is to discuss the failure of the so-called *Hintikka strategy* concerning the implementation of the accessibility relation while constructing tableau systems for non-normal modal logics. This problematic seems to apply too to the "propositionalisation" techniques of frame conditions such as practised in hybrid languages.

Let us first present the tableaux which result from our dialogic.

3.1 Dialogical tableaux for non-normal modal logics

As discussed in the appendix mentioned, the strategy dialogical games introduced above furnish the elements for building a tableau notion of validity where every branch of the tableau is a dialogue. Following the seminal idea at the foundation of dialogic, this notion is attained via the game-theoretical notion of *winning strategy*. X is said to have a winning strategy if there is a function, which, for any possible Y-move, gives the correct X-move to ensure the winning of the game.[28]

Indeed, it is a well known fact that the usual semantic tableaux in the tree-shaped structure we owe to Raymond Smullyan are directly connected with the tableaux for strategies generated by dialogue games, played to test validity in the sense defined by these logics. E.g.

(**O**)-cases	(**P**)-cases
Σ,(**O**)$A \rightarrow B$ -------------------------------- Σ,(**P**)A, ... | Σ,<(**P**)A> (**O**)B	Σ,(**P**)$A \rightarrow B$ ------------------- Σ,(**O**)A, Σ,(**P**)B

The vertical bar "|" indicates alternative choices for **O**, **P**'s strategy must have a defence for both possibilities (dialogues). Σ is a set of dialogically signed expressions. The signs "<" and ">" signalise that the formulae within their scope are moves but not formulae which could be attacked. The elimination of expressions like <(**P**)*A*> and the substitution of **P** by **F(alse)** and **O** by **T(rue)** yields the signed standard tableau for the conditional.

However, strictly speaking, as discussed in Rahman/Keiff 2003, the resulting tableaux are not quite the same. A special feature of dialogue games is the notorious formal rule (SR-ST4) which is responsible for many of the difficulties of the proof of the equivalence between the dialogical notion and the truth-functional notion of validity. The role of the formal rule, in this context, is to induce dialogue games which will generate a tree displaying the (possibly) winning strategy of **P**, the branches of which do not contain redundancies. Thus the formal rule actually works as a filter for redundancies, producing a tableau system with some flavour of natural deduction. This role can be generalised for all types of tableau generated by the various dialogics. Once this has been made explicit, the connection between the dialogical and the truth-functional notion of validity becomes transparent.

Let us see first the dialogical tableaux for normal logic as presented in Rahman/Rückert 1999 and improved in Blackburn 2001, though the notation there diverges slightly from the present one:

(O)-cases	**(P)-cases**
$(O)\ \nabla A\ \mathtt{m}$	$(P)\nabla A\ \mathtt{m}$
-------------------	--------------------
$\langle(P)?_{\nabla}\ \mathtt{n}\ \#\rangle(O)A_{Ls}\ \mathtt{n}$	$\langle(O)\ )?_{\nabla}\ \mathtt{n}\ \rangle(P)\ A_{Li}\ \mathtt{n}$
the context $\mathtt{n}$ does not need to be new	the context $\mathtt{n}$ is new
$(O)\not\subset A\ \mathtt{m}$	$(P)\not\subset A\ \mathtt{m}$
---------------------	-------------------------
$\langle(P)?\rangle(O)\ A\ \mathtt{n}$	$\langle(O)?\rangle(P)A\ \mathtt{n}\#$
the context $\mathtt{n}$ is new	the context $\mathtt{n}$ does not need to be new

"m" and "n " stand for contexts; "#" restricts the choices of **P** according to the properties of the accessibility relation which define the corresponding normal modal logics. Dialogical contexts always constitute a set of moves. These contexts may have a finite number, or a countable infinity of elements, semi-ordered by a relation of succession, obeying the very well known rules which define a tree. The thesis is assumed to have been stated at a dialogical context which constitutes the origin of the tree. The initial dialogical context is numbered 1. Its *n* immediate successors are numbered 1.*i* (for *i*=1 to *n)* and so on. An immediate successor of a context *m.n* is said to be *of rank +1*, the immediate predecessor *m* of *m.n* is said to be *of rank -1*, and so on for arbitrarily higher (lower) degree ranks. I will leave the discussion of how to specify # for the next section and display now the tableaux for non-normal dialogics:

(O)-*cases*	**(P)-*cases***
(O)∇A m ------------------- $<(P)?_{\nabla}\ n\ \#/\ Ls>(O)\ A_{Ls}\ n$ the context n does not need to be new the logic at m is the standard logic *Ls*	(P)∇A m -------------------- $<(O)\ )?_{\nabla}\ n\ /\ Li>(P)\ A_{Li}\ n$ the context n is new the logic *Li* is different from the standard one *Ls* iff m is non-normal
(O)$\not\subset A$ m --------------------- $<(P)?>(O)\ A_{Li}\ n$ the context n is new the logic *Li* is different from the standard *Ls* iff m is non-normal	(P)$\not\subset A$ m ------------------------- $<(O)?>(P)\ A_{Ls}\ n\#$ the context n does not need to be new the logic at m is the standard logic *Ls*

We need the following rule concerning closure:

Closing branches:
No branch can be closed with the moves (P)*a* and (O)*a* if these moves correspond to games with different logics.

To produce S.05 add to the adequate implementation of the accessibility relations the following:

SO5 normality conditions:

1. The dialogue's initial context has been assumed to be normal. No other context than the initial one will be considered as being normal.
2. The standard logic chosen by **P** is classical logic ***Lc***.
3. The Proponent may not:
    (a) choose a context where the logic is different of the standard one;
    (b) change the logic of a given context **m** if **m** has been generated from a non-normal context.

To produce S2 add to the SO5-rule the following:

(S2--normality conditions):

If **O** has stated in a context **m** a formula of the form $\Box A$ (or if **P** has stated in **m** a formula of the form $\Diamond A$), then the context **m** can be assumed to be normal**.**
A normal context can only be generated from a(nother) normal context.

The construction of the other tableaux is straightforward.

3.2 On how not to implement the accessibility relations

In dialogics, the properties of the accessibility relation could be implemented in the following way:

(SR-ST9.2K) (*K*): **P** may choose a (given) dialogical context of rank +1 relative to the context he is playing in.
(SR-ST9.2T) (*T*): **P** may choose either the same dialogical context where he is playing or he may choose a (given) dialogical context of rank +1 relative to the context he is playing in.
(SR-ST9.2B) (*B*): **P** may choose a (given) dialogical context of rank -1 (+1) relative to the context he is playing in, or stay in the same context.
(SR-ST9.2S4) (*S4*): **P** may choose a (given) dialogical context of rank >+1 relative to the context he is playing in, or stay in the same context.
(SR-ST9.2S5) (*S5*): **P** may choose any (given) dialogical context.

Moreover we could e.g. build the transitivity part of the rule for S4 in the tableau rule in the following way.

$$\frac{\begin{array}{l}(O)\nabla A\ \mathtt{m}\\ \mathtt{n} = \mathtt{m}\ {>}{+}1\end{array}}{<(P)?_{\nabla}\ \mathtt{n}>(O)A_{Ls}\ \mathtt{n}}$$

Actually, there is another technique to implement this and which is connected with the idea of finding in the object language formulae which express frame conditions: the idea has been used by Hintikka for the construction of tableaux and is thus known today as *Hintikka's strategy*. The idea is a bold one and captures the spirit of the axiomatic approaches. Let us formulate the rule in Hintikka's style leaving aside for the moment the choice of the logic:

(O) ∇A m
n = m >+1

<(P)?$_\nabla$ n >(O) ∇A n

That is, if ∇A holds at m then it should also hold at the context n provided n is accessible from m. The rule stems from the idea that transitivity is associated with the validity of the formula: $\nabla A \rightarrow \nabla\nabla A$.

The "up-wards" transitivity of S5 can be formulated similarly. Actually, the only device one needs is the one concerning K. Then, as soon as context has been "generated" the rules defining the other modal logics tells what formulae can be used to fill the opened context - Hintikka speaks of "filling rules. The simplicity and conceptual elegance of this strategy had made it very popular[54] and it is connected with a more radical formalisation strategy such as that of *hybrid languages*.[55] In the latter, the point is to fully translate the properties of accessibility relations into the object language of propositional modal logic, which has been extended with a device to "name contexts" such as "@m ". The idea behind the @ operator is to distinguish the assertion that a given formula *A* can be defended in the dialogical context m from the dialogical context n where the assertion has been uttered – which could be different from m.. Properties of the accessibility relation can in this case be formulated as propositions. One problem for the general application of Hintikka's strategy is that there are some frame conditions like irreflexivity, asymmetry, antisymmetry, intransitivity and trichotomy which are not definable in orthodox modal languages. The aim of hybrid languages is to close this gap by enriching the modal language and apply then Hintikka's strategy.

The hybrid strategy seems at first sight, very appealing to our interpretation of non-normal modal logic where the concession of normality actually amounts to the concession of a rule defining the corresponding standard logic. If the standard logic is a modal one, then the concession, when formulated in the style of hybrid languages, amounts to add a premise. Now, if it is indeed a premise (stating frame conditions) then it seems a good

[54] See for example Fitting [1983], 37; Fitting/Mendelsohn [1998], 52, Girle [2000], 32-34.
[55] Cf. Blackburn [2001] and Blackburn/de Rijke/Venema [2002].

idea to have this premise expressed in the same language as the other premises. For example in the following way:

$$
\begin{array}{l}
\quad (O)\ \nabla A\ @\mathrm{m} \\
\quad \not\subset\not\subset \mathrm{n} \rightarrow \not\subset \mathrm{n}\ \ @\mathrm{m} \\
\text{-+-----------------------------} \\
\quad <(P)?_{\nabla}\ \ \mathrm{n}\ >(O)\ A\ @\mathrm{n}
\end{array}
$$

However, the application of both the Hintikka and the hybrid strategy in the context of non-normal logic should be done very carefully. If not we might, say in the S3, convert a non-normal context into a normal one by the assumption that the accessibility relation is transitive.[56] Moreover, we would come to the result that every non-normal logic with transitivity collapses into normality. But normality is a condition qualifying worlds and not about accessibility. In fact the point of logic as S3 is that we could have transitivity without having necessitation. Certainly, defenders of Hintikka's and hybrid strategies might fight back introducing the proviso that their rules apply under the condition that the contexts in question are normal. In fact, Fitting uses such a strategy in his book of 1983 (274).

Anyway, this loss of generality awakes, at least to the author of the paper, a strange feeling. A feeling of being cheated: Transitivity talks about accessibility between contexts and not about necessitation in normal contexts. Hybrid languages seem to be the consequent and thorough development of a notion akin to Hintikka's strategy and perhaps pay the same price. Indeed, in the language of dialogics we would say that the propositionalisation of frame conditions amounts to producing a new (extension of a) logic without really changing either the local or the global semantics. It is analogue to the idea of producing classical from intuitionistic dialogic just by adding *tertium non datur* as a concession (or axiom) determined by the particular circumstances of a given context. Indeed, with this technique we can produce classical theorems within the intuitionistic local and global (or structural) semantics. Assume now that we are in the modal dialogic K and that in a given (dialogical) context the Opponent has attacked a necessary formula $a \vee b$ of the Proponent. Assume further that the Proponent has at his disposal a filling rule which allows him to "fill" this very context with a necessary formula of the Opponent, say, b.[57] Then

[56] Cf. Girle [2000], 187 where the exercise 3.3.1. 2(a) shows how such a mistake slipped into his system.

[57] . Moreover, if the thesis were $\nabla b \rightarrow \nabla\nabla(a \vee b)$ it would be valid.

obviously, **P** will win and strictly speaking, from the dialogical point of view, he always remains in K. One other way to see this is to realise that, what the "filling rules" do, is to allow appropriate "axioms" to be added to some contexts specified by these rules in order to extend the set of theorems of K without changing its semantics. As already acknowledged, the idea is elegant and perspicuous but it simply does not work so straightforwardly if non-normal contexts are to be included.[58]

I would like to finish with an open problem which follows from the connection of the games of enquiry to the considerations discussed above. Games of enquiry which deal with the search of frame conditions demand the propositionalisation of such conditions. But such a propositionalisation seems to be harder than expected if the game is supposed to be about seeking the adequate logic for a given context. More generally, knowledge about structural conditions defining the global semantics of a logic does not seem to be easily made explicit.

Acknowledgements

I would like to thank André Laks, director of the Maison des Sciences de l'Homme du Nord-Pas de Calais (MSH). In fact this work is part of the research projects *Preuve* and *La science et ses contextes*, attached to the MSH-Nord-Pas de Calais. Also thanks to Jaakko Hintikka (Boston) and John Symons (University of Texas at El Paso) for interesting and lively discussions concerning an earlier version of the paper presented at the meeting of the *Institut International de Philosophie* on "Open Problems in Epistemology" in Paris 20-21 November 2003.

[58] It could be fruitful to relate this problematic with *tonk*. From the dialogical point of view, *tonk* produces an extension into triviality because it has been introduced without semantic support (see Rahman/Keiff [2003]). Here, if the semantics concerning the accessibility relation is not adequately changed, the logic will collapse into just another normal modal logic.

APPENDIX:

A brief survey of dialogic:

The aim here is to introduce very briefly the conceptual kernel of dialogic in the context of the dialogical reconstruction of first-order propositional calculus, in its classical and intuitionist versions. [59]

Let our language **L** be composed of the standard components of first-order logic (with four connectives $\wedge$, $\vee$, $\rightarrow$, $\neg$, and two quantifiers $\forall$, $\exists$), with small letters (a, b, c,…) for prime formulæ, capital italic letters (A, B, C, …) for formulæ that might be complex, capital italic bold letters (***A***, ***B***, ***C***, …) for predicators, let our constants be noted τ_i, where i $\in$ N, and our variables the usual (x, y, z, …). We will also need some special force symbols: $?_{\ldots}$ and $!_{\ldots}$, where the dots stand for indices, filled with some adequate information that will be specified by appropriate rules. An *expression* of **L** is either a term, a formula or a special force symbol. **P** and **O** are two other special symbols of **L**, standing for the players of the games. Every expression *e* of our language can be augmented with labels **P** or **O** (written **P**-*e* or **O**-*e*, called (*dialogically*) *signed expressions*), meaning in a game that the expression has been played by **P** or **O** (respectively). We use X and Y as variables for **P**, **O**, always assuming X≠Y. Other more specific labels will be introduced where needed.

An *argumentation form* or *particle rule* is an abstract description of the way a formula, according to its principal logical constant, can be criticised, and how to answer the criticisms. It is abstract in the sense that this description can be carried out without reference to a determined context. In dialogic we say that these rules state the *local semantics*, for they show how the game runs locally, in the sense that what is at stake is only the critic and the answer to a given formula with one logical constant rather than the whole (logical) context where this formula is embedded. Hence, the particle rules fix the dialogical semantics of the logical constants of **L** in the following way:

[59] Cf. Lorenzen [1958] and Lorenzen/Lorenz [1978]. The present more modern version stems from Rahman/Keiff [2003].

	$\wedge$	$\vee$	$\rightarrow$
assertion	X-$A \wedge B$	X- $A \vee B$	X-$A \rightarrow B$
attack	Y-$?_L$, or Y-$?_R$	Y- $?_\vee$	Y-A
defence	(respectively) X-A or X-B	X-A, or X-B	X-B

	$\forall$	$\exists$	$\neg$
assertion	X-$\forall xA$	X-$\exists xA$	X-$\neg A$
attack	for any τ Y may choose, Y- $?_{\forall/\tau}$	Y- $?_\exists$	Y-A
defence	for any τ chosen by Y, X-A(x/τ)	for any τ X may choose, X-A(x/τ)	— (i.e. no defence)

(Where A and B are formulæ, and $A(x/\tau)$ is the result of the substitution of τ for every occurrence of the variable x in A.)

One more formal way to stress the locality of the semantics fixed by the particle rules is to see these rules as defining a state of a (structurally not yet determined) game. Namely:

Definition (*state of the game*): A *state of the game* is an ordered triple <ρ, σ, A> where:

ρ stands for a role assignment either R, from players X, Y to only one element of the set {?(attack), !(defence)} determining which player happens to occupy the challenger and which the defender role, or R', inverting the role assignment R of both players (e.g. if R(X)=? and R(Y)=!, then R'(X)=! and R'(Y)=?). The players perform their assigned role as challengers (defenders) by stating an attack (or asserting a defence) fixed by the corresponding rule.

σ stands for an assignment function, substituting as usual individuals by variables.

A stands for a dialogically labelled subformula A with respect to which the game will proceed.

Particle rules are seen here as determining which state of the game S' follows from a given state S without yet laying down the (structural) rules which describe the passage from S to S'. What state follows of S=<R, σ, F>for the X-labelled formula F?

Negation particle rule: If F is of the form $\neg A$ then S'=<R', σ, A>, i.e. Y will have the role of defending A and X the role of (counter)attacking A.

Conjunction particle rule: If F is of the form $A \wedge B$ then S'=<R, σ, A> or S'' =<R, σ, B>, according to the choice of challenger R(Y)=? between the attacks $?_L$ and $?_R$.

Disjunction particle rule: If F is of the form $A\vee B$ then S'=<R, σ, A> or S'=<R, σ, B>, according to the choice of defender R(X)=!, reacting to the attack $?_{\vee}$ of the challenger R(Y)=?.

Subjunction particle rule: If F is of the form $A\rightarrow B$, then S'=<R', σ, A> and the game might proceed to the state S''=<R'', σ, B>, or even the other way round according to the choice of the defender and reacting to the attack A of the challenger R(X)=?.

Universal quantifier particle rule: If F is of the form $\forall xAx$ then S'=<R, σ(x/τ), A> for any constant τ chosen by the challenger R(Y)=? while stating the attack $?_{\forall/\tau}$.

Existential quantifier particle rule: If F is of the form $\exists xAx$ then S'=<R, σ(x/τ), A> for any constant τ chosen by the defender R(X)=! reacting to the attack $?_{\exists}$ of the challenger R(Y)=?.

A dialogue can be seen as a sequence of labelled expressions, the labels carrying information on the game significance of these expressions. Dialogues are processes, so they are dynamically defined by the evolution of a game, which binds together all the labels mentioned. In other words, the set of expressions which is a complete dialogue can be dynamically determined by the rules of a game, specifying how the set can be extended from the original thesis formula. Particle rules are part of the definition of such a game, but we need to set the general organisation of the game, and this is the task of the *structural rules*. Actually structural rules can, while implementing the local semantics of the logical particles, determine a kind of game for a context where e.g. the aim is persuasion rather than logical validity. In these cases dialogic extends to a study of argumentation in a broader sense than the logical one. But *when the issue at stake is indeed testing validity*, i.e. when **P** can succeed with the use of the appropriate rules in defending the thesis against all possible allowed criticism by **O**, games should be thought of as furnishing the branches of a tree which displays the games relevant for testing the validity of the thesis. As a consequence of this definition of validity, each split of such a tree into two branches (dialogue games) should be considered as the outcome of a propositional choice of **O**. In other words when **O** defends a disjunction, he reacts to the attack against a conditional, and when he attacks a conjunction, he chooses to generate a new branch (dialogue). Dually **P** will not choose to change the dialogue (branch). In fact, from the point of view of games as actual (subjective) procedures (acts), it could happen that the subject playing as **O** (**P**) is not clever enough to see that his best strategy is to open (not to open) a new dialogue game (branch) anytime he can, but in this context where the issue is an inter-subjective concept of validity, which should lead to a straightforward construction of a system of tableaux, we simply assume that **O** makes the best possible move.

(SR-ST0) (*starting rule*): Expressions are numbered and alternately uttered by **P** and **O**. The thesis is uttered by **P.** All even-numbered expressions including the thesis are **P**-labelled, all odd-numbered expressions are **O** moves. Every move below the thesis is a reaction to an earlier move with another player label and performed according to the particle and the other structural rules.

(SR-ST1) (*winning rule*): A dialogue is closed iff it contains two copies of the same prime formula, one stated by **X** and the other one by **Y**, and neither of these copies occur within the brackets "<" and ">" (where any expression which has been bracketed between these signs in a dialogue either cannot be counterattacked in this dialogue, or it has been chosen in this dialogue not to be counterattacked). Otherwise it is open. The player who stated the thesis wins the dialogue iff the dialogue is closed. A dialogue is finished if it is closed or if no other move is allowed by the (other) structural and particle rules of the game. The player who started the dialogue as a challenger wins if the dialogue is finished and open.

(SR-ST2I) (*intuitionist ROUND closing rule*): In any move, each player may attack a (complex) formula asserted by his partner or he may defend himself against *the last not already defended* attack. Defences may be postponed as long as attacks can be performed. Only the latest open attack may be answered: if it is **X**'s turn at position n and there are two open attacks m, l such that $m < l < n$, then **X** may not at position n defend himself against m.

(SR-ST2C) (*classical ROUND closing rule*): In any move, each player may attack a (complex) formula asserted by his partner or he may defend himself against *any attack* (including those which have already been defended).

(SR-ST3/SY) (*strategy branching rule*): At every propositional choice (i.e., when X defends a disjunction, reacts to the attack against a conditional or attacks a conjunction), X may motivate the generation of two dialogues differentiated only by the expressions produced by this choice. X might move into a second dialogue iff he loses the first chosen one. No other move will generate new dialogues.

(SR-ST4) (*formal use of prime formulæ*): **P** cannot introduce prime formulæ: any prime formula must be stated by **O** first. Prime formulæ can not be attacked.

(SR-ST5) ***(no delaying tactics rule)***:

While playing with the <u>classical structural rule</u> **P** may perform once a new defence (attack) of an existential (universal) quantifier using a different constant (but not new) iff the first defence (attack) compelled **P** to introduce a new constant. No other repetitions are allowed.

While playing with the <u>intuitionistic structural rule</u> **P** may perform a repetition of an attack if and only if **O** has introduced a new prime formula which can now be used by **P.**

Definition (***Validity***): A tableau for (**P**)A (i.e. starting with (**P**)A) proves the validity of A iff the corresponding tableau is closed. That is, iff every dialogue generated by (**P**)A) is closed.

Examples: In Fig. 1 the outer columns indicate the numerical label of the move, the inner columns state the number of a move targeted by an attack. Expressions are not listed following the order of the moves, but writing the defence on the same line as the corresponding attack, thus showing when a round is closed. Recall, from the particle rules, that the sign "—" signalises that there is no defence against the attack on a negation. In this example **P** wins because, after the **O**'s last attack in move 3, **P**, according to the (classical) rule ST2C, is allowed to defend himself (once more) from the attack in move 1 (in the same dialogue). **P** states his defence in move 4 though, actually, **O** did not repeat his attack – this fact has been signalised by inscribing the unrepeated attack between square brackets.

O			**P**		
				$a\vee\neg a$	0
1	$?_\vee$	0		$\neg a$	2
3	a	2		—	
[1]	[$?_\vee$]	[0]		A	4

Fig. 1. SDC rules. **P** wins.

O			**P**		
				$a\vee\neg a$	0
1	$?_\vee$	0		$\neg a$	2
3	a	2		—	

Fig. 2. SDI rules. **O** wins.

In the game of Fig. 2, **O** wins because, after the challenger's last attack in move 3, **P**, according to the intuitionistic rule SR-I, is not allowed to defend himself (once more) from the attack in move 1.

Philosophical remarks: games as propositions.

Particle rules determine dynamically how to extend a set of expressions from an initial assertion. In the game perspective, one of the more important features of these rules is that they determine, whenever there is a choice to be made, who will choose. This is what can be called the pragmatic dimension of the dialogical semantics for the logical constants. Indeed, the particle rules can be seen as a proto-semantics, i.e. a game scheme for a not yet determined game which when completed with the appropriate structural rules will render the game semantics, which in turn will build the notion of validity.

Actually by means of the particle rules games have been assigned to sentences (that is, to formulæ). But sentences are not games, so what is the nature of that assignment? The games associated to sentences are meant to be *propositions* (i.e. the constructions grasped by the (logical) language speakers). What is connected by logical connectives are not sentences but propositions. Moreover, in the dialogic, logical operators do not form sentences from simpler sentences, but games from simpler games. To explain a complex game, given the explanation of the simpler games (out) of which it is formed, is to add a rule which tells how to form new games from games already known: if we have the games A and B, the conjunction rule shows how we can form the game $A \wedge B$ in order to assert this conjunction.

Now, particle rules have another important function: they not only set the basis of the semantics, and signalise how it could be related to the world of games – which is an outdoor world if the games are assigned to prime formulæ, but they also show how to perform the relation between sentences and propositions. Sentences are related to propositions by means of assertions, the content of which are propositions. Assertions are propositions endowed with a theory of force, which places logic in the realm of linguistic actions. The forces performing this connection between sentences and propositions are precisely the attack (?) and the defence (!). An attack is a demand for an assertion to be uttered. A defence is a response (to an attack) by acting so that you may utter the assertion (e.g. that A). Actually the assertion force is also assumed: utter the assertion that A only if you know how to win the game A.

Certainly the "know" introduces an epistemic moment, typical of assertions made by means of judgements. But it does not presuppose in principle the quality of knowledge required. The constructivist moment is only required if the epistemic notion is connected to a tight conception of what means that the player X *knows that there exists* a winning game or strategy for A.

Soundness and completeness of the tableaux systems

The tableau systems for non-normal logics presented above are essentially those of Fitting [1983], Girle [2000] and Priest [2001] without the use of Hintikka's strategy for the accessibility relation of the first two authors. I will not rewrite the proofs here and rely on the proofs of Fitting[1983] and Priest [2001]. What I will do is to show how to transform the dialogical tableaux into the ones of the authors mentioned above. To see this notice that if the Opponent (=**T** in the signed non dialogical version of the tableau) is clever enough, on any occasion where he may choose a logic he will choose one, where he assumes that the Proponent (=**F** in the signed non dialogical version of the tableau) will lose. In fact, if the tableau systems are thought as reconstructing the usual notion of validity of non-normal modal logic we must assume that it will be always the case that if **O** chooses a logic then **P** will lose – however, notice that dialogically we must not assume this: **O** might lack some information and choose the wrong logic. One way to implement the assumption of the cleverness of the Opponent slightly more directly is to forbid **P** to answer to an attack on a necessary formula (or to attack a possible° formula of the Opponent) stated at a context m unless this context is normal. Moreover, if we are interested in freeing ourselves from the interpretation of the contexts as representing situations where logic could be different, or more generally from any interpretation concerning the "structural inside" of non-normal contexts, the rules will amount to the following simplified formulation:

(O=T)-*cases*	**(P=F)-*cases***
(O=**T**)⊄*A* m --------------------- <(P)?>(O=**T**) *A* n the context n is new the rule is activated iff m is normal	(P=**F**)∇*A* m -------------------- <(O))?$_{\nabla}$ n > (P) n the context n is new the rule is activated iff m is normal

Furthermore, if we delete from the tableau the expressions <(P)?>(and <(O))?$_{\nabla}$ n >, which have only a dialogical motivation, we have the usual tableau systems mentioned above.

REFERENCES

Blackburn P. "Modal logic as dialogical logic". In S. Rahman and H. Rückert [2001], 2001, 57-93.

Blackburn P., de Rijke M., and Venema Y, *Modal Logic*, Cambridge, Cambridge University Press, 2002.

Creswell M. J. "Intensional logics and truth". *Journal of Philosophical Logic*, vol. 1, 2-15, 1972.

Fitting M. *Proof Methods for Modal and Intuitionistic Logic*, D. Reidel, Dordrecht, 1983.

Fitting M. and Mendelsohn R. L. First-Order Modal Logic, Dordrecht, Kluwer; 1998.

Girle R. "Epistemic logic; language and concepts", *Logique et Analyse*, vol. 63-64, 359-373, 1973.

Girle R. *Modal Logics and Philosophy*, Montreal, McGill-Queen's University Press, 2000.

Grattan-Guinness, I. "Are other logics possible? MacColl's logic and some English reactions, 1905-1912". *Nordic Journal of Philosophical Logic*, vol. 3, 1, 1998, 1-16.

Hintikka J. “Impossible Possible Worlds Vindicated”. *Journal of Symbolic Logic*, 4, 1975, pp. 475-484; modified and reedited in Hintikka J. and M.B., *The Logic of Epistemology and the Epistemology of Logic*, Dordrecht, Kluwer, 63-72, 1989.

Kripke S., "Semantical Analysis of Modal Logic II; non-normal modal propositional calculi." In J. W. Addison et alia (eds), *The Theory of Models*, Amsterdam, N. Holland, 202-220, 1965.

Lorenzen P. “Logik und Agon”. *Arti del XII Congresso Internationale de Filosofia*, Venezia. 187–194, 1958. (Reprinted in Lorenzen and Lorenz, 1–8 1978.)

Lorenzen P. and Lorenz K. *Dialogische Logi*k. WBG, Darmstadt, 1978.

McCall S., *Aristotle's Modal Syllogisms*. Amsterdam: North-Holland, 1963.

McCall S., "MacColl". In: P. Edwards (Ed.): *The Encyclopedia of Philosophy*, London: Macmillan, Vol. 4, 545-546, 1967

MacColl H. *Symbolic Logic and its applications*, London, 1906.

Priest G. "What is a Non-Normal World? *Logique et Analyse*, vol. 139-140, 291-302, 1992.

Priest G. "Editor's introduction". Special issue on "Impossible Worlds" of the *Notre Dame Journal of Formal Logic*, vol. 3/1, 481-487, 1998.

Priest G. *An Introduction to Non-Classical Logic*. Cambridge, Cambridge University Press, 2001.

Rahman; S."Hugh MacColl – eine bibliographische Erschliessung seiner Hauptwerke und Notizen zu ihrer Rezeptionsgeschichte". *History and Philosophy of Logic*, vol. 18, 165-183, 1997.

Rahman; S. "Ways of understanding Hugh MacColl's concept of symbolic existence". *Nordic Journal of Philosophical Logic*, vol. 3, 1, 1998, 35-58.

Rahman; S. "Hugh MacColl and George Boole on Hypotheticals". In J. Gasser (ed.), *A Boole Anthology*, Dordrecht, Synthese-Library Kluwer, 287-310, 2000.

Rahman S. and Keiff L. "On how to be a dialogician", to appear in D. Vandervecken (ed.), *Logic and Action*, Dordrecht, Kluwer, 2003.

Rahman S. and Rückert H. (eds.) “New Perspectives in Dialogical Logic”. Special issue of *Synthèse*, 127, 2001.

Rahman S. and Rückert H. “Dialogische Modallogik (für T, B, S4, und S5)”. *Logique et Analyse*, vol. 167-168. 243-282, 2001a.

Rantala V. “Urn Models: a new kind of non-standard model for first-order logic.” *Journal of Philosophical Logic*, 4, 455-474, 1975.

Read S. "Hugh MacColl and the algebra of implication". *Nordic Journal of Philosophical Logic*, vol. 3, 1, 1998, 59-84.

Read S. *Thinking About Logic*. Oxford, Oxford University Press, 1994.

Restall G. "Simplified Semantics for Relevant Logics (and Some of their Rivals)", *Journal of Philosophical Logic*, vol. 22, 481-511, 1993.

Routley R, Pluwood V., Meyer R. K. and Brady R. *Relevant Logics and their Rivals*, Atascadero, Ridgeview, 1982.

Wolenski I. "MacColl on Modalities ". *Nordic Journal of Philosophical Logic*, vol. 3, 1, 1998, 133-140.

PROBABILISTIC FEATURES IN LOGIC GAMES

Johan F. A. K. van Benthem
University of Amsterdam
Amsterdam, Netherlands

1. INDEPENDENCE IN QUANTUM MECHANICS

Hintikka's main idea in his recent work on quantum logic is that quantum mechanics deviates from classical mechanics in the dependencies and independencies which it postulates between physical variables like position and momentum. In particular, Heisenberg's uncertainty relation constrains the values that can be taken by the latter quantities, and in an intuitive sense, they both depend on the other. And if this is the crux, *logic* – to be sure, in its liberated *IF* version – might be involved essentially in the transition. We would need quantifier patterns then allowing for *mutual dependence*, whereas the linear quantifier order of standard logic always makes one variable independent from the other. Note that this is the inverse of the original cases for going *IF*, where variables needed to be mutually *independent* – but I will not pursue this asymmetry here. Instead, I will just focus on one new idea that Hintikka is led to, which I find of particular interest given the game-theoretic nature of *IF*-semantics. Eventually, he advocates the use of probabilistic *'IF functions'* over domains of pure objects, introducing extended objects that are maps assigning probability values to the original objects An example would be the position of an electron, which – in some intuitive sense given by the 'cloud pictures' that we learn as physics students – is a probabilistic mixture of classical clear-cut positions. Or more domestically, in addition to Jaakko and me, there might be objects of the form *3/4•Hintikka + 1/4•van Benthem* – even though I do not care to speculate which particular colleague this mixed object might be embodied in. More generally, generalized Skolem functions might map mixed objects to mixed objects, and so on. My aim in what follows is extremely simple. I see this move as tightening up connections with game theory, and indeed, as a way of

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 189-194.

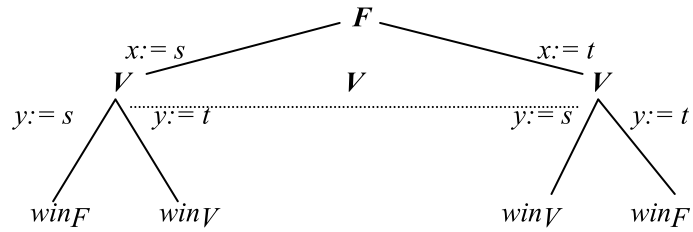
addressing a genuine concern about the game-theoretical foundations of *IF* logic.

2. MIXED EQUILIBRIA IN LOGIC GAMES

As observed in my own chapter in the *Library of Living Philosophers* volume (forthcoming), there is something strange about the connection between logic and game theory found in *IF* logic. Games with imperfect information no longer have the simple determined character that either Verifier has a winning strategy or Falsifier has. As logicians, Hintikka & Sandu then fold this into a 'three-valued' view – with 'True' being the first case, 'False' the second, while an intermediate 'third truth value' stands for all remaining cases. But this uses just one uninformative label to cover what is in fact the most interesting and most common situation of strategic interaction in games, viz. neither player's having a winning strategy. Moreover, we know by the basic theorems of Von Neumann and Nash that, at least over finite models, each *IF* game has Nash equilibria describing optimal play.

Take the example from my chapter, also ubiquitous in the *IF* literature. On a two-object domain *{**s**, **t**}*, the game for the *IF* formula

$\forall x\ \exists y/\mathbf{x}\ \neg x{=}y$ is:



This is just a game-theoretic classic, viz. 'Matching Pennies', as was observed by Rohit Parikh. The later game is not determined, but it has an optimal Nash value *(1/2, 1/2)*, achieved by players using their uniform strategies with probability *1/2*. In other words, the mixed object $1/2\bullet\boldsymbol{s} + 1/2\bullet\boldsymbol{t}$ is the optimal way to go for both Verifier and Falsifier. Or in more biological terms, a population with 50% ***s***'s and 50% ***t***'s is in equilibrium in this game – just as we find stable percentages in other simple games like 'Hawk versus Dove' or the 'Stag Hunt'. Likewise, other mixed equilibria, with other probabilistic weights over pure objects, can be found with *IF* games for different formulas over different object domains. In some sense, these equilibria carry much more refined information about players' powers in a game, and it would be really exciting if this also had logical import. The idea of importing some of this equilibrium structure into logic games, in order to refine logical notions of consequence, has been developed recently for propositional logic in Paul Harren-stein's dissertation *Logic in Conflict* (department of computer science, Utrecht University, SIKS dissertation series 2004-14). There, logical games serve to model agents' differential 'control' over the setting of different variables in a joint task.

3. GAME-THEORETIC SOLUTION METHODS FOR IF GAMES

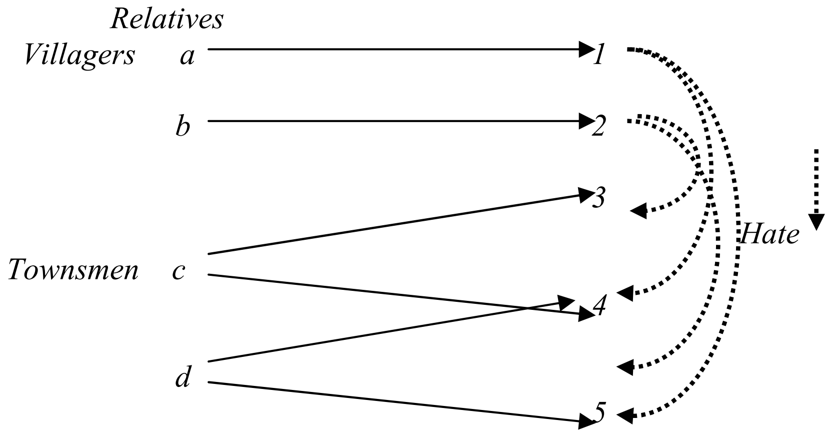
The analogy between logical *IF* structure and game-theoretic solution concepts runs even further. When games have more consecutive moves than the two of $\forall x\ \exists y/\mathbf{x}\ \neg x{=}y$, we can look at their strategic form, listing all strategies, and then search for its structure using any of the known game-theoretic solution methods. The latter need not yield just Nash equili-bria. For instance, 'Iterated Removal of Strictly Dominated Strategies' or even of just weakly dominated strategies leads to a restricted area of strategy pairs that represent some improvement from the viewpoint of rational play, though not necessarily in equilibrium. Again, this structure can be highly suggestive from a logical point of view. Merlijn Sevenster (working paper, ILLC Amsterdam, spring 2004) has defined a 'Weak Dominance Semantics' for propositional languages where he shows how such elimination procedures can lead to alternatives for propositional *IF* logic as investigated by Hintikka & Sandu. But this idea can also be lifted to quantified cases. Purely for the sake of illustration, consider Hintikka's classical example

"Some relative of each villager hates some relative of each townsman".

This is often taken to exemplify the *IF* form

$\forall x\ \exists y\ \forall z\ \exists u/\mathbf{x}\ Hyu$

with quantifiers bounded by the obvious atoms *'villager(x)', 'relative(x, y)', 'townsman(z)' , 'relative (z, u)'.* Here is a situation where this formula is neither true nor false: the formula does not hold with a Skolem function for *u* independent from *x* – but it is true without the slash, and hence it cannot be false in *IF* semantics.



Looking at best strategies for players, it is clear that ***V*** must pick *1* and *2* for *a, b*, respectively, so that her strategies are only the choices as *c, d*. The strategies for ***F*** are only the choices for *x, z.* Tabulating outcomes as usual in game theory, we get:

	V	*34*	*35*	*44*	*45*
F	*ac*	–	–	+	+
	ad	+	+	+	+
	bc	+	+	–	–
	bd	–	+	–	+

Here, + stands for a win for ***V*** and – for a loss. By the way, the full horizontal row of +'s for '*ad*' does not mean that ***V*** has a winning strategy in this game, but rather that ***F*** has a *losing strategy*. Now, iterated removal of weakly dominated strategies removes successive columns or lines as follows (the rationale should be clear):

	V	*34*	*35*	*44*	*45*
F	*ac*	–	–	+	+
	bc	+	+	–	–
	bd	–	+	–	+

	V	*35*	*45*
F	*ac*	–	+
	bc	+	–
	bd	+	+

		V	*35*	*45*
F	*ac*		–	+
	bc		+	–

The end result is that ***F*** should play either *a* or *b*, but always *c*, while ***V*** should always play *5* for *d*, but either *3* or *4* for *c* In an obvious sense, this gives us much more information about the roles of the players, the dependency structure of the game, and of course, their interaction with the shape of the model that we play on. Outcomes of the strategy elimination procedure will shift, e.g., when we change the *Hate*-pattern – and we may have to separate out the role of the different factors: dependency structure, and underlying facts of the model.

My only point with this illustration here is that there is much more game-theoretic content to famous *IF* examples in such settings than just 'not true and not false'. Taking advantage of this might uncover new logical concepts waiting to be invented.

4. PROBABILIZING FIRST-ORDER MODELS

But now back to mixed strategies and probabilistic aspects of solving games. Here are a few more points about the connection. I think there are at least two ways of taking Hintikka's *IF* functions.

The less radical way is in terms of the above mixed strategies. Given the dependency structure of quantum-mechanical assertions, probabilistic objects $\Sigma_{i\in I}\, p_i{\bullet}d_i$ and the associated mixed strategies may be a useful sort of behaviour for Verifier and Falsifier in evaluating such physical statements. This is like the initial role of mixed strategies in game theory: introducing them smoothens existence results for strategic equilibria. More generally, in logical terms, we would be working with our old object domains, but the universe of Skolem functions is probabilized to some extent.

But a second, more radical approach would take mixed strategies *themselves* as new forms of behaviour. Then we need (probabilized) Skolem functions taking these new objects as arguments and values – corresponding to a much more radical probabi-lization of ordinary first-order semantics. In particular, then, one needs to lift the meanings of basic properties and relations to the new probabilistic objects, leading naturally to weighted sums of truth values (*IF* functions of truth values?), and presumably, some sort of many-valued logic. It remains to be seen how far one would have to go for the purposes of analyzing the logical structure of quantum mechanics. But either way, the prospect looks intriguing.

5. LOGIC AND PROBABILITY

Finally, these ideas live on a broader canvas. Logic and probability have met in many ways, from Carnap's inductive logic to probabilistic logics from the 1960s, and then all the way to the 1990s. *IF* logic and its connections with game theory may provide another point of contact here – where it may be no coincidence that probabilistic assertions typically involve dependence and independence of variables. More generally, I feel that the true probabilistic analysis of first-order logic may have hardly begun. Around 1970, Per Lindström proved his famous characterization of first-order logic in terms of its qualitative meta-properties like Löwenheim-Skolem and Compactness. Many people saw this as the 'end of history'. But only a few years later, Ron Fagin proved the Zero–One Law, the first significant statistical property of first-order logic. It says that, for any first-order formula ϕ, the percentage of finite models of domain size n making ϕ true must go to either *0* or *1* with increasing n By now, other intriguing statistical properties of first-order deduction are emerging in computational logic, including physical 'phase transitions' in average-time complexity for proof search. We may be still at the beginning of seeing the true probabilistic side of first-order logic, and exploring *IF* logic in ways that takes probability seriously may provide a powerful vantage point.

ON SOME LOGICAL PROPERTIES OF 'IS TRUE'

Jan Woleński
Jagiellonian University
Kraków, Poland

I think that it is very proper to include a paper about truth in a volume dedicated to Jaakko Hintikka even if it does not deal directly with his highly original ideas about this topic that are related to IF logic. This essay considers the problem of truth in a conservative way, so to speak way. I will focus on properties of the predicate 'is true' derived from very elementary logical insights. It is convenient to start with Frege's view (see Frege 1979a, Frege 1979b) that logic is the science of truth. This thesis should be understood to mean that logic concerns formal principles of truth (the distinction between formal truth and material truth was very common in German philosophy of the 19th century). Putting this in more contemporary terminology, formal truth consists in truth defined by purely structural criteria and transmitted by principles of logical entailment (logical consequence) as codified by logical calculus. These principles always guarantee the truth of conclusion if premises are true; technically speaking, logical entailment preserves truth, that is, transmits it from premises to conclusion.

Is the view that logic is the science of truth the only possibility? Surprisingly enough, the answer is no, because it is not difficult to define the consequence relation as preserving falsehood (see also Woleński 1995). Consider the formulas

(1) (a) $A \wedge B \Rightarrow A$;
(b) $A \Rightarrow A \wedge B$

The first is truth-preserving. Assume that $A \wedge B$ is true. Thus, A is true and B is true. Assume that A is false. Thus, $A \wedge B$ is false too. Summing up, it is impossible for the antecedent of (1a) to be true if its consequent is false. However, the situation is different in the case of (1b). Clearly, A can be true, but $A \wedge B$ is false if B is false, and (1b) is not truth-preserving. However, it preserves falsehood, because if A is false, $A \wedge B$ is false too. Further, $A \wedge B$ cannot be true, if A is false. It means that it is impossible the antecedent of

D. Kolak and J. Symons (Eds.), Quantifiers, Questions and Quantum Physics, pp. 195-207.

(1b) false, but its consequent true. The last statement shows that the formula (1b) is not truth-preserving, but falsehood preserving.

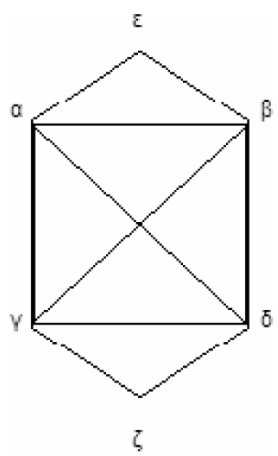
Similarly, the first of formulas:

(2) (a) $A \Rightarrow A \vee B$;
(b) $A \vee B \Rightarrow A$.

preserves truth, but the second one preserves falsehood. Now, (1a) and (2b) are mutually dual (the symbol $\wedge$ in (1a) is replaced by the symbol $\vee$ in (2b)); the same goes for (1b) and (2a). Thus, (1b) and (2b) are examples of principles of dual logic, that is, logic preserving falsehood. The duality of $\wedge$ and $\vee$ is syntactic. It is parallel to the semantic duality of truth and falsehood. This duality goes even further, namely to pragmatics. We can say that if the antecedents of (1a) and (2a) are assertible, the relative consequents are also assertible on purely logical grounds. However, this description fails with respect to (1b) and (2b). Instead we should say that if the antecedents of (1b) and (2b) are rejectable, then their logical consequences are rejectable as well. Thus, assertibility and rejectability are examples of the next duals related to our "alternative" logic. The idea of dual logic can be developed formally in all details, but I will not enter into this question (see Woleński 1995 and the literature quoted in this paper).

Thus, from the theoretical point of view, logic can be based on the falsehood as the distinguished logical value. The dual logic is an exact mirror of the normal logic, that is, based on the concept of truth and assertibility. Why do humans choose truth as the basic semantic concept? The explanation is that truth reveals facts. An additional factor is that logic based on truth accumulates information, while dual logic disperses it. The conjunction is a connective playing an important role in cumulating information. If $A \wedge B$ is assertible, then A and B are assertible too; the reverse link holds as well. The symmetry is broken in dual logic, because if $A \wedge B$ is rejectable, I cannot say whether A or B contributes to this situation. Thus, normal logic smoothly transmits information, but the dual logic contributes to its dispersion. Perhaps this observation enlightens the problem of the genesis of logic. It seems to open the way to a naturalistic and evolutionary account of logic as related to accumulating the knowledge about facts. The most important thing to note in the context of dual logic is that logic is not the science of truth or falsehood, but the science, if one wants to use this label, about the relation of logical consequence.

Another problem concerning the relation of logic and truth focuses on the formal properties of the latter. Theories displaying such properties are called logics of truth (see Turner 1990, Turner 1990a, Von Wright 1999). Several formal aspects of the concept of truth find their precise display in the following diagram (D) (its part determined by the points α, β, γ, δ is similar to the well-known square of oppositions for categorical sentences, usually symbolized as *SaP* (every *S* is *P*), *SeP* (no S is P), *SiP* (some *S* are *P*) and *SoP* (some *S* are not *P*)):



Interpret the Greek letters as follows (I use the same letters for the set of true sentences and truth-predicate; the same concerns falsehood):

α – A is true ($\mathrm{TR}(A)$);
β – $\neg A$ is true ($\mathrm{TR}(\neg A)$);
γ – $\neg(\neg A$ is true$)$ ($\neg\mathrm{TR}(\neg A)$);
δ – $\neg(A$ is true$)$ ($\neg\mathrm{TR}(A)$);
ε – A is true or $\neg A$ is true ($\alpha \vee \beta$; $\mathrm{TR}(A) \vee \mathrm{TR}(\neg A)$);
ζ – $\neg(\neg A$ is true$) \wedge (\neg A$ is true$)$ ($\gamma \wedge \delta$; $\neg\mathrm{TR}(\neg A) \wedge \mathrm{TR}(\neg A)$.

The diagram (D) interprets truth as a modal concept. We have several logical dependencies summarized in

(3) (a) $\alpha \Rightarrow \varepsilon$; $\mathrm{TR}(A) \Rightarrow (\mathrm{TR}(A) \vee \mathrm{TR}(\neg A))$;
(b) $\alpha \Rightarrow \gamma$; $\mathrm{TR}(A) \Rightarrow \neg\mathrm{TR}(\neg A)$;
(c) $\beta \Rightarrow \varepsilon$; $\mathrm{TR}(\neg A) \Rightarrow \mathrm{TR}(\neg A) \vee \mathrm{TR}(A)$;
(d) $\beta \Rightarrow \delta$; $\mathrm{TR}(\neg A) \Rightarrow \neg\mathrm{TR}(A)$;
(e) $\neg(\alpha \wedge \beta)$; $\neg(\mathrm{TR}(A) \wedge \mathrm{TR}(\neg A))$;
(f) $\gamma \vee \delta$; $\neg\mathrm{TR}(\neg A) \vee \neg\mathrm{TR}(A)$;
(g) $\neg(\alpha \Leftrightarrow \delta)$; $\neg(\mathrm{TR}(A) \Leftrightarrow \neg\mathrm{TR}(A))$;
(h) $\neg(\beta \Leftrightarrow \gamma)$; $\neg(\mathrm{TR}(\neg A) \Leftrightarrow \neg\mathrm{TR}(\neg A))$;

(i) $\zeta \Rightarrow \gamma$; $\neg TR(\neg A) \wedge \neg TR(A) \Rightarrow \neg TR(\neg A)$;
(j) $\zeta \Rightarrow \delta$; $\neg TR(\neg A) \wedge \neg TR (A) \Rightarrow \neg TR(A)$;
(k) $\neg(\varepsilon \Leftrightarrow \zeta)$; $\neg((TR(A) \vee TR(\neg A)) \Leftrightarrow (\neg TR(\neg A) \wedge \neg TR(A)))$;
(l) $\alpha \vee \beta \vee \zeta$; $TR (A) \vee TR(\neg A) \vee \neg TR(\neg A) \wedge \neg TR(A)$;
(m) $\varepsilon \vee \zeta$; $(TR(A) \vee TR(\neg A)) \vee (\neg(TR\neg A) \wedge \neg TR(A))$.

The particular points in (3) are justified by appealing to various grounds. (3a), (3c), (3g), (3h), (3i), (3j), (3k) and (3m) are simple applications of tautologies of propositional calculus, for example, the formula $A \Rightarrow A \vee B$ in the case of (32a) and (32c). The proper use of the logical square occurs in the remaining formulas. They establish (I also include (g) and (h)):

(4.) (a) α entails γ (in the traditional vocabulary: γ is subordinated to α) (3b);
(b) β entails δ (in the traditional vocabulary: δ is subordinated to β) (3d);
(c) α and β are contraries, that is, their conjunction is always false (3e);
(d) γ and δ are subcontraries, that is, their disjunction is always true (3f);

α and δ are contradictories, that is, if one is true, the other is false; the same holds for β and γ (3g), (3h);

(f) the disjunction of α, β and ζ exhausts all possible cases (3l).

Since all formulas (3a) – (3m) are theorems, they can be prefixed by the universal quantifier. For example, (l) becomes

(5) $\forall A(TR(A) \vee TR(\neg A) \vee \neg TR(\neg A) \wedge \neg TR(A))$.

Interpret now β as '*A* is false' (FL(*A*)). Since ε is not a theorem of the logic of the diagram (D) (D-logic), that is, logic, which generates principles related to $\alpha - \zeta$), the formula

(6) $\forall A(TR(A) \vee FL(A))$.

is also not a theorem.

Now the problem arises of how to express the fundamental metalogical claims, like the law of bivalence (BI), the law of contradiction (CO) and the law of excluded middle (EM). Typical traditional wordings of these principles in ordinary language are (a) 'there are exactly two logical values applicable to sentences', (b) 'no sentence is true or false', (c) 'no sentence is true and not true', (d) 'every sentence is true or false', (e) every sentence is true or not-true'. We have the following symbolic translations of (a)–(e):

(7) (a) $\forall A(\mathrm{v}(A) \in \{\mathrm{v'}, \mathrm{v''}\}$;
(b) $\neg\exists A(\mathrm{TR}(A) \wedge \mathrm{FL}(A))$;
(c)$\neg\exists A(\mathrm{TR}(A) \wedge \neg\mathrm{TR}(A))$;
(d)$\forall A(\mathrm{TR}(A) \vee \mathrm{FL}(A))$;
(e)$\forall A(\mathrm{TR}(A) \vee \neg\mathrm{TR}(A))$.

The status of these formulas is different. (7c) and (7e) are (after dropping quantifiers) instances of tautologies of propositional logic, namely $\neg(A \wedge \neg A)$ and $A \vee \neg A$; the presence of TR and FL indicates that we are dealing with metalogic. If one adopt the equivalence of $\mathrm{FL}(A)$ and $\neg\mathrm{TR}(A)$, (7b) becomes (7c) and (7d) becomes (7e). However, (D) does not validate this move. It follows that logic validates the identification of $\mathrm{TR}(\neg A)$ and $\neg\mathrm{TR}(A)$. To sum up, (7b) and (7c) are reducible to logical validities by extralogical claims, by purely logical considerations. Eventually, one can treat them as rules of applied logic obtained by suitable identifications related to pragmatic insights, namely that it is sound to equate 'to be false' with 'to be not-true'. The status of the rule (7a) appears as completely different. It says only that the set of logical values consists of two different elements. One can protest that I am unfair to the received meaning of the law of bivalence in logic. And yet the term 'bivalence' suggests nothing more except we have two logical values. It does not suggest how these values are related. Of course, we can make (7a) closer to the rest of (7) by saying that 'there are exactly two logical values, truth and falsehood' or 'there are exactly two logical values, truth and non-truth'. Since the disjunction connective is not exclusive in (7d) and (7e), both formulas can be interpreted as versions of (BI). However, this move makes this principle logically equivalent to (EM). What is perhaps important is that (BI) and (EM) on these readings are purely logical truths on their one reading, but theorems of applied logic under another interpretations. It seems that one and only one interpretation should be adopted as correct.

Let us see what happens when we deny particular formulas included in (7). The related denials as specifies in ((8a) omits the case in which the set of logical values is a singleton, because it immediately leads to inconsistency):

(8) (a) $\forall A(\mathrm{v}(A) \in \{\mathrm{v'}, \mathrm{v''}, \mathrm{v'''}, \ldots\}$;
(b) $\forall A(\mathrm{TR}(A) \vee \mathrm{FL}(A))$;
(c) $\forall A(\mathrm{TR}(A) \vee \neg\mathrm{TR}(A))$;
(d) $\exists A(\neg\mathrm{TR}(A) \wedge \neg\mathrm{FL}(A))$;
(e) $\exists A(\neg\mathrm{TR}(A) \wedge \mathrm{TR}(A))$.

(8b) and (8c) have the same status as (7d) and (7e), (8e) is a contradiction unless we admit paraconsistencies, but (8d) admits many-valueness or truth value gaps (the sentences which are neither true nor false), unless we identify ¬TR and FL. Also in this case we have a loose connection between (8d) and (8a), because each of them could be considered as forming a background for the second. Yet (7e) (or (8b) does not exclude the situation in which we have more than two truth values, because one can argue as follows. Nothing precludes the case that ¬TR covers a whole set of logical values. Thus, TR refers to truth, but ¬TR the set being the complement of {TR}. This reading says only that the set {TR} is disjoint with −{TR}. On the other hand, FL looks like a positive singular denomination. This gives reason to prefer (7d) (or (8b)) over (7e) (or (8c)) and to identify FL with ¬TR (eventually FL, ¬TR and ¬TR), because (BI) has a very precise meaning in this case. Thus, we are entitled to consider (7b) as (BI) and to link (7a) as a companion.

A disadvantage of this move consists in equating (BI) and (EM). In fact, both these principles were considered as expressing the same in history, but it is no decisive point. (7d) misses another point of the traditional interpretation of (BI). On this account, (BI) indicates not only that we have only two values, but also that they are exclusive. Colloquially speaking, this says that every sentence is true or false and no sentence is true and false. This means tha (BI) is the conjunction of (7b) and (7d), that is we have

(9) (BI) $\Leftrightarrow$ (EM) $\wedge$ (CO).

Still another wording of (BI) under this reading is that the division of logical values into truth and falsehood is simultaneously disjoint and exhaustive. Perhaps we should say that (9) gives an exact meaning to semantics associated classical two-valued logic in which the set of tautologies is Post-complete. This reservation has its justification in the fact that one can claim that also other logics, for example, paraconsistent systems are also bivalent in a sense.

The status of (9) is derivative of the character of its components, namely (EM) and (CO). Since (EM) (as (7b)) is not a logical tautology, its conjunction with (CO) preserves this feature. This allows us to maintain that a fundamental principle of classical metalogic is not a tautological claim, although it is closely related to logical theorems. I consider this situation as highly intuitive, because it would be strange to expect that the essential properties of logic are generated by tautologies. The non-tautological character of (BI) stems from the decision to identify the predicate ‘to be false’ with the attribution ‘to be not-true’. The diagram (D) also suggests that ‘to be false’ is co-extensive with ‘to be true not’. In fact, if these identifications are accepted, that is, if (BI) becomes a new principle of the logic of truth, the diagram (D) is automatically reduced to its segment α–β. In particular, we have then the equivalence

(10) $\mathrm{TR}(A) \Leftrightarrow \neg\mathrm{TR}\neg(A)$,

which can be considered as another version of (BI). Decomposition of (10) gives

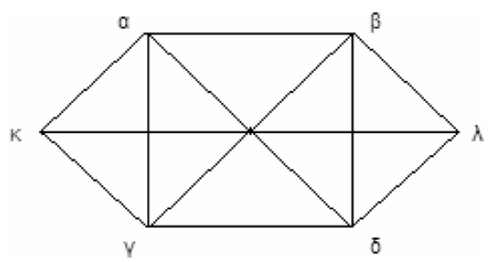
(11) $\mathrm{TR}(A) \wedge \neg\mathrm{TR}(\neg A) \vee \neg\mathrm{TR}(A) \wedge \mathrm{TR}(\neg A)$.

Applying the definition of FL(A), we obtain

(12) $\mathrm{TR}(A) \wedge \neg\mathrm{FL}(A) \vee \neg\mathrm{TR}(A) \wedge \mathrm{FL}(A)$,

which also expresses (BI).

Now I consider the next diagram (D1):



The first interpretation of this diagram consists in understanding

α as 'it is necessary that A' ($\Box(A)$);
β as 'it is impossible that A' ($\Box(\neg A)$;
γ as 'it is possible that A' ($\Diamond(A)$;
δ as 'it is possible that $\neg A$' ($\Diamond(\neg A)$);
κ as $\mathrm{TR}(A)$;
λ as $\mathrm{TR}(\neg A)$.

The dependencies between α, β, γ and δ remains as in (32). In addition, we have

(13) (a) $\alpha \Rightarrow \kappa$; $\Box(A) \Rightarrow \mathrm{TR}(A)$;
(b) $\beta \Rightarrow \lambda$; $\Box(\neg A) \Rightarrow \mathrm{TR}(\neg A)$;
(c) $\kappa \Rightarrow \gamma$; $\mathrm{TR}(A) \Rightarrow \Diamond(A)$;
(d) $\lambda \Rightarrow \delta$; $\mathrm{TR}(\neg A) \Rightarrow \Diamond(\neg A)$;
(e) $\neg(\kappa \Leftrightarrow \lambda)$; $\neg(\mathrm{TR}(A) \Leftrightarrow \mathrm{TR}(\neg A))$.

(13e) minus the definition of FL (A) as TR($\neg A$) is essentially weaker then (EM), because it is reduced to:

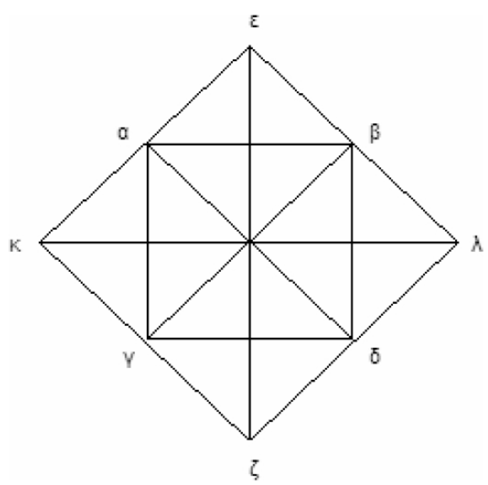
(14) $\mathrm{TR}(\neg A) \vee \neg\mathrm{TR}(\neg A)$.

An essential feature of (14) is that although it falls under the scheme $A \vee \neg A$, it is still weaker than the usual excluded middle, that is

(15) $\mathrm{TR}(\neg A) \vee \neg\mathrm{TR}(A)$,

unless the reductions between ¬TR, TR¬ and FL are adopted. In consequence, (14) is fairly acceptable in logics without bivalence. Using FL(A) in place of TR($\neg A$) sends D1-logic to the classical system. The diagram (D1) shows what happens when the concept of necessary truth is introduced alongside the truth *simpliciter*.

(D1) can be further extended to (D2):



The point ε refers to 'it is necessary that *A* or it is impossible that *A*' ($\Box(A) \vee \Box(\neg A)$), the point ζ to 'it is possible that *A* and it is possible than $\neg A$ ($\Diamond(A) \wedge \Diamond(\neg A)$; another reading it is contingent that *A*). The universal generalization of ε, that is:

(16) $\forall A(\Box(A) \vee \Box(\neg A))$.

This is the view that every truth is necessary or impossible (briefly, but not quite accurately: every truth is necessary). On the other hand, the formula:

(17) $\forall A(\Diamond(A) \wedge \Diamond(\neg A))$

displays the opinion that every truth is contingent. If we say that what is represented in the diagram (D2) governs our thinking about truths in the sense that we have necessary and contingent truths, then the statement '*A* is true' is ambiguous, because it can mean either that '*A* is necessarily true' or '*A* is possibly true'. On the other, hand γ and δ rather play an auxiliary role in this framework and serve as devices to define the concept of accidental truth. Assume that *A* is a possible truth, that is, $\Diamond(A)$ holds. If *A* is also necessary on separate grounds, further consideration stops. Thus, it remains to review what happens when *A* is possible, but not necessary. Under our assumptions, *A* cannot be impossible. Thus, *A* is either possible and true or possible and false (it does not matter whether *A* is false or $\neg A$ true). The latter case immediately implies a contingency of *A*, because we have that $\mathrm{TR}(\neg A)$ implies $\Diamond(\neg A)$. In the former case, if *A* is not necessary, its negation is possible and it is also contingent. Incidentally, these facts are important for the analysis of logical determinism, but I will not enter into details.

We return now to (D1), but in a yeat different interpretation in which:

α refers to $\mathrm{TR}(A)$;
β refers to $\mathrm{TR}(\neg A)$ ($\Leftrightarrow$ $\mathrm{FL}(A)$);
γ refers to $\neg\mathrm{TR}(A)$;
δ refers to $\mathrm{TR}(\neg A)$;
κ refers to *A*;
λ refers to $\neg A$.

All dependencies listed in (13) and holding between α, β, γ and δ remain without any change, except replacing the old symbolism by a new one. The formula

(18) $\mathrm{TR}(A) \Rightarrow A$

is of a special interest. It holds due to (D1)-logic, but its converse, that is:

(19) $A \Rightarrow \mathrm{TR}(A)$

has no justification in the considered interpretation. The heuristic reason is that inferences from the point κ to the point α are invalid, except the case when we conclude $\Box(A)$ from A, provided that A is a tautology. However, a more substantial support is required for the view that (19) does not hold for purely logical reasons. Suppose that A is neither true nor false, that is, it has another logical value or represents a logical gap, that is, a formula which is devoid of any value. It is obvious that A cannot imply $\mathrm{TR}(A)$ in this case. It is not quite clear how to understand A as the an antecedent of any inference in such a case.

The observation about the status of (19) is very important because the conjunction of (18) and (19) gives:

(20) $\mathrm{TR}(A) \Leftrightarrow A$,

This is the well-known scheme of T-biconditionals that plays a fundamental role in many truth-theories, including Tarski's semantic theory. The fact that only a half of (20) has a logical justification is an informal demonstration that T-biconditionals are not tautologies as is sometimes claimed (see Putnam 1985-86 and Woleński 2001 for criticism).

Our diagrams do not generate all principles for T-logic (the logic of truth). If one accepts (see Turner 1990, p. 25) the following formulation for the classical case (note that I focus only on the propositional part and omit the rules of inference):

(TA1) $\mathrm{TR}(A) \Leftrightarrow A$, for all atomic A;
(TA2) $\mathrm{TR}(A \wedge B) \Leftrightarrow \mathrm{TR}(A) \wedge \mathrm{TR}(B)$;
(TA3) $\mathrm{TR}(\neg A) \Leftrightarrow \neg\mathrm{TR}(A)$;
(TA4) $\mathrm{FL}(A) \Leftrightarrow \mathrm{TR}(\neg A)$;
(TA4) $\neg(\mathrm{TR}(A) \wedge \mathrm{FL}(A)$,

an extended (with comparison to D1-logic) T-logic is obtained in which all T-biconditionals become theorems.

(TA2) can be replaced by:
(21) $\mathrm{TR}(A \Rightarrow B) \Rightarrow (\mathrm{TR}(A) \Rightarrow \mathrm{TR}(B))$.

T-logic allows us to clarify a point raised by Wittgenstein. He maintains, in his earlier philosophy (see Wittgenstein 1922), that tautologies and its negations (logical inconsistencies) are senseless. According to him, only contingent sentences have meaning. Let τ be an arbitrary tautology and A an arbitrary contingent sentence. Since tautology is implied by everything, we have

(22) $A \Rightarrow \tau$.

Applying (21) we obtain

(23)(a) $\mathrm{TR}(A \Rightarrow \tau) \Rightarrow (\mathrm{TR}(A) \Rightarrow \mathrm{TR}(\tau))$.

$\mathrm{TR}(\neg A \Rightarrow \tau) \Rightarrow (\mathrm{TR}(\neg A) \Rightarrow \mathrm{TR}(\tau))$.

Since A is contingent, it is meaningful, according to Wittgenstein. Thus, A is true or false; in this second case, $\neg A$ is true. A simple reasoning shows that τ is true independently whether A is true or false. Thus, if something is true, tautologies are true too. Wittgenstein's view is then not correct, at least the T-logic outlined above holds and governs the concept of truth.

Another system of T-logic was proposed by Von Wright (see Von Wright 1999). He proposes the following axioms for the basic truth-logic CS (I omit the rules of inference also in this case):

(TA'1) All tautologies of classical propositional calculus interpreted by formulas of the type $\mathrm{TR}(A)$, their negations, conjunctions, etc.;
(TA'2) $\mathrm{TR}(A) \Leftrightarrow \mathrm{TR}(\neg\neg A)$;
(TA'3) $\mathrm{TR}(A \wedge B) \Leftrightarrow \mathrm{TR}(A) \wedge \mathrm{TR}(B)$;
(TA'4) $\mathrm{TR}\neg(A \wedge B) \Leftrightarrow \mathrm{TR}(\neg A) \vee \mathrm{TR}(\neg B)$.

This logic has the weak rule of excluded middle in the form

(24) $\mathrm{TR}(A) \vee \neg\mathrm{TR}(A)$,

which follows from the formula $A \vee \neg A$. The system CS can be supplemented by adding (20) to its axioms. The result is that non-truth and falsehood are not distinguishable (the strong excluded middle is valid). A further extension arises when (18), that is, the obvious part of the T-scheme is added. These brief remarks show how the diagrams (D1) – (D3) are related to axiomatic approaches to truth-logic.

T-logic has an obvious application for any form of minimalism in the theory of truth, that is, deflationism, disquotationalism or prosententialism, because our diagrams and their axiomatic extensions clearly show that the concept of truth behaves very regularly from the formal point of view. However, it would be very misleading to maintain that minimalism is vindicated by T-logic. First of all, T-logic must be supplemented in a way to block semantic paradoxes, and that can be done by rejecting (19), because the fixed-point lemma does not hold in the resulting system. Other devices consist in isolating paradoxical sentences as representing truth-value gaps. Another problem for deflationism stems from the undecidable sentences or the undefinability of truth. Although these defects can be repaired by additional supplements, the idea of the minimalist truth loses its simplicity, which is usually proclaimed as the main argument for its support.

Finally, let me add that are several other problems related to the problem of the relation between logic and truth. I have only mentioned the problem of more than two logical values. It is related to the question which logic, classical or one of the non-classical variety (many-valued, intuitionistic, paraconsistent) is the "right" logic. As far as the matter concerns intuitionistic logic there is a question of the relation between proof and provability. This shows that, although we can have doubts as to whether logic is the science of truth, certainly there are many interesting issues to be investigated about logic and truth, when both are taken together.

REFERENCES

Frege, G.: 1979a, "Logic", in Frege 1979c, pp. 1-9.

Frege, G.: 1979b, „Logic", in Frege 1979c, pp. 126-151.

Frege, G.: 1979c, *Posthumus Writings*, ed. by H. Hermes, K. Kambartel and F. Kaulbach, tr. by P. Long and R. White, Basil Blackwell, Oxford.

Putnam H. 1985-86, „On Comparison of Something with Something Else", *New Literary History* 17, pp. 61-79; repr. in Putnam 1994, pp. 330-350.

Putnam, H. 1994, *Words and Life*, ed. by J. Conant, Harvard University Press, Cambridge, Mass. 1994.

Turner, R.: 1990, *Truth and Modality for Knowledge Representation*, Pitman Publishing, London.

Turner, R.: 1999a "Logics of Truth", *Notre Dame Journal of Formal Logic* 31, pp. 308-329.

Von Wright, G. H. : 1999, "Truth-Logics", in G. H. von Wright, *Six Essays in Philosophical Logic*, Societas Philosophica Fennica, Helsinki, pp. 71-91.

Wittgenstein, L.: 1922, *Tractatus Logico-Philosophicus*, Routledge and Kegan Paul, London.

Woleński, J.: 1995, "Logic and Falsity", in *Logica '94. Proceedings of the 8th International Symposium*, ed. by T. Childers and O. Majer, Filosofia, Praha 1995, pp. 95-105.

Woleński, J.: 2001, In Defence of the Semantic Definition of Truth, *Synthese* 126(2001), 1-2, pp. 67-90

THE RESULTS ARE IN: THE SCOPE AND IMPORT OF HINTIKKA'S PHILOSOPHY

Daniel Kolak and John Symons

Jaakko Hintikka is more like a scientist or a mathematician than most philosophers in that his greatest contributions derive less from his views than from his results.[60] Hintikka, probably more than any other major philosopher, works at the intersection of traditional philosophical questions and the technical results of mathematical logic, physics, neuroscience and computer science.[61] In this respect, he is sometimes compared with the great American philosopher W.V. Quine. However, there are many significant differences between the two. While Quine has admitted to regretfully having produced no major contribution to real logical theory, Hintikka brings his mathematical creativity to bear directly on philosophical questions, using logical techniques to reach philosophical results that, once they are understood, are as extraordinary as his technical results are indisputable.

There is another related difference between Hintikka and Quine that helps illuminate Hintikka's unique place in contemporary philosophy. Most readers recognize that Quine's philosophy is shaped by his commitment to the all-encompassing metaphysical framework of philosophical naturalism. By contrast, if there is an all-encompassing framework that future readers will associate with Hintikka's work, they will find it as one of the conclusions, rather than as a premise of his many investigations. Hintikka's philosophy is driven principally by what he can prove. This makes reading Hintikka difficult (if satisfying) work. The lack of any obvious and familiar big-picture assumptions is an obstacle facing many of his readers, but it also makes his philosophy less of a defensive action in support of a particular view and more of a progressive accumulation of insights. Unfortunately, Hintikka has been impatient to explain the philosophical import of technically-driven results to philosophers who may not be able immediately

[60] In saying this, we are to a certain extent modifying Hintikka's own occasional claims to have adopted a kind of Kantianism. While, the question of the relation between Hintikka's and Kant's philosophy is quite fascinating and well worth further attention, we do not believe that much light is shed on Hintikka's work merely by seeing it through the lens of Kant's transcendental philosophy. To the contrary, in our view, such an interpretation may (and indeed, often does) serve to conceal some of the more radical aspects of Hintikka's philosophy. Since arguing this point here would require us to defend a reading of Kant in addition to Hintikka, we leave this matter for another occasion.

[61] Thus as Wiebe van der Hoek notes in the *Knowledge, Rationality & Action* special issue of *Synthese*, "Epistemic logicians in computer science acknowledge Hintikka as their origin," (Volume 139, **2**, March 11, 2004, p. v).

to grasp them. (As one of us recently put it: "What the machinery conceals is what the machinery reveals."[62]) Hintikka's reluctance is not due to modesty on Hintikka's part but through the sheer doggedness of his devotion to inquiry. He has rarely engaged in retrospective consideration or explanation of his work as a whole because, as he often says to friends and colleagues, there are too many other interesting things to work on.

Often, Hintikka will end an article or a lecture by introducing a new line of inquiry or a new set of open questions rather than settling on some easily digested philosophical conclusion. Hence, our task in this essay is to make the philosophical import of his work as clear as possible. It is difficult to do justice to Hintikka's results, while at the same time making them accessible. Sometimes we have found that by beginning with some analogy to a point in the history of philosophy, some remark from Aristotle or Leibniz on modality, Peirce on language or quantification, we can see more clearly what Hintikka might be up to. Our strategy in this essay is to begin in the relatively familiar territory of the history of philosophy, mentioning the way some of Hintikka's technical insights figure into his reading of history, before embarking on a more detailed exposition of some of those results.

Section One begins by sketching some of the main features of his approach to the history of philosophy. We can only examine a selection of prominent cases where his technical work is set in dialogue with his interpretation of his predecessors. Hintikka's reading of Aristotle's logic is probably familiar to most philosophers, but we will also discuss some of his less famous studies of Newton, Hume and the Bloomsbury Group. These are likely to surprise readers familiar with Hintikka's work and are, in some ways, more representative of his approach to history than, for example, the more prominent work on Descartes' Cogito argument. The historical topics we discuss in Parts One and Two are (with one exception) those which we can directly connect to our more technical overview of Hintikka's work later in this essay. We hope thereby to reinforce our claim that Hintikka's historical discussions are by and large continuous with his technical work. We hope also to show that his systematic work illuminates, and in turn is illuminated by, his forays into the history of philosophy.

Hintikka's contributions to the philosophy of language have been quite prominent. However, to get a clear sense for the general import of his work in this field, it is useful to see the interplay of historical scholarship and technical investigation. His views on language and ineffability are a clear case where he mixes a reading of the history of early analytic philosophy with a set of theses concerning the nature of logic and semantics. In Section

[62] Daniel Kolak, *On Hintikka* Belmont: Wadsworth.

Two we examine his famous argument for the language as calculus view and discuss some of its implications.

In the third section of this paper we provide an extended discussion of Hintikka's epistemic logic. In addition to outlining his seminal early work from the 1960's, we also include a discussion of his more recent second-generation epistemic logic as well as his interrogative approach to epistemology generally. Much of this is, of course, difficult to discuss properly since his more recent views on epistemological questions constitute something of a moving target. Hintikka, after all, is still developing the implications of the interrogative approach and the new logic for himself. Nevertheless, we take a stab at introducing readers to this new line of investigation and look forward to seeing Hintikka present his new ideas in print in the near future (e.g., his forthcoming *Socratic Epistemology*[63] with Cambridge University Press).

Section Three also presents the dispute between Kripke and Hintikka over the nature of trans-world or cross- identification. There has been a great deal of confusion in the ensuing discussions of these questions over what precisely is at stake in their disagreement. Much of this debate has taken place by proxy, which makes the situation muddier than it ought to be. We try to show here the extent to which Hintikka's work in epistemic logic encroaches on territory occupied by metaphysicians in order to untangle an unnecessarily confused discussion.

Section Four provides a relatively straightforward, what is quickly becoming the standard, introduction to Hintikka's independence-friendly (IF) logic. In Section Five, we provide a similar overview of Hintikka's game theoretical semantics (GTS). This leads us to an analysis of the implications of GTS and IF logic for set theory and in particular for the all-important notion of truth in Section Six. We conclude in Section Seven with something of an open vista. Here we consider Hintikka's recent claim to have grasped the true logic of quantum theory. This aspect of Hintikka's work holds great promise, but raises almost as many questions for us as it answers.

Following the pattern of Hintikka's own essays, we will not close this overview with some definitive statement of Hintikka's place in the history of philosophy, or of the lasting importance of his work. This would be both premature and out of character for an approach to philosophy that never rests on its many laurels. It would be especially premature, given many exciting recent developments and the number of unfinished projects that Hintikka has embarked upon. This essay is meant to encourage readers to appreciate the character and quality of Hintikka's philosophy, to suggest some possibilities

[63] This is Hintikka's working title as we go to press.

for future work, to provoke some critical reaction and to highlight the many deep and interesting open questions that Hintikka's work poses.

1. HINTIKKA ON THE HISTORY OF PHILOSOPHY

Hintikka has been criticized for engaging with the work of great historical figures as though they were his contemporaries. There is a sense in which the charge of anachronism inadvertently gets to the heart of what is most interesting about his historical investigation. Critics are right to remind us that Hintikka's approach to the arguments and ideas of his predecessors bears little resemblance to what usually falls under the rubric of "the history of ideas." Rather, his historical inquiry is unabashedly continuous with his purely conceptual work. The refinement of central concepts and methods in the history of philosophy is integral to Hintikka's program of redrawing the traditional notions of analysis, induction, intuition and the principle of plenitude to name but a few. Especially noteworthy are his interpretations of Plato, Aristotle, Descartes, Leibniz, Peirce, Husserl, Hilbert, Wittgenstein, Tarski, and Gödel, about whom and from whom he has drawn important insights.

Traditionally, inquiry into the history of ideas involves analysis of the context and content of technical terms, their corresponding concepts, and the role they play in the views of the particular schools or traditions of a period. For instance, historians might track the social and moral influences that acted on certain key players or they might work to understand the role of key ideas against the intellectual backdrop of philosophical periods and movements.

By contrast, Hintikka is less an historian of ideas than an expositor of the development of ideas *qua* ideas, viewed quite independently of the particular philosophers with whom they are associated or the historical events within which or out of which they can be viewed as emerging. Hintikka's approach is premised on the idea that the history of philosophy is practiced most fruitfully by philosophers. While one can disagree with Hintikka as matter of historiographical principle, it is difficult to deny that when great philosophers read the history of philosophy *qua* philosophers, it is likely to lead to interesting results.

A revealing example of Hintikka's approach to the history of ideas is his interpretation of the origins of formal logic itself. For Hintikka, Aristotle's logic is the result of an investigation into the nature of questioning and specifically of a reflection on the nature of Socratic *elenchus*. In Socratic *elenchus* answers are (at a certain point) clearly necessitated by the interlocutor's response to earlier questions. In some sense, according to Hintikka, the necessity of an inference originally derives from its place within an interrogative context. This sequence of necessary answers to questions can be represented formally via Aristotle's syllogistic logic and, according to Hintikka, this was precisely the original purpose of the syllogism. The syllogism then is a notion that appears as part of a general theory of questioning. Essentially, Aristotle saw logical and scientific reasoning as occurring within an interrogative framework rather than as an abstracted process of deducing propositions from premises.

Of course, the interrogative approach to logic has been central to Hintikka's own systematic work for many years and so this aspect of his interpretation of Aristotle is clearly filtered through technical observations in that endeavor. This is not the place to get too far into the details of his reading of Aristotle, however a sympathetic reader can find a great deal of textual evidence in its support. For instance, in *Posterior Analytics* (A vi, 75a 22-27), after having laid out the necessary steps in the process of scientific reasoning, Aristotle seems to confirm Hintikka's claim that even in what appears to be a strictly deductive context, we are still within an interrogative framework:

> Yet one might perhaps puzzle why we should ask questions…when the conclusion is not necessary; for one might as well ask any chance questions and then say the conclusion. [The answer is that] we must ask the question not because what is asked is necessary, but because necessarily whoever says them says them, and says something true if it is true.

Aristotle can be read as emphasizing that not all steps in a scientific questioning process are implied, in a strictly deductive sense. Hintikka draws on similar passages in defense of his interrogative interpretation of Aristotle.[64] It is beyond the scope of this essay to take sides for, or against the interrogative reading of Aristotle's logic. However, as we discuss the details of Hintikka's interrogative approach to logic in later sections, it is worth keeping the good Aristotle in mind. Better historians than the present authors are likely to see that even if Hintikka is only partially correct, it is likely to lead to significant changes in the way we understand Aristotle's philosophy.

[64] See for example his "Socratic Questioning, Logic, and Rhetoric," Revue Internationale de Philosophie 47, (1993), 5-30 and more directly his "On the Development of Aristotle's Ideas of Scientific Method and the Structure of Science," in *Aristotle's Philosophical Development: Problems and Prospects,* William Wians, editor, Rowman and Littlefield, Savage, Maryland, 1996, 83-104. There you will find the textual evidence for the interrogative reading of Aristotle's logic.

Unlike his more recent essays on Aristotle, where Hintikka's own results are enlisted in an effort to rethink the history of philosophy, his papers on Leibniz and Aristotle from the 1960's show Hintikka drawing on the history of philosophy in order to form a clear picture of necessity and possibility. Especially noteworthy in this respect are his "Leibniz, Plentitude, Relations and the 'Reign of Law'," [65] as are his many essays on Aristotle's conception of modality from the 1960's early 1970's.

Returning to cases where Hintikka is applying technical results to historical considerations, we find another important example in his reading of Frege and Russell on the supposedly unavoidable ambiguity of the word "is." By applying game-theoretical semantics to natural languages, Hintikka shows that we do not need to live with this apparent ambiguity. Frege and Russell thought otherwise, which is why they built the machinery necessary to handle the distinction into their logical notation.[66] Hintikka shows decisively that—in spite of Russell's claim that this is the greatest advance in logic since the Greeks—we do not have to distinguish the *is*es of identity, existence, predication and the general conditional (subsumption). In some cases it is quite impossible to make the distinction in any natural way. Different uses of *is* are distinguished not by reference to different meanings of the operative word but by reference to context. Hintikka's systematic approach to the logic of ordinary language reveals that the traditional or received logic of quantifiers from Frege and Russell is not the only possible model of the semantics of natural language nor is it the most faithful. We will have more to say about this below.

Hintikka's attention to the fit (or failure thereof) between ordinary language and received first-order logic has a number of other important consequences and has served as an important argumentative strategy in much of his work. One prominent case in point is Hintikka's criticism of Chomsky's use of conventional logical form as a representation of the logical form of natural-language sentences. Were Chomsky's account of the nature of quantification in natural language correct, we would be compelled to conclude that no generative methods can fully account for the acceptability of English sentences.[67]

[65] "Leibniz, Plentitude, Relations and the 'Reign of Law'," *Ajatus* **31**, (1969), 117-144.

[66] See Hintikka's paper, "'Is,' Semantical Games and Semantical Relativity," *Journal of Philosophical Logic* vol,. 8 (1979), pp. 433-468, reprinted in *Paradigms for Language Theory and Other Essays*, vol. 4 of his *Selected Papers*.

[67] See for example "Quantifiers in Natural Languages: Some Logical Problems II," *Linguistics and Philosophy* **1**, (1977), 153-172, and "Quantifiers in Logic and Quantifiers in Natural Language," in *Philosophy of Logic. Proceedings of the 1974 Bristol Colloquium*, Stephan Körner, editor, Basil Blackwell, Oxford, 1976, 208-232. Quantifier phrases behave in natural languages rather like other denoting noun phrases. This fact is not accounted for by using the

Returning to the history of philosophy, if one approaches the work of pre-Fregean philosophers with Hintikka's criticism of the ambiguity thesis in mind, it will actually change how one reads one's predecessors. After all, prior to the 19th century the Frege-Russell ambiguity thesis played no significant role. However, since this thesis is built into our received elementary logic, common applications (by most philosophers and historians of philosophy) of the received logic to pre-nineteenth century work are both dubious and misleading. Why then is our received first-order logic still used as grist for the mill of historians' and philosophers' interpretations of early modern, medieval, and ancient philosophies?

The anachronism of the Frege-Russell ambiguity thesis and, with it, our received first order logic, is not itself a condemnation of the application of logical and semantical analysis of the history of philosophy. Hintikka's work from the 1980's makes this clear, when for example he and Jack Kulas developed their game-theoretical semantics for English quantifiers and anaphoric pronouns.[68] This treatment relies in no way on the Frege-Russell ambiguity thesis and strikingly, the resulting theory is remarkably similar to Aristotle's theory of categories.

Many have been puzzled by Aristotle's wavering description of his categories: e.g. as widest genera and as etymological categories. Aristotle himself correlated the distinction by using different question words as labels of different categories; his verb for being, for instance, *einai*, is used differently in the different categories. Hintikka argues that Aristotle did not mean just one of these distinctions but rather, all of the above, because in a natural game-theoretical treatment of ordinary-language quantifiers such different distinctions must go together. Aristotle's theory of categories reveals the logical structure of ancient Greek and his categories are an ontological dramatization of this *Sprachlogik*.

To take another example, consider the *historical* development of the notion of induction, specifically, its role in the history and philosophy of science. For instance, many historians of science have found it strange that Newton claims to have derived or *deduced* the most general laws of physics from particular phenomena. Newton's methodology, after all, is strictly experimentalist, in that it relies on controlled experiments. Once we understand that among Newton's "phenomena" are outcomes of controlled

---

usual first-order logic as one's canonical notation. Hintikka contends that a game-theoretical treatment explains the similarity: each quantifier phrase will denote one particular individual, but only relative to a play of a semantical game. Moreover, the values (denotations) of existential and universal quantifiers are selected by a different player.

[68] Jaakko Hintikka and Jack Kulas, *The Game of Language* Dordrecht: D. Reidel (1983) and Jaakko Hintikka and Jack Kulas, *Anaphora and Definite Descriptions: Two Applications of Game-Theoretical Semantics*, Dordrecht: D. Reidel (1985).

experiments and, moreover, that what Newton means by *induction* is not making inferences from particulars to general laws but, rather, *extrapolation*, *interpolation* and other combinations of partial generalizations, Newton's claim is made quite clear.

This, Hintikka suggests, has a certain resemblance to Aristotle's methodologically similar assumption that we each have immediate access to certain general truths in so far as we are capable of realizing the relevant forms in our own souls. Thus medieval nominalists, who gave up the Aristotelian idea of a full-fledged realizability of universals, did not have to resort to inductive inference; instead, they postulated suitable "innate ideas" in the mind, thereby demonstrating how it is possible to make up for a paucity of available answers to a given question by strengthening our initial premises. Hintikka's contention is that the problem of induction became a problem as such only after both the metaphysics of forms and innate ideas were discarded.

Not only was "Hume's problem" not a problem before Hume, the reason Hume had a problem to begin with stemmed from a misunderstanding of the nature of the experimentalist methodology in Newton's system. According to Hintikka, Newton did not rely on inference from particulars to general laws. Rather, his methodology presupposed the generalizations and consisted in the extrapolation, interpolation, and integration of already reached partial generalizations. Newton's notion of induction is a quantitative version of Aristotle's puzzling notion of *epagoge*.

Hintikka's historical work is not restricted to the philosophical literature in the narrow technical sense, but includes *belles-lettres*, theology, and aesthetics. This point is easily overlooked because Hintikka is known and admired for looking at the history of ideas from the vantage point of logic and epistemology. Nevertheless, both in his lectures and in a few of his publications, his broader attention to the role and evolution of philosophical ideas outside technical philosophy is revealed. Consider, for instance, his essays on the Bloomsbury intellectuals, whose titles alone reveal quite a bit of the story: "The Longest Philosophical Journey: Quest of Reality as a Common Theme in Bloomsbury" (1995), and "Virginia Woolf and Our Knowledge of the External World" (1979).

Moore and Russell claimed, famously, or infamously—depending on your metaphysical presuppositions—that we *do* have direct access to reality in virtue of the fact that in an experience we can, at least in principle, distinguish the experience as an event in your consciousness, from the object of this experience. The object experienced is not merely subjective. Rather, it belongs, or better to say *is part* of reality in Moore and Russell's view. What, then, are the "objective objects," given to you in different kinds of experience? Hintikka explains a parallelism between, on the one hand, the quest by Moore and Russell of the objects of perceptual experience and, on the other hand, the search by Bloomsbury's art theorists for the basic objects of aesthetic experience.

This search is illustrated by Rickie, the protagonist of E. M. Forster's *The Longest Journey*. Forster's novel opens with a parody of sophomoric Cambridge philosophy undergraduates questioning the reality of external objects: does the world exist when I do not perceive it? Does the cow? As the novel unfolds we discover that this is in fact the theme of the novel; Rickie's story is a prolonged quest for immediate contact with fellow humans and the world, in brief, a quest for reality. What Rickie hopes to avoid is the stultifying effect of conventional social norms and institutions, including conventional marriage and family life, which separate him from others. Forster sometimes described the effect of the kind of marriage that Rickie manages to avoid as being like an "astonishing glass shade" that falls between the couple and the world. Rather than facing the doomed longest journey towards death in an unhappy marriage – to echo the original home of the phrase in Shelley's *Epipsychidion* – Rickie rejects exclusivity in favor of immediate and unrestricted connection with other people.

The members of Bloomsbury were desperate to 'connect' with the world without the intrusion of any kind of mediating factors. Hintikka suggests that knowledge by acquaintance, in Russell's sense, has the same basic character. When Rickie reaches his goal, he finds that life has a new and refreshed meaning, "Because, as we used to say at Cambridge, the cow is there. The world is real again. This is a room, that is a window, outside is the night —." This sentence, Hintikka points out, is almost a paraphrase of G.E. Moore's (in)famous "proof of the existence of the external world" before the British Academy where he held up his hands and said, "This is a hand that is a hand, hands are external objects, hence the external world exists." The search for the objects of immediate awareness is part of the Bloomsbury Group's overall quest for authenticity and immediacy.

Likewise, in his other "Bloomsbury" paper, Hintikka reveals parallels between Russell's construction of the physical world in *Our Knowledge of the External World*, consisting in the experiences of real and possible observers, and Virginia Woolf's construction, through her fictional characters' stream of consciousness, of fictional worlds. These and other such essays by Hintikka are examples of how his rendering of philosophical problems, ideas and concepts reaches well beyond the narrow limits of technical philosophy.

As we turn in earnest to some of the technical details of Hintikka's work in logic and semantics, we will return to some of the historical claims touched upon here. For now, we are merely pointing to some of the highlights of Hintikka's conceptual engagement with the history of philosophy. We have, of course, left most of Hintikka's historical work out of our story. In fact, we have omitted his two most prominent historical studies, namely, his performative reading of Descartes' Cogito argument and his extensive work on Wittgenstein.[69] However, a comprehensive survey of this kind is well beyond the scope of this essay.

2. NO EXIT? HINTIKKA AND THE LIMITS OF LANGUAGE

Hintikka's view of the nature of language is informed by a significant distinction between two contrasting views of the relationship of language, reality and human knowledge. While the distinction goes back to Leibniz's contrast between two different projects in logic, namely, *lingua universalis* vs. *calculus ratiocinator* it was articulated in its most influential modern form by Jean van Heijenoort in his paper, "Logic as Language, Logic as Calculus."[70] Unlike Leibniz and van Heijenoort, Hintikka calls these two contrasting views either *language as the universal medium* vs. *language as calculus* or sometimes the idea of the *universality of language* and of the *model-theoretical view of language*.

These terms are anything but self-explanatory. By "universality" in e.g. "language as the universal medium," Hintikka does not mean some universal features of actual languages. He means, rather, a kind of "inescapability." For the universalist, language is an "iron curtain" between reality and us. We cannot avoid the medium nor can we change it by means of language for the simple reason that everything we say already presupposes the meanings of our language. We thus cannot by-pass the iron curtain and, as it were, speak to what is on "the other side." As such we are simply incapable of seeing how language is related to nonlinguistic reality.

Readers should recognize immediately the Kantian and Wittgensteinian quality of this universalist view. Symptomatic of the universalist conception of language are, for instance:

1) the continuing fascination in certain philosophical (and broader) circles with the notion of ineffability
2) the rejection of metaphysics as nonsense

and

3) the failure of the broader philosophical community to recognize the usefulness of model theoretic techniques in philosophy.

[69] See e.g. Chapter 8, "Hintikka's Wittgenstein," in Daniel Kolak, *On Hintikka*.

[70] van Heijenoort, Jean, 1967: 'Logic as language and logic as calculus.' *Synthese*, vol. 17, pp. 324-330.

The core of the received universalist conception is the view that the semantics of a language is inexpressible in that language. And because meaning relations of a language are inexpressible in that same language, the crucial semantical concept of *truth* is indefinable. That is, according to idea of language as the universal medium, the notion of truth applied in your working language cannot be defined in that language. Consequently, universalists have great difficulty accepting any sort of correspondence theory of truth. Just as seriously, a universalist cannot describe how meaning relations of his or her language might systematically vary. Since the fundamental idea of model theory is the study of what happens as a consequence of such variation, we can see why, according to this view, model theory has little to contribute to the philosophy of language. For universalists, there simply cannot be any systematic model theory for ordinary discourse. We thus cannot speak about any but our actual world in our language, since trying to speak about some other possible world would presuppose a linguistic shift in the references of our expressions. Consequently, we have to speak as if only the actual world were relevant to our language and its semantics. Nor is there a place within a universalist position for the notions of metalanguage or metatheory. This is what Hintikka dubs the "one-world view."

For a universalist, then, logical truths are truths about the actual world, not about all possible worlds, as Leibniz or Carnap supposed. Russell expressed the same point by saying that the truths of logic are as much about the constituents of reality, i.e., the actual world, as are the truths of zoology, the only difference being that they are a good deal more abstract than are the birds and bees.

Thus, to take another example, Wittgenstein defends the ineffability of semantics in the *Tractatus*, without subscribing to the one-world-view. What he does is to adopt instead the lesser, but not unrelated, view that when we speak of different states of affairs we are nevertheless in each case speaking of the same objects, in so far as all possible states of affairs consist in the same simple objects, the same "substance." Tarski similarly showed us the means for defining truth for explicit first-order languages using a richer metalanguage, providing us with a model theory for such languages while at the same time denying the possibility of our ordinary, "colloquial" language having any consistent notion of truth

Since, for the universalist, the semantic aspects of language cannot be discussed in language and hence cannot be theorized about, the universalist is forced to cultivate a syntactical, i.e., purely formal, study of language. This, in spite of the fact that some universalists, most notably Wittgenstein, suggest definite relations between language and reality. These relations, a proponent of the language as calculus view could argue, make it possible for us not only to speak, but to provide us with something to speak about, within well defined limits. As such, they surely form the basis for precisely the kind of theorizing that the universalist wishes to block.

The alternative view, "language as calculus", can be understood as a view which embraces precisely those lines of inquiry deemed impossible or illegitimate in the universalist view. However it must be kept in mind that by the phrase "language as calculus" we should not be read as endorsing the notion of language as an *uninterpreted* calculus. Rather, according to this view even our interpreted language, like a calculus, is freely re-interpretable.

The definitive source for studying this fundamental contrast between two opposing philosophies of language is Hintikka's second volume of selected papers, *Lingua Universalis vs. Calculus Ratiocinator* (Kluwer, 1997). Here Hintikka demonstrates and explains how the universalist view has dominated analytic philosophy for well over a century, and why it held sway over Frege, Russell, Wittgenstein, Quine and Church. For a while it held in its grips the entire Vienna Circle, as evinced by their preference for what they dubbed the "formal mode" of speech vs. the "material mode." Chomsky's preference for syntax over semantics may well be another case in point. Gradually, however, logicians were inspired by the various advances contained in Gödel's incompleteness theorems to move beyond the universalist bent for the primacy of syntax. Attempts to force even these theorems to purely formal and computational frameworks persisted for a number of years, but eventually the calculus or model-theoretical view has gained more philosophical respectability, while earlier defenders of the then unpopular calculus view, such as most notably Charles Peirce, have of late grown in stature.

Although Hintikka's publications on the universalist vs. calculus views are focused on the analytic and pragmatist traditions, his broader understanding in relation to the so-called continental philosophies are well known by his students and followers, whom he has inspired to build philosophical rapprochement. Martin Kusch, for instance, one of Hintikka's students, has applied Hintikka's distinction brilliantly to illuminate historical differences in the continental tradition. In his *Language as Universal Medium vs. Language as Calculus: A Study of Husserl, Heidegger and Gadamer* (1989), one of the most significant bridges between the analytic and continental traditions of the past several decades, Kusch dramatically illustrates how different philosophical stances toward Hintikka's distinction helped shape the development of phenomenology since Husserl.

The illuminating distinction between language as a universal medium and language as calculus exemplifies Hintikka's fusion of historical and systematic analysis. The two ways of understanding language have clear parallels at the level of what might seem to be dry and abstract logico-epistemological results. By examining the conceptual situation in the technical context, we can arrive at a precise and clear way to understand and take a principled stand on one of the grand themes in Twentieth Century philosophy.

Some of the best evidence for the universalist view was once thought to be Tarski's theorem showing that explicit first-order language can only be defined in a richer metalanguage. Since no metalanguage beyond or above our actual working language exists, it was widely believed that our applied "colloquial language" cannot provide a definition of truth, such that the semantics of our own language is to a great extent bound to be inexpressible. But now Hintikka's IF (independence friendly) logic, as we shall show in Section Four below, has illuminated the reason Tarski's result holds: Tarski restricts his analysis to languages with an arbitrarily restricted logic. As we shall see, by overcoming this artificial restriction on first-order logic, Hintikka's technical advances help establish the case for the "language as calculus" view. The implications of this shift are significant. For example, in "Contemporary Philosophy and the Problem of Truth" (1996), Hintikka contends that the expressibility of semantical concepts such as "truth in the same language," renders hermeneutical approaches to language and thought unnecessary.

Similarly, Hintikka's systematic criticisms of Quine (e.g. "Three Dogmas of Quine's Empiricism,"[71] and "Quine's Ultimate Presuppositions,"[72]) illustrate the broader consequence of Hintikka's perspective for ideas currently central to the work of many leading Anglo-Saxon philosophers. First is the one-world assumption, according to which "the only purpose of our factual discourse . . . is to represent things as they are in this one actual world of ours," which Hintikka's analysis contends is, on the one hand, far too ontologically ambitious and, on the other, too naively realistic. English speaking (and thinking) philosophers who, like Quine, know only the "real world" know of it very little, as Hintikka quotes Kipling's famous lament: "What do they know of England who only England know." Hintikka's serious point, which he makes light of, is that we cannot do justice to our epistemic practice if we insist on using logic as if there were but one all-comprehensive domain of discourse. It is important to point out too that what blocked realistically interpreted modal logics for Quine was none other than this one-world assumption.

Another important Quinean commitment that comes under critical scrutiny is what Hintikka calls the "atomistic postulate." This is the notion that the input of information into an epistemic system will always take the form of particular, quantifier-free truths. According to Hintikka, if we

[71] *Revue Internationale de Philosophie,* 1997.
[72] *Theoria*, 1999.

actually examine the formation of scientific theories or even simple informal claims to knowledge, we will find that the "atomistic postulate" is not only defective but misleading. The problem with the atomistic postulate is that it grossly misrepresents actual scientific practice, where nature's answers to our questions—Hintikka's apt characterization of the experimental method—take the form of results from controlled experiments. The results of controlled experiments, as Hintikka argues persuasively, offer a counter-example to the atomistic postulate, since there is no to express them without including some reference to generality. We will have much more to say about the atomistic postulate below.

The fourth Quinean notion that Hintikka criticizes is the view that logic, in the sense of formal inference relations, plays the role of holding our theoretical structures together. If one drops a purely syntactical conception of logic and cognition, then Quine's web of belief must be made of stronger stuff than mere rules for the transformation of schemata. Hintikka has argued that logical relations between propositions cannot be reduced to formal rules of inference. And in a sense, this lesson can already be drawn from Gödel's incompleteness theorems. Quine's attempt to understand logical inference purely formally or schematically runs counter to the entire model-theoretical tradition in logic. One could reject the model-theoretic tradition but, in doing so, one would need to ignore the fact that Gödel's incompleteness theorems seem to make the model-theoretical approach indispensable.

3. HINTIKKA'S EPISTEMIC LOGIC

Hintikka is best known among philosophers, logicians and computer scientists as the creator of modern epistemic logic. His 1962 book *Knowledge and Belief: Introduction to the Logic of the Two Notions* has served as the basis for all subsequent work in this important field. Originally, epistemic logic simply involved the addition of an epistemic operator K to ordinary first-order logic. The relatively formal nature of this work should not be disconnected from what Vincent Hendricks calls "the epistemological ambition of the early Hintikka."[73]

The semantics of this supplemented first-order logic are modal in nature insofar as to talk about what a person knows is to specify a set of possible scenarios. This space of possible scenarios is divided between those that are compatible with what an agent knows, and those that are not. This is a relation between a knower a in the scenario $w1$ and those scenarios that are compatible with everything the knower knows in $w1$. a knows S in $w1$ iff it

[73] See Vincent Hendricks *Forcing Epistemology*, forthcoming, Cambridge University Press.

is true that S in all scenarios w^* accessible to a from $w1$. w^* is the set of epistemic alternatives to $w1$ for a, they are what Hintikka calls a's knowledge worlds in $w1$. The epistemic operator Ka therefore functions as a universal quantifier ranging over all a's knowledge worlds. So, not only is one's attitude towards the notions of possibility and necessity important to one's view of epistemic logic, but perhaps even more importantly, the behavior and nature of quantifiers becomes appreciable in Hintikka's presentation as one of the most critical topics in the development of epistemic logic.

We will return to some of the details of the epistemic logic below. However, Hintikka's contribution to epistemology is not restricted to the development of a useful formalism. He has begun to rethink all of epistemology in a strikingly simple and intuitive manner. Rather than focus on traditional epistemological debates over various modifications to the justified true belief model, Hintikka has developed an approach that models knowledge-seeking and belief formation as a questioning process. In a sense this approach is not radically new, for it can be thought of as an updated version of the Socratic method of questioning. However, the approach allows analyses and applications in a completely precise manner once we have an explicit logic of questions and answers in place. A completely general logic of this kind has recently been formulated as a part of his "second-generation epistemic logic."

Hintikka is in the process of applying the resulting "interrogative model of inquiry" to different epistemological problems. In a series of papers that will appear within the next year or so, Hintikka will argue for the irrelevance of philosophers' notions of knowledge and belief to the actual processes of knowledge-seeking (See, for instance, his forthcoming "Epistemology Without Knowledge and Without Belief"). According to Hintikka, philosophers would benefit by adopting a more pragmatic approach to epistemological theorizing. We use the term 'knowledge,' he suggests, as an honorific label that we attach to information that we are entitled to act on. Information rather than knowledge is the stuff of epistemology, according to Hintikka and, in this new model, the notion of acceptance replaces that of belief.[74] Additionally in this recent work, Hintikka urges us to revise our view of the varieties and different uses of the notion of information (including its uses in computer science and neuroscience), the presuppositions of questions as revealing the presuppositions of inquiry, the presuppositions of answers as revealing the a priori element in empirical inquiry, the logic of experimental inquiry, the different senses of the notion

[74] See the brief abstract of his forthcoming paper on this topic in the annotated bibliography in this volume.

of induction, and the notion of explanation (including "how possible" explanation).

An especially intriguing application is to the famous theory of cognitive fallacies developed by Tversky and Kahneman. Hintikka has argued that the so-called conjunctive fallacy is not necessarily fallacious at all. He is extending this point to a general refutation of the Tversky-Kahneman theory, including the other alleged fallacies, especially the so-called base rate fallacy and including the Bayesian presuppositions of the theory. Another application of the interrogative model concerns the question whether omitting data in experimental science is always a violation of scientific methods.

Of course, what makes it difficult, if not impossible, to present the full scope of Hintikka's view of epistemology is that the most philosophically dramatic claims of his second-generation epistemic logic have not yet appeared in print. We have only been able to sketch some of that material here from lectures, conversations and some unpublished material. Again, Hintikka's forthcoming *Socratic Epistemology*[75] will provide a detailed and unified presentation of these developments.

This having been said, it should also perhaps be pointed out that, from a historical point of view, Hintikka's epistemological revolution in the making might seem so traditional as to be downright counter-revolutionary. Hintikka sees the entire knowledge seeking enterprise as a related series of questioning procedures put to different sources of information. Scientific knowledge is the quest for answers from nature in the form of observations resulting from controlled experiments. This is what Hintikka means when he calls his "the conception of inquiry as inquiry." Knowledge as inquiry means knowledge resulting from interrogation, modeled after the Socratic *elenchus*. The Socrates of Plato's early dialogues claims he asks people questions because he knows nothing. This usually leads, irony of ironies, to Socrates' interlocutors realizing their own ignorance. Perhaps it's not just misery but ignorance too that loves company. But in Plato's middle and late dialogues *elenchus* ceases to be subtly deconstructive and becomes often not very subtly constructive, as when Socrates strategically interrogates Meno's slave toward the expression of a geometrical truth. The model of knowledge-seeking as questioning is a natural product of the spirit of *elenchus*.

Aristotle's *Topics* and *On Sophistical Refutations*, systematic studies of the Socratic questioning games practiced in Plato's Academy, both used question techniques that included the search of the first premises of different sciences keenly tuned on the winning strategies. Just as every trial lawyer knows that success in questioning a witness depends crucially on being able

[75] Forthcoming with Cambridge University Press.

to anticipate the answers one is likely to receive, Aristotle according to Hintikka is drawing our attention to the art of predicting answers we might get in a questioning game played against various "oracles."

There is a class of answers that any rational person must give, answers that are logically implied by the same answerer's earlier responses. By enumerating the conditions on such answers and their relation to their antecedents, Aristotle discovered systemic relations of logical consequence. In this way Hintikka establishes his view that logic itself originated as result of the study of questioning games. One crucial difference from the traditional Socratic method is that Hintikka's method of questioning requires that the predetermined answers, which he calls logical inference steps, be clearly distinguished from genuine interrogative steps for the simple reason that even if they are responses to questions, what matters is not the interlocutor's identity or attributes but, rather, that the premises occur earlier in the dialogue. As Aristotle put it, *ad argumentum*, not *ad hominem*, is how we must judge our logical inference steps. But now one might wonder why, if the fundamentals of the interrogative approach has been with us so long, why has it not been perfected long ago? The reason is that to use it successfully, one must be armed with an explicit logical theory of questions and answers. No such theory existed before Hintikka's groundbreaking work on the subject.

We do not mean to imply that the logic of questions and answers has not been duly studied. Indeed it has, but without arriving at a satisfactory, fully general, theory. But what might one mean here by *satisfaction*? In this case, satisfaction presupposes solutions to such problems as concern the logical form of questions and the question-answer relation. In other words, there must be distinct parameters, clearly expressible in logical notation, of when a given response is in fact a fully satisfactory answer to a given question? Likewise, *generality* in this case presupposes our being able to analyze all the different forms of questions. That Hintikka has fully solved these problems with a fruitful theory is no less remarkable than the fact that we are presently forced to piece it together from various notes and writings primarily addressed to other subjects. Once again, Hintikka provides no full-scale systematic presentation.

The first step we must take is to approach the logic of questions and answers in view of the obvious truth that they are not statements, whereas our usual logic is one of statements. Here is how Hintikka suggests we solve this problem. We start by noting that questions are themselves primarily and essentially epistemic, insofar as a question expresses the purpose of our coming to know some particular truth. That is why the logical properties of questions is determined, by and large, by their epistemic aim expressed as such by the statement specifying the epistemic state which we want any given answer to bring out. This Hintikka calls the *desideratum* of a particular question.

Consider, for example, the desideratum of the following question:

(3.1) Is Hintikka going to Paris, Helsinki, or Martha's Vineyard?

This question can be translated into the following statement:

(3.2) I know that Hintikka is going to Paris or I know that Hintikka is going to Helsinki or I know that Hintikka is going to Martha's Vineyard.

But of course, 3.2 is but an extraordinarily clumsy way of stating that

(3.3) I know whether Hintikka is going to Paris, Helsinki or Martha's Vineyard.

Now, what Hintikka terms the *desideratum* of the question,

(3.4) Who is the author of *Knowledge and Belief: Introduction to the Logic of the Two Notions?*

is

(3.5) I know who the author of *Knowledge and Belief: Introduction to the Logic of the Two Notions* is.

That is how Hintikka reduces the study of questions to the study of their desiderata, which because they are statements can be studied by using our usual traditional logical methods. Now, let us beware that desiderata differ importantly from their corresponding direct questions in the following crucial way: desiderata contain a subordinate question with "know," "knows," etc., as the governing verb, which means that the logic of questions and answers must be a part of epistemic logic.

As introduced at the beginning of this section, the original feature of Hintikka's innovation that went well beyond the scope of traditional first-order logic was the subscripted operator K_a. This operator corresponds, in ordinary language, to "*a* knows *that*." Hintikka's research program in epistemic logic strives to express other ordinary language constructions with *knows* as the main verb in terms of the K-operator. Unfortunately, because the agent-indicating subscript is not in the scope of the operator, Hintikka's innovative way of writing out the K–operator is potentially confusing. And to merely say, in response, that *it works*, fails to do it justice. It is important to note that the K–operator does not in fact receive its meaning from its counterpart in ordinary discourse. More than that, Hintikka's way of dealing with the meaning of the K-operator is very straightforward and admits of an elegant formal treatment. In Hintikka's view, if you specify what Smith knows, you are thereby specifying the entire class of the *scenarios* compatible with what Smith knows, what Hintikka calls "epistemic b–alternatives" to the actual states of affairs. It will then be true to say that b knows that S if and only if S is true in all those alternatives, which corresponds quite well with what ordinary people mean when they say that someone knows something. His characterization of the key crucial concept of knowledge is explicit and well enough defined to serve as our basis for a full-blown logic of knowledge.

Epistemic logic does not solve all, or even most traditional epistemological problems. In fact, it quite explicitly leaves a number of questions open. This should not be read as a weakness of the formal treatment of the concept of knowledge; on the contrary, it actually helps us maneuver around some traditionally thorny problems, such as defining the class of scenarios compatible with what someone knows, itself tantamount to the problem of defining explicitly the concept of knowledge and related concepts. What Jones believes, for instance, likewise determines and is determined by the class of scenarios, called *doxastic alternatives*, compatible with everything Jones believes. The obvious key difference, of course, is that whereas knowledge is assumed to be true, beliefs need not be. This mirrors other similar sorts of pair relations, such as the necessary condition that the actual world must be one of its own epistemic alternatives but not one of its own doxastic alternatives, the distinction between the notion of information vs. belief, and so on.

Hintikka has not always made these points clear in his work, nor has he always been consistent about what he has said about these similar but different kinds of logic in methodological practice. Oftentimes he seems to imply that what he means is that realistic applications such as are involved in scientific reasoning revolve around epistemic logic rather than doxastic logic or the logic of information. This unfortunately is misleading with regard to his own interrogative methodological model of scientific reasoning, wherein he explicitly leaves open the possibility that some of the tentatively accepted propositions are not true. This means, rather revealingly, that what Hintikka must here have in mind is something quite different from the logic of knowledge. These qualifications and complications go some way towards explaining why Hintikka seems to prefer the term *information* to the term *knowledge*.

Hintikka's scenarios are what in common technical philosophical parlance is known as possible worlds, and the translation of the semantics of epistemic logic into a variant of possible-worlds semantics is rather straightforward. The technicalities do not need concern us here, since the main problems and their solutions are more easily explained using the examples such as, for instance, *wh*-questions as exemplified by (3.4) and whose desideratum is (3.5). We can quite easily express this desideratum in Hintikka's K–notation as follows:

(3.6) $(\exists x)KA(x,k)$,

where A(x,k) is but shorthand for "x is the author of *Knowledge and Belief*." The reason that the subscript K has been omitted is that the particular knower is irrelevant, so that the naked K can be read "it is known that." It is of course assumed that the relevant values of the x are persons. The important point is to understand exactly what (3.6) involves, which is best seen by comparing (3.6) with (3.7):

(3.7) $K(\exists x)A(x,k)$,

which says that it is known that someone is the author of *Knowledge and Belief*, where as what (3.6) says is that it is known of some particular person, x, that it is x who wrote *Knowledge and Belief.* This, clearly, is what ordinarily we mean when we say that we know who the author of *Knowledge and Belief* is. And, clearly, although the meanings of (3.6) and (3.7) are both straightforward, from the point of view of possible-world semantics there is a striking difference between them. (3.7) says that in each epistemic alternative someone wrote *Knowledge and Belief*. (3.6) says that there is some particular individual *x* who in each alternative wrote *Knowledge and Belief*, which presupposes something not presupposed in (3.6), namely, that it makes sense to speak of the same individual in different scenarios or "possible worlds." This brings us to one of the most important points in Hintikka's approach not only to epistemic logic but more generally to any logic whose semantics involves possible worlds, namely, that unless we have somehow been given a principle of cross-identification—a principle that tells when the denizens of two different scenarios or possible worlds are identical manifestations, in other words, of the same individual—we cannot understand such a logic. Moreover, this requirement is relevant only in cases where these principles are not themselves consequences of the principles determining the references of our terms in different possible worlds. Hintikka's important result here is that *such a reduction of identification principles to principles of reference is not possible in our actual conceptual system*. This means, for instance, that proper names—our most firmly targeted singular terms—do *not* fix the identity of their references. If I do not know who Jaakko Hintikka is, there must be scenarios among my epistemic alternatives in which the name "Jaakko Hintikka" refers to different people.

The necessity of cross-identification principles is shown, from a purely formal point of view, by the failure of some of the rules of inference readily found in first-order logic, such as the rule of existential generalization exemplified by an inference to (3.6) from a sentence having the form,

(3.8) KA(f,k)

where "f" is shorthand for, say, "the Finnish philosopher at Boston University." (3.8) says it is known that *Knowledge and Belief* was written by the Finnish philosopher at Boston University. However, if it is not known who the Finnish philosopher at Boston University is, (3.8) might be true while (3.6) is false. In all the relevant alternatives it is true that the Finnish philosopher at Boston University wrote *Knowledge and Belief*. But the Finnish philosopher at Boston University might be a different person in some of the different scenarios, such as for instance Georg Henrik Von Wright, and so since there is no specific person who is known to be the author of *Knowledge an Belief*, in order for us to be able to infer (3.6) from (3.8) we need an extra premise guaranteeing this identity, which can be expressed by (3.9):

(3.9) (∃x)K(f = x)

which says, in English, "It is known who the Finnish philosopher at Boston University is."

The relationships just expressed have a clear counterpart in the theory of questions, answers and their presuppositions. (3.7) is the presupposition of (3.4), the question whose desideratum is (3.6). If "f" is offered as a response to (3.7) so as to make (3.8) true, this satisfies the requirements of the questioner provided that (3.9) is true, and thus (3.9) is what Hintikka calls the *conclusiveness condition* of (3.4), which are the most significant notions in the theory of simple *wh*-questions. The theory of epistemic logic, which enables us to define all these important concepts for simple wh-questions, is already contained in *Knowledge and Belief*, which Hintikka only applied to questions and answers.

The problem nevertheless remains: how do we generalize these notions to other kinds of questions? To explain how Hintikka achieves this generalization, we must first ask: What are the principles of cross-identification? Clearly, they are both complicated and multifarious in real life. Hintikka rejects with counterexamples David Lewis' argument that cross-identification is based on a number of weighted similarity principles; in their jointly authored paper, "Toward a general theory of individuation and identification,"[76] Jaakko and Merrill Hintikka argue that in typical cases, including the identification of physical objects, cross-identification depends on continuity. Additionally, the Hintikkas imply that the kind of mathematics best suited for such cross-identification tasks is the stability theory of differential equations.

This brings us to one of the most tangled and misunderstood aspects of Hintikka's philosophy, namely, his dispute with Saul Kripke over the nature of reference and rigid designation. Hintikka's distinction between principles of reference and principles of cross-identification seem to be directly opposed to Kripke's "new theory of reference." However, it is important to untangle Kripke's famous claim that trans-world identification between possible worlds is implemented by stipulative rigid designations from Hintikka and Sandu's concerns about quantification and reference. Their criticism of Kripke's theory in their 1995 paper, "The Fallacies of the New Theory of Reference,"[77] is directed primarily at Kripke's assumption that quantifiers range over a fixed set of values. The problem, in a nutshell, is that Kripke in effect misses completely the difference between (3.7), wherein x ranges over the individuals of some one possible world, and (3.6), where x "ranges over" only such individuals as can be identified in all the relevant alternatives.

It is worth taking some time to place these issues in some historical context. Over the past three decades, philosophical discussions of identification have followed the metaphysical path mapped out by Kripke in his *Naming and Necessity*. Kripke's starting-point is familiar. Any object is identical with itself and itself alone. No two objects can be identical. True identity statements are true necessarily. According to Kripke, true identity statements holding between names, for example, "Cicero is Tully" are markers of *de re* necessity. If it is true that Cicero is Tully then it is necessarily true, and this necessity stands apart from how anyone happens to come to know the true proposition.

This basic move permits Kripkeans to contend that arguments presented in *Naming and Necessity* have somehow overcome traditional Kantian objections to non-epistemic treatments of identity and have cleared the way for a revival of metaphysical inquiry free from the critical constraints of

[76] Jaakko Hintikka and Merrill B. Hintikka, "Towards a General Theory of Individuation and Identification," in Werner Leinfellner et al., editors, *Language and Ontology, Proceedings of the Sixth International Wittgenstein Symposium*, Hölder-Pichler-Tempsky, Vienna, 1980, pp. 417-22.

[77] Jaakko Hintikka and Gabriel Sandu, "The Fallacies of the New Theory of Reference," *Synthese* vol. 104 (1995), pp. 245-283. Reprinted in Jaakko Hintikka, *Paradigms for Language Theory and Other Essays*, Dordrecht: Kluwer Academic Publishers 1998, pp. 175-218.

epistemology. The necessity of identity is, after all, a straightforward theorem of modal logic and would be, according to Kripke, no matter what the state or sources of our knowledge. Such a view implies that the necessity of identity precedes any particular identification and, more significantly perhaps, that it is possible to examine the implications of some metaphysical propositions apart from all epistemological considerations. Kripke's argument is powerful and its basic premises seem incontrovertible. However, in order to understand how to build upon this metaphysical insight in order to actually conduct an investigation or application of the notion of identity or identification, one has to turn to other sources. Kripke's basic move is brilliant, but at its heart it is extremely thin.

Hintikka and Sandu understand Kripke's basic insight in *Naming and Necessity* as the claim that quantification in a modal or intensional context presupposes identity conditions that do not reduce to descriptive conditions. It is important to be clear about the target of their criticisms. They acknowledge, of course, that what they take to be Kripke's basic insight is undeniable. The theory of rigid designation that follows from the necessity of identity is criticized by Hintikka and Sandu not on metaphysical grounds, but because, by characterizing rigid designation as a relationship between names and objects, Kripke has arbitrarily restricted the sense of what it is for us to identify an individual. In fact, their criticism is intended to show that questions of reference are orthogonal to questions of identification. Once the distinction between reference and identification is established, it becomes easier to understand their criticism of Kripke's restriction on the behavior of quantifiers and the theory of rigid designation that follows from it.

To repeat the basic point already broached above: Kripke understands quantifiers as ranging over a fixed set of values. He therefore excludes the difference between identifying that, and identifying what or who, between saying for example,

'it is known that someone paid Ann,'

which has the form

$$K(x)P(x,a),$$

where x ranges over individuals of some one possible world, and

'it is known who paid Ann,'

which has the form

$$(x)\ \mathrm{KP}(x,a).$$

Here the x will pick out only such individuals as can be identified in all the relevant or accessible knowledge worlds. In the second case, the individual being spoken of is identifiable in all possible worlds that are compatible with the agent's knowledge.

Once we turn to the question of understanding an agent's knowledge one must decide on whether, for example, a Kripkean account of the quantifier is appropriate. One is basically asking whether a distinction of the kind presented above, is worth retaining in one's formal apparatus. It should be obvious that this is a separate matter from the stand one takes on, for example, the metaphysical necessity of identity. One's criteria for deciding between different treatments of the quantifier will inevitably be drawn from some source other than our reflections on the *de re* necessity of identity. While Hintikka and Sandu argue that the necessity of identity is not enough to enforce the treatment of quanitifiers underlying Kripke's theory of rigid designation, Kripke may be able to defend it on other grounds.

The point here is that when we consider how one might go about quantifying-in in epistemic and other modal contexts, it is clear, even in the relatively straightforward example mentioned above, that logical connectives, quantifiers and all the rest are not sufficient for giving an account of the cross-identification for individuals.

In the case of cross-identification, what Hintikka and Sandu have suggested is that once criteria for cross-identification are specified, quantification into modal or intensional contexts becomes manageable via the specification of the relevant set of worlds and the fate of their members. This is precisely the reverse of what Kripke understood his work to have demonstrated. For Kripke, the basic insight that a thing is identical with itself and itself alone, that Nixon is Nixon, (even if he had been named something other than Nixon) is evidence that something like rigid designation is called for. While Hintikka and Sandu would certainly agree with the necessity of self-identity, they do not see this as grounds for the introduction of rigid designators. Instead, they argue that true identifications of the kind that hold any real interest for us are drawn between different ways of specifying the same thing. How one determines the appropriate critieria for such cross-identifications is not a matter for logic alone to accomplish, however, once these criteria are in place, then the ordinary quantificational infrastructure can do all the necessary work. In order to engage in any modal or intensional reasoning whatsoever, one must be able to cross-identify. Since cross-identification is conceptually prior to quantification in a modal context, it cannot be explained without moving beyond the resources of our logic *per se*.

Given its importance, Hintikka has had strikingly little to say about the non-logical principles governing cross-identification. The closest he comes is an hypothesis concerning continuity in the joint paper with Merrill Hintikka discussed above where they outline an account of how one might use the stability theory of differential equations as the mathematical framework for cross- identifications. Such a view contrasts sharply with Kripke's claim that our having stipulated the possible worlds eliminates the problem of trans-world identification. It also contrasts sharply with other attempts to understand what is involved in cross-identification. David Lewis, for example, famously suggested that we cross-identify or more accurately that we pick out counterparts across possible worlds via subjective similarity measures. Of course, Lewis was keen to point out that counterparts are never actually identical with one another. Plantinga too, in a very different way, and in a way opposed to Lewis, is also eager to point out that the denizens of possible worlds are fundamentally different (this time in kind) from those of the actual world.

While the way one understands identity may well be influenced by one's attitude towards metaphysical questions in general, the difference between Hintikka and Kripke over the existence of rigid designators as we have indicated above stems from differences concerning the nature of logic and specifically from differences concerning the nature of quantification. Logical, metaphysical and perhaps even empirical considerations of identity are thoroughly entangled. This entanglement is the site of a range of open-problems for philosophers; however, getting clear on the dispute between Kripke and Hintikka helps us to make some headway on the issue.

Individuation and identification involve us in a mess of problems. However, as we shall see, some of these problems admit of progress. For instance, one reason that there is such a problem of generalizing the treatment of simple *wh*-questions represented by the examples (3.4) – (3.9) is clarified via examination of questions involving several quantifiers. Let us ask:

(3.10) Who is each person loved by?

The desideratum of (3.10) is

(3.11) I know who each person is loved by.

The presupposition of (3.10) is

(3.12) I know that each person is loved by someone.

The logical form of (3.12) is

(3.13) $K(\forall x)(\exists y)L(y,x)$

where "L(y,x)" means that y loves x. Responses to (3.10) are of the form

(3.14) $K(\forall x)L(g(x),x)$

where g(x) is the person who loves x.

Now, let us ask: What is the logical form of (3.11)? This critical question leads us directly to the questions of quantifier dependence and independence that are the sum and substance of Hintikka's IF (independence-friendly) logic, extended in the present case to include epistemic operators as well. The connection is easily seen from looking at the form revealed by (3.11), where the truth-making choice of the lover must be a known person, the same in all my epistemic alternatives, and hence independent of K_I, that is, independent of the choice of any alternative possible world. At first glance, you might think to express this by having $(\exists y)$ precede K. But if that's what you think, then look again: $(\exists y)$ depends on $(\forall x)$, and $(\forall x)$ cannot precede K. Were $(\forall x)$ to precede K, then (3.11) would speak only on individuals known to me! Once you understand this, you can see immediately how problems can be solved by means of Hintikka's slash notation, (this will be explained in much more detail below). For of course the logical form of (3.11) is

(3.15) $K(\forall x)(\exists y/K)L(y,x)$

Likewise, and by the same token, the conclusiveness condition for (3.14) has to be

(3.16) $K(\forall x)(\exists y/K)(g(x)=y)$.

But (3.16) is equivalent both to

(3.17) $K(\exists f/K)(\forall x)(g(x)=f(x))$

and to

(3.18) $(\exists f)K(\forall x)(g(x)=f(x))$

which is analogous with (3.9). Once you see this you can see also quite readily that the form of (3.6) and (3.9) can just as easily be expressed, instead, as

(3.19) $K(\exists x/K)A(x,k)$

(3.20) $K(\exists x/K)(f=x)$

thus showing the treatment of (3.4) and (3.11) to be strictly parallel, the only difference being that while (3.19) and (3..20) have slash-free synonyms, (3.15) and (3.17)-(3.18) do not. One is almost tempted to put it like this: the difference that makes no difference in logic makes all the difference in the world. In any case, what should be obvious to all is that the generalization to all *wh*-questions is such that the general form of the desideratum of a question is expressed by

(3.21) KS

where S is a proposition that is first-order and in the negation normal form except for that some existential quantifiers are slashed $(\exists x / K)$ and so some disjunctions $(\vee/K)$ may be as well.
To now find the presupposition corresponding to (3.21), we omit all the slashes, such that a response to the corresponding question has a form in which we replace each subformula of S of the form

(3.22) $(\exists x/K)F[x]$

in context by

(3.23) $F[g(y_1,y_2,\ldots)]$

keeping in mind that $(\forall y_1)$, $(\forall_2)$,… are all the universal quantifiers within whose scope (3.22) occurs in (3.21), and the conclusiveness condition is but the straightforward conjunction of all statements of the form

(3.24) $K(\exists f/K)(\forall y_1)(\forall y_2)\ldots (g(y_1,y_2,\ldots) = f(y_1,y_2,\ldots))$

Our treatment here can easily be extended to propositional questions as well as to mixed ones simply by replacing some disjunctions $(S_1 \vee S_2)$ in (3.21) by

(3.25) $(S_1(\vee/K)S_2)$

and then treating (3.25) as one would treat

(3.26) $(\exists x/K)((S_1 \&(x=0)) \vee (S_2 \&(x\neq 0)))$

These are just some of the results of Hintikka's logical theory of questions and answers. Not only does Hintikka's theory provide a uniform treatment of all the most general notions concerning questions and answers, it is the main tool of Hintikka's new epistemology. In combination with his second-generation epistemic logic, Hintikka's theory of questions and answers provides both analysts and synthesizers a powerful new conceptual tool that we are now free to use even more generally, as Hintikka has himself used it in collaboration with Ilpo Halonen in their application of logic to the philosophical analysis of *why*- and *how*- questions. Hintikka and Halonen show that to tell why something happens, why it is the case that S, and so on, S must be clearly derived interrogatively from whatever initial premises are available which, in scientific discourse, includes prominently some background theory. A suitably normalized interpolation sentence in the sense of Craig's interpolation theorem is a summary of the argument leading from explanatory premises to the explanandum. If a normalized interpolation sentence exists, it answers the why-question and if not, that is, when the relevant interpolation theorem does not apply, the entire unsummarized argument remains, which is an answer to a *how*-question but not to a *why*-question. Hintikka and Halonen's account of why-questions is revealing from a methodological point of view. It makes use of nontrivial logical results and evinces the relevance of Hintikka's revitalized epistemic logic to other applications.

Some of these applications are already well underway in Hintikka's own epistemological work. His logic of questions and answers allows him to formulate his interrogative model of knowledge acquisition, a novel approach to epistemology. To understand the full philosophical impact of his innovative approach, one must see his interrogative model from the standpoint of *epistemic strategies.* Hintikka understands his interrogative model as a game against nature, or against whatever (or whoever) it is that provides the answers to our epistemic inquiries. He distinguishes two different kinds of rules or principles characteristic of a game. The *definitory* rules define the game. In a game of chess, for instance, the definitory rules tell us which moves are permitted and which not, what "checkmate," "castling," etc., mean, and so on. These rules define the game of chess. If a player makes a move not allowed by the definitory rules, say by moving a pawn three spaces forward, it is not a chess move and the player must take it back. We can thus describe the definitory rules of any game or rule-governed, goal-oriented activity. However, knowing the definitory rules of a game does not mean you know how to play. You must also know what Hintikka calls the *strategic* rules (or principles) of a game. In chess, for instance, you must plan your moves, select the best course of action, make judgments as to which moves will serve you better than others, and so on. These rules are not merely heuristic. They can be formulated as precisely as the definitory rules. This is well explained by the crucial role of *complete strategies* in von Neumann's game theory.

The results of applying Hintikka's distinction to the interrogative "games" of inquiry are striking. First, the standard rules of an interrogative game—the rules for logical inference moves as well as interrogative moves—are definitory. They tell us nothing about what to do in a logical or epistemological game. The rules for making both logical inference moves and interrogative moves merely define our game. For example, the so-called rules of inference in deductive logic are neither descriptive nor prescriptive but merely permissive, in so far as they do not tell us which particular inference or set of inferences we should draw from a given number of potential premises. The rules may tell us which inferences we are allowed to draw, for instance without our in the process of so doing committing any fallacies. But which rules? It is highly misleading even to call these rules of *inference.*

What we need, if our inquiry is going to be successful, is more than the definitory rules of inquiry. We need *strategic rules*. Indeed, the better our strategic rules, the better our inquiry. The best player in a game of inquiry is the player with the best strategy, which corresponds in game theory to what happens where values, i.e., "utilities," are associated not with moves themselves but, rather, with *combinations* of *strategies*, as in von Neuman's game theoretical notion of a *complete strategy*.

Likewise, what determines whether or not some particular inquiry is successful is not well the players follow the definitory rules but, rather, on how well the players play, namely, success depends upon a player's choice of strategy. Now, although it is highly unlikely that applying the definitory rules will by itself even further the aims of an inquiry, in so far as it fails to lead you to the desired information, but it may itself be strategically valuable. For instance, it may open up a new way of information or knowledge acquisition, say by providing presuppositions for questions that could not have been asked earlier. This also reveals something about the task of an epistemologist. You can't try to capture, say, a scientist's epistemic behavior using the definitory rules of logic. You can only do it using the strategic rules of the suitable game. Thus, when some scientific process is modeled using logic, it is not the definitory rules of that game, i.e., the rules of inference, that should correspond to what the scientist does. Rather, the strategic rules of that logic should mirror the strategic rules of a scientist's inquiry. It is for instance simply a category mistake to think that when a scientist is presented with contradictory evidence that somehow the laws of paraconsistent logic will illuminate the scientist's behavior. These logical laws are definitory, not strategic. We should likewise now be able to look at induction in the same light: induction is a process that cannot be captured or analyzed using the inference rules of inductive logic. Induction, too, must be understood with strategies of inquiry.

In the beginning of this essay we said that to understand Hintikka we must look not to his views but at his results. These, then, are exactly the sorts of *results* of Hintikka's work that require us to make radical changes in our approach to epistemology, which the way it has traditionally been practiced of late has been severely handicapped. Even today, most analytical epistemologists are simply ignore the possibility that a revolution may be taking place in our understanding of the relationship between logic and epistemology through Hintikka's groundbreaking insights. Thus, typically, what those contemporary epistemologists who have not understood Hintikka's results do, and that is the majority of them, is to set up some sort of rule-governed processes within which the definitory rules are supposed to mimic the knowledge seeker's behavior.

Hintikka's interrogative model has other similarly strong, albeit more general, implications. Most epistemologists and philosophers of science have taken for granted the distinction between so-called "contexts of discovery" and "contexts of justification." Supposedly, rational logical and epistemological terms could only be useful in contexts of justification, not contexts of discovery. The notion of genuine rules for discovery seems an oxymoron; there cannot exist a logic of discovery, it was often repeated *ad nauseum*, only a logic of justification. Over the last several decades, there have been occasional denials both the distinction and the justification problem, but just about all actual work in epistemology has been surrounding the problem of justification. Even theories of belief change have had little affect, since therein belief change is based not on a rational agent's prospects of acquiring new information but, rather, on what the rational agent knows now. The general consensus has been that because seekers of knowledge make one move at a time, their moves cannot be understood using strategies. Such old presuppositions were built into the still generally accepted hypothetico-deductive model of the scientific knowledge seeking enterprise.

Hintikka's interrogative model changes all of this. In Hintikka's model, a context of pure discovery corresponds to a questioning game in which all the answers are known to be true. According to the received view, this is impossible to deal with from an epistemological point of view. But Hintikka shows that this is the most paradigmatic case of interrogative inquiry, which leads to another extraordinary result. First of all, in the wider sense in which logic is not restricted to deductive logic, it decisively refutes once and for all the claim that there cannot be a logic of discovery. What Hintikka has shown, to put it most simply, is this: there *can* be, because there *is*.

This case of pure interrogative discovery, the "all answers true" case, presents us with a wonderful object for study because it has even in non-leaner cases a clearly delineated structure and, in this case, the aspects of interrogative inquiry that come most into play are partly analogous to deductive reasoning. On paper, the logical step from the presupposition of a question to the conclusive answer looks like a logical inference from a premise to a conclusion. As a result, as Hintikka has shown, we can extend the metatheorems valid in first-order logic to the case of interrogative inquiry.

Hintikka's logic of interrogative discovery can easily be illustrated by asking what the optimal strategies are in the case of pure interrogative discovery. It is extremely difficult to find an absolute, general answer. Even in the limiting case of purely deductive reasoning, it is generally not possible to compute the optimal strategies. Nevertheless, Hintikka shows how nevertheless it is possible to reach an extremely fruitful *relative* answer.

Let us suppose we are in an interrogative game of pure discovery. Suppose we've come to a number of propositions. Then, the question—the crucial, most important *strategic* question—is this:

> Which of these propositions should become the presupposition for the next question?

The counterpart to this question in the purely deductive case is:

> Which of these same propositions should become a premise for the next logical inference?

There is no computational answer, that is, in neither case is there any general mechanical rule for computing the answer. However, what Hintikka shows, is this. The two questions have something incredible in common, namely, the answer to both questions is the same!

With some minor technical qualifications, we can express this astonishing result most simply by saying that in the case of pure discovery the best strategies of interrogative inquiry and the best deductive strategies in the parallel situation are, remarkably, one and the same. This reveals the real role of logic is the game of empirical inquiry. The notion that the secret of all good reasoning lies in "logic" and "deduction," what Hintikka calls "the Sherlock Holmes conception of logic," cannot be true in so far as it refers to the definitory rules of logic, that is, to the usual rules of logical inference. These rules are, necessarily, truth-preserving, which means that they cannot introduce new information to reasoning. And so although they cannot serve as vehicles of discovery, if we switch our focus from definitory rules to strategic rules, the situation is quite different. To the extent that there are any guides, logic in the strategic sense is our guide to pure discovery. Sherlock Holmes, as Hintikka so aptly puts it, was quite right: strategically speaking, Watson, what is truly elementary is that the secret of all discovery lies in logic.

Now for some technical qualifications and explanations. There is no need for us to explain the parallelism between deductive inferences and question-answer steps mentioned above in the case of simple wh-questions, since in that case the use of a proposition of the form

(3.27) $K(\exists x)S[x]$

as a presupposition of a question yields a response in the form of a proposition

(3.28) $K\ S[b]$

where b is the individual specified by the answer. Now, the answer cannot possibly be conclusive unless

(3.29) $K(\exists x/K)(b=x)$

which allows us to substitute for b a universally quantified variable falling within "K"'s scope. This important step, from (3.27) to (3.28), which parallels existential instantiation, allows us to go from

(3.30) $(\exists x)S[x]$

to

(3.31) $S[\beta]$

where β stands for the "dummy name" or the name of "an arbitrary individual," like "John Doe" on a legal form. (3.28) and (3.31) are analogous, and the rest of the interrogative argument will preserve this analogy, which is the basis of the parallelism between deductive and questioning strategies. Now, we can extend this parallelism to more complex cases by generalizing the rule of existential instantiation. The extended form, which allows us to move from a first-order sentence $S_o = S_o[(\exists x)S_1[x]]$, is in the negation normal form. It contains the subformula

(3.32) $(\exists x)S_1[x]$

to a sentence where (3.32) is replaced by a sentence having the form

(3.33) $S_1[g(y_1, y_2,\ldots)]$

where $(\forall y_1)$, $(\forall y_2)$,… are the universal quantifiers within whose scopes (3.32) occurs in S_o. This extension restores the strategic parallelism between deduction and questioning. What we've just shown, in other words, is that we *can* speak about strategic rules of discovery. In the case of pure discovery, strategic rules of discover are closely connected with the strategic rules of deductive logic, which Hintikka shows cannot be recursive.

None of this is meant to suggest that we don't need to study interrogative games that correspond to contexts of justification. We do. In such games, some of the answers we get as a result of our inquiry can be false, which means that the true ones must be sifted out by further questioning. The complexity of such a process may seem beyond reach of our interrogative model. In point of fact, however, the complexity of the process pertains only to the strategic rules of such uncertain interrogative inquiry, not the definitory rules. All we need therefore is to make sure that the inquiry is nonmonotonic, that is, that the inquirer can reject any particular answer, what Hintikka calls "bracketing." The only difficulty is that the knowledge seeker must then also bracket all the steps that depend on the rejected one. Unbracketing, of course, is also a legitimate move.

This general case is more complicated than the pure interrogative discovery case, and here Hintikka succeeds by in effect turning the received view upside down. We can formulate the logic of discovery, and the interrogative logic of discovery is far simpler than the interrogative logic of justification. Moreover, as game theory clearly shows, it is possible to rationally evaluate not just particular moves but entire strategies. Actual scientific inquiry involves both discovery and justification.

For instance, a typical scientific paper presents the evidence leading to particular results. This same evidence is used in the paper to justify the results. The logic of justification cannot therefore be considered on its own, independently of the logic of discovery, because the ultimate goal of epistemological evaluation is to find the strategies used both in discovery and in justification. It might be a good strategy for instance to try and uncover the truth in a given situation by initial reasoning unbacked by strong justificatory evidence, simply because the very discovery of the truth can help in the quest for justificatory evidence. In science discoveries are often made on the basis of sketchy evidence and are then confirmed only with the help of that very discovery. Regardless of how shaky the initial evidence, further investigation might not have been possible without it. This may explain the false appeal of the hypothetico-deductive model. Scientific discoveries do not have to be thought of purely hypothetically; however, they are nevertheless often reached before we have enough evidence to fully justify them.

Thus, Hintikka's new epistemology may serve to redirect the work of contemporary epistemologists. Most of that work nowadays has been with the nuances of the justification of particular beliefs in relation to the available evidence, which requires that one single step in the epistemic process be considered at a time, the one based on that particular body of evidence. Several epistemologists have considered what "warrant" one might have for some particular inference. General rules concerning such particular steps can only be definitory for some "game" of warranted inference. Hintikka, on the other hand, has shown that epistemological evaluation pertains only to strategies, not to the particular moves or to the definitory rules governing them. If he is correct, then Hintikka's epistemological results make much of the current work in epistemology moot. One point is quite clear: Hintikka's line of argument steers epistemology much closer to actual scientific practice and so, one might suggest, it serves to put the work of philosophers back on a progressive course.

Bottom line, the fact is that only up to a point do working scientists rely on a given body of evidence. A working scientist wants more evidence to answer questions that are still open, such as for instance which new experiment might help choose between competing theories, experiments that by and large can be identified only on the basis of the theory being tested and which up until that point typically is not yet backed by a lot of evidence.

Another fascinating implication of Hintikka's interrogative model of inquiry for epistemology and philosophy of science stems from the fact that in his model an interrogative game is not fully defined until one specifies what questions the respondent, be it a human being or nature, is supposed to answer. This allows us to characterize different sorts of inquiry on the basis of the nature of the available answers. The theory of quantificational complexity (or simplicity) of possible answers is one such particularly fruitful classification.

What, then, is a good scientific empiricist to do? The answer, one might think, is to stick to particular propositions under the assumption that empirical inquiry requires that all answers must be in the form of particular propositions, what we discussed in Section One under the heading of the *atomistic postulate*, the view that the world will not tell you what happens always and everywhere but only what happens here and now. Using his interrogative approach, Hintikka analyzes the implications of the atomistic postulate and shows it is indeed equivalent to assuming that the answers nature can provide to our questions are particular propositions, and as we mentioned previously, this has played an important role in epistemology and philosophy of science. Such a restriction imposed on nature's answers to our inquiry implies that we can reach general conclusions, e.g. scientific theories, only in virtue of initial premises that are themselves already general. This means that if we assume the atomistic postulate, we must modify or complement our interrogative model in one of the following three ways:

(i) our model must include strong *a priori* assumptions
(ii) our model must be broadened with additional rules of inference introduced over and above the deductive ones
(iii) we cannot derive general conclusions from the data, not even interrogatively.

There are other possible modifications, such as Larry Laudan's suggestion that we should choose between competing theories by comparing their question-answering and problem-solving power. Hintikka's point is that each option can be seen as motivation of a major tendency in the philosophy of science. The first leads to a rationalistic construal of the scientific method. The a priori assumptions that can serve as the initial premises of inquiry might include such assumptions as the uniformity of nature. The second leads to an inductivist conception of science or to the idea of abduction providing the additional rules of scientific inference. The third option is the hypothetico-deductive model that, filtered (or perhaps we should better say *augmented*) through Hintikka's insights, should now be called the hypothetico-interrogative model.

Hintikka's interrogative model is nothing less than a multiperspectival framework for the comparative study of various apparently incommensurate approaches to scientific inquiry and the insights motivating them. Which is not to say that modifications (i)-(iii) are complete or satisfactory after all. For the consistent empiricist, strong a priori assumptions are unacceptable. Moreover, inductive and other sorts of ampliative[78] reasoning are generally not truth-preserving. Option (ii) requires further explanation as to how such rules lead to actually true conclusions. Instead of cleaning up the really crucial epistemological problems, the hypothetico-deductive model sweeps them under the rug.

Hintikka's primary insight here is that since all these different views are based on the assumption of the atomistic postulate, we should instead of adopting any of them give up the postulate. Only then can we understand the true nature of our actual scientific inquiry as working scientists practice it.

[78] Roughly speaking, the analytic/synthetic distinction can be thought of in terms of the explicative/ampliative distinction, since in an ampliative judgment (or proposition) the predicate adds something not already contained in the (meaning of) the subject-term.

Because of the nonatomic inputs into the scientific process, nature according to Hintikka can provide nonatomic answers to our questions. To find out what such answers are like, we must ask what questions scientists can ask whose answers are logically complex.

Controlled experiments, in Hintikka's view, are themselves best understood as questions put to nature. Rarely is the outcome of a successful controlled experiment a singular datum but, rather, involves the discovery of dependence. For example, we discover how an observed variable depends on the controlled one, a dependence that Hintikka has shown can, from a logical point of view, be expressed only by means of quantifiers. Answers to our experimental questions must therefore be considered as nonatomistic answers to our experimentalist's questions. This means that the motivation of (i)-(iii) disappears. Epistemologists in general and philosophers of science in particular must learn to get along without them. This is another major way in which Hintikka's interrogative approach can revolutionize both epistemology and philosophy of science.

To take the example from the history and philosophy of science discussed in Section One, consider how historians of science have grappled with the problem of how Newton could have claimed to have derived, even "deduced," his general laws from observed phenomena. Since Newton's methodology was experimentalist within the strict sense of controlled experiments, Hintikka observes, Newton's own view of his method should be surprising, once we come to understand, as Hintikka does, that Newton included among his "phenomena" outcomes of controlled experiments. Hintikka shows the sense in which Newton's statement can be seen as literally true, which is consistent with Newton's view on induction. What Newton means by *induction* is not the making of inferences from particulars to general laws but, rather, extrapolation, interpolation and other such combinations of partial generalizations. This is the same sense in which Aristotle, in Hintikka's analysis, thought we have immediate access to certain general truths by realizing within our own souls the relevant forms.

In a similar vein, Hintikka shows how we can compensate for a narrow range of available answers by formulating sufficiently strong initial premises. This, in his view, is what allowed medieval nominalists to avoid having to make inductive inferences even after they gave up the notion that the mind has direct access to the Aristotelian notion of a full-fledged realizability of forms, i.e., universals, with the addition of one sufficiently strong postulate, namely, that God placed into our minds the right innate ideas. It was only after both the metaphysics of forms and innate ideas were eliminated from the then-current canons of inquiry that the problem of induction became, as it were, a *problem*. Thus, to further extend the case in point into the modern era, the reason "Hume's problem" was not a problem as such before Hume, and the reason it became a problem in the first place, is that Hume completely misunderstood Newton's experimentalist methodology. Not only does Hintikka's interrogative model force us to reexamine the conceptual issues in epistemology and philosophy of science, it forces us to rethink and retool the fundamental principles of the knowledge seeking enterprise then and now.

4. A TIMELY REVOLUTION: HINTIKKA'S NEW LOGIC

In his soon-to-be-published autobiography, Jaakko Hintikka tells about his dissatisfaction with most of the current trends in philosophy, especially in logic and epistemology, and about his efforts to reform large parts of the subject. One of the tools he is using in this enterprise is what has been called independence-friendly (IF) logic. Contrary to what this name might suggest, IF logic is not a new branch of logic and not a new "nonclassical" logic. (Indeed, Hintikka has himself recently suggested that a better name would be *hyperclassical logic*.[79]) It is what the traditional basic logic was supposed to be but is not, that is, a general theory of quantifiers and propositional connectives. This received logic is variously known as first-order logic, predicate logic, or quantification theory. It is only a part of the real story of logic, for it overlooks one important aspect of the role of quantifiers. This role is to express actual dependence relations between variables by means of the formal dependence relations between the quantifiers to which they are bound. Once this is realized, it is seen that we cannot represent all possible patterns of dependence and independence among variables in the received logic, which is therefore defective in an important respect.

IF logic differs from the received first-order logic in that all these patterns are representable by its means. IF logic thus marks the first substantial general improvement on basic logic since the days of Frege and Peirce and opens important new avenues for research. Among other things, it puts the

[79] In "Independence-Friendly Logic and Axiomatic Set Theory," *Annals of Pure and Applied Logic* 126 (2004) 313-333, Hintikka writes:
The most important fact about this "new" logic is that in a deeper sense it is not new. It is not just another "non-classical logic." It was a mistake to give it a special name. Or if a nametag is absolutely necessary, the best suggestion I now can offer is *hperclassical logic*. . . . it is the so-called ordinary first-order logic that should be given a special epithet, not IF logic. (Is "dependence-handicapped" logic too abusive?) If a name is absolutely necessary, perhaps IF logic should be called "hyperclassical" in view of its retaining all the classical rules for semantical games.

concept of negation to a new light by showing that in all sufficiently rich languages there are two different negations present.

In his earlier work, Hintikka has shown how IF logic makes possible truth definitions that were earlier thought of as being impossible. Currently, Hintikka is engaged in showing how IF logic with its different ramifications forces us to reconsider the entire foundations of mathematics. He is in the process of showing how all mathematical reasoning can be carried out on the first-order level, that is, without quantifying over any higher-order entities. Among other novelties, Hintikka is engaged in showing how the consistency of elementary arithmetic based on IF logic can be proved by elementary means. This is a partial realization of the grand project of Hilbert's which has mistakenly been thought of as being discredited by Gödel's results. It is also a positive solution of the second one of Hilbert's famous list of open problems in mathematics.

It is of course well known that Kant's philosophical revolution did not include logic. Aristotle's logic according to Kant was as it were a view, if one could even speak of it as such, with no room, that is, no room for improvement. Since then, in case anyone has been asleep for the past century and a half, Aristotelian logic has gone the way of classical physics. Yet, just as there some physicists and unfortunately many philosophers with a shall we say Newtonian attitude toward relativity and quantum mechanics, there is something of an Aristotelian attitude pervasive throughout the knowledge seeking enterprise, mutated into the notion that Frege and Russell have the last definitive word on what generally is called first-order logic, quantification theory or predicate calculus, and generally recognized as the core area of logic, what sometimes is called "elementary logic."

What are the preconditions of the applicability of a first-order language? They of course include prominently the specification of a domain of individuals (e.g. "universe of discourse") over which all the individual variables range and a contextual elimination of all other singular noun phrases (e.g. Russell's "denoting terms").

What makes such a language first-order? Logically speaking, quite simply, the values of the variables are always individuals. When a first-order theory is thus devised for other kinds of entities such as *sets*, they must therefore themselves be reified into individuals. Thus the leading role played by quantifiers among the symbols of first-order logic is, from a semantical point of view, typically explained in terms of their "ranging over" the entire domain of individuals. The reason for this leading role is that dependencies between different variables can only be expressed in a first-order language by the dependence of the quantifiers on each other to which the variables are bound. When we consider, for instance, the sentences

(4.1) Someone loves everyone,

and

(4.2) Everyone is loved by someone,

we can easily see that in the latter the truth-making value (if any) of the variable *someone* depends on the value of the variable *everyone*, while the truth of the claim about the variable *someone* depends importantly on the value of the variable *everyone*. Whereas in the latter the truth-making value of the variable *someone* depends not on the value of the variable *everyone* but, rather, the truth of this claim about the variable *everyone* depends importantly on the value of the variable *someone*. This is most easily seen by showing the respective forms of (4.1) and (4.2), as follows:

(4.3) $(\exists x)((\forall y)(x \text{ loves } y))$

(4.4) $(\forall y)((\exists x)(x \text{ loves } y))$

Notice that in (4.3), $(\exists x)$ does not depend on $(\forall y)$, whereas in (4.4) it most clearly does. Now, the quantifiers that are dependent on a given one, say $(\forall z)$ are the ones that lie in its scope, indicated by a pair of parentheses following the quantifier in question, like so:

(4.5) $(\forall z)\ (\underbrace{\rule{3cm}{0.4pt}}_{\text{scope}})$

All this is well enough familiar to anyone who has but glimpsed any logic textbook and, in any case, until Hintikka's breakthrough was of interest only to logicians. Hintikka has however discovered a remarkable link between these concepts and the ways in which logical concepts serve the purpose of representing reality, for instance when logical and mathematical concepts are used in science. Hintikka asks: What must a language be able to express to be adequate for the representation of reality? Of all the many answers, the most relevant if not the most obvious is that a language must at the very least be able to represent any possible pattern of dependence and independence between variables. This is in fact the launching point of some of Hintikka's most profound work in logic. In traditional first-order logic, the dependence of a variable on another is expressed by the dependence of the quantifier to which it is bound on the quantifier to which the other one is bound. That is, the dependence and independence of variables is expressed by the dependence and independence of quantifiers. Thus in a sentence having the form

(4.6) $(\forall x)(\exists y)\ S[x,y]$

the truth-making value of the y (regardless of whether such an individual exists) depends on the value of x, a relationship expressed by the dependence of the existential quantifier $(\exists y)$ on the universal quantifier $(\forall x)$. Quantifier dependence is expressed in the received quantification theory by the nesting of the syntactical scopes of the different quantifiers. A quantifier $(\exists y)$ depends on $(\forall x)$ in a formula S if and only if it occurs in the scope of $(\forall x)$.

For simplicity sake, let us take S as being in a negation normal form. We can then see the major flaw in the received Frege-Russell logic: all possible patterns of dependence and independence between quantifiers cannot be expressed in it because not all such patterns can be captured by the nesting of scopes, which transitive and asymmetrical and hence incapable of codifying intransitive or symmetrical dependence relations between quantifiers. Many patterns of dependence and independence among variables are therefore inexpressible in a language whose logic is the ordinary first-order logic.

This flaw is built into the formation rules of our received quantification theory. Hintikka has thus taken it upon himself to extend our usual first-order logic so as to remove this flaw. In "No Scope for Scope?"[80] Hintikka shows how such an extension can be carried out most simply by as it were *liberalizing* the way in which we use the parentheses to define the scopes of different quantifiers.

Parentheses serve two entirely different purposes in first-order languages, expressing respectively what Hintikka calls *priority scope* and *binding scope*. On the one hand these parentheses express through their nesting relations the relative priorities of the different quantifiers. On the other hand, they mark the segment of the formula in question where a variable is bound to the given quantifier. Hintikka is perhaps the first logician in history to realize that because there is no reason why these two should always go together in the semantics of natural language these two functions of scope have to be distinguished from each other. In an elegant way, he has shown that this simple distinction solves in one fell swoop the problem of so-called donkey sentences that has occupied theoretical linguists. Consider, for instance, the notorious donkey sentence, "If Peter owns a donkey, he beats it," whose intended logical form is

$$(\forall x)((D(x)\ \&\ O(x) \supset B(x)).$$

[80] "No Scope for Scope?" *Linguistics and Philosophy* vol. 20 (1997), pp. 515-544.

The problem is how the existential quantifier that is the indefinite article in the donkey sentence can be transformed into the universal quantifier, as above. Hintikka shows that

> All problems concerning such simple donkey sentences in fact disappear in one fell swoop as soon as we acknowledge the difference between binding scope and priority scope. All that needs to be done is to assume that the priority scope of *a donkey* comprises only the antecedent of [the donkey sentence above] while its binding scope comprises also the consequent as is spelled out in [the donkey sentence expressed in the logical notation, as obove]. This is eminently natural. ("No Scope for Scope?" p. 26.)

But in formal logic as well, simply by separating these two functions of parentheses and by liberalizing the requirements on the binding scope, we can build a much stronger first-order logic than the received one. As it turns out, however, such liberated use of parentheses is apt to be highly confusing. Hintikka therefore coined a new item of notation, the slash, "/", which serves to express the independence of a quantifier from another one in whose (syntactical) scope it occurs. Thus in a sentence of the form

(4.7) $(\forall x)(\forall y)(\exists z)(\exists u)\ S[x,y,z,u]$

the truth-making choice of z depends on both $(\forall x)$ and $(\forall y)$, and likewise for u, what is sometimes referred to as choices of *witness individuals* that vouchsafe the truth of a sentence in question. In contrast to (4.7), in the sentence

(4.8) $(\forall x)(\forall y)(\exists z/\forall y)(\exists u/\forall x)\ S[x,y,z,u]$

the choice of a truth-making value of z depends only on $(\forall x)$ and the choice of a truth-making value of u depends only on $(\forall y)$. It can easily be seen that (4.8) cannot be expressed in ordinary first-order logic by showing that such a pattern of dependence relations cannot be captured by any linear ordering of the four quantifiers. Since $(\exists z)$ is independent of $(\forall y)$ but dependent on $(\forall x)$, in ordinary first order logic it must be placed after $(\forall x)$ but before $(\forall y)$, and vice-versa for $(\exists u)$. But then there is no adequate linear ordering of the four quantifiers $(\forall x)$, $(\forall y)$, $(\exists z)$ and $(\exists u)$. Thus (4.8) cannot be expressed in ordinary first-order logic without Hintikka's slash notation (or some such device). Consider the sentence

(4.9) Some relative of each villager and some friend of each townsman hate each other.

Here the quantifiers are understood as illustrated by (4.8), the value of "some relative" is only dependent on the value of "each villager" and the value of "some friend" is dependent only on that of "each townsman." Under this reading what (4.9) asserts is that there is a set S of relatives of villagers and friends of townsmen such that

(i) for every villager, some relative of this villager is in S
(ii) for every townsman, some friend of this townsman is in S, and
(iii) every relative of a villager in S hates every friend of a townsman in S and vice versa.[81]

Critics who challenge Hintikka's strong reading of (4.9) claim that the only legitimate reading is one on which (4.9) asserts that conditions (i), (ii) and the much weaker than (iii) condition (iii*) are satisfied: for every villager-relative in S there is some townsman-friend in S such that this villager-relative and this townsman-friend hate each other, and vice versa. But it can easily be seen that (4.8) can also be written as:

(4.8)* $(\forall x)(\exists z)(\forall y)(\exists u/\forall x) \quad S[x,y,z,u]$

Here the dependencies and independencies between the different quantifiers in can also be illustrated by writing it in a "branching quantifier" notation:

(4.8)** $$\begin{matrix} (\forall x)(\exists z) & \searrow & \\ & & S[x,y,z,u] \\ (\forall y)(\exists u) & \nearrow & \end{matrix}$$

[81] Hintikka, "Quantifiers vs. Quantification Theory," *Linguistics and Philosophy* 5 (1974), pp. 153-77. See Gilles Fauconnier, "Do Quantifiers Branch?" *Linguistic Inquiry* 6 (1975), pp. 555-67 and Jon Barwise, "On Branching Quantifiers in English," *Journal of Philosophical Logic* 8 (1979), pp. 47-80. Matti Eklund and Daniel Kolak present a vigorous defense of Hintikka on this point in, "Is Hintikka's Logic First-Order?" *Synthese* 131(3):371-288 June 2002; see, in the same issue, Hintikka's illuminating, "Response to Eklund and Kolak."

Now the same slash (independence) notation can be extended to propositional connectives, both in their relation to each other and in their relation to quantifiers. In the simplest cases, as in

(4.10) $(\forall x)(A[x]\ (\vee/\forall x)\ B[x])$

Hintikka's new notation does not add to the expressive power of ordinary first-order logic. The reason is that since the choice of a disjunct is independent of $(\forall x)$, it might as well be done before the choice of the value of x. Hence (4.10) is logically equivalent to

(4.11) $(\forall x)A[x] \vee (\forall x)B[x]$

In certain other cases formulas with independent disjuncts do not reduce to ordinary first-order logic. This is the case with

(4.12) $(\forall x)(\exists z)(\forall y)(A(x,y,z)(\vee/\forall x)B(x,y,z))$

Now, look back at (4.8)*. What its form reveals is that it is parallel with (4.12), with the disjunction in (4.12) playing the same role as the existential quantifier $(\exists u/\forall x)$ in (4.8)*. And just as (4.8)* is inexpressible in the Frege-Russell notation, so too (4.12) is likewise inexpressible in the Frege-Russell notation.

This is just the briefest sketch of the main ideas of Hintikka's new IF (Independence-Friendly) logic. The reader can fill in the details. Hintikka sees it as a liberated version of first-order logic, free from all the unnecessary restrictions that limit the expressive power of ordinary first-order logic. Hence Hintikka's new logic ought to be called, simply, first-order logic, without qualifications, and the so-called "ordinary," or "classical," first-order logic should bear in its name a restriction, such as *dependence-handicapped* logic or, politically correctly, "independence-challenged" logic.

In Hintikka's new logic we can express types of *dependence* (for instance mutual dependence) that were not expressible earlier and, as such, can be thought of as "unrestricted first-order logic." The special case of mutually dependent quantifiers is especially interesting, exemplified by sentences of the following form:

(4.13) $(\forall t)(\forall x)(\forall y)(\exists z/\forall x)(\exists u/\forall y)((x=z)\ \&\ (y=u)\ \&\ S[t,x,y])$

Such sentences with mutually dependent quantifiers behave differently from other first-order sentences. Their semantics cannot be dealt with in any usual way, not even thought of in the usual way. The familiar way is to think of the meaning of quantifiers as consisting in their "ranging over" a class of values, a way of thinking that underlies both Frege's idea of quantifiers as higher-order predicates and Tarski's way of defining truth by reference to valuations. There is of course something to the "ranging over" notion, in so far as the use of a quantifier obviously presupposes a range of values for it.

But that's hardly the whole story. A moment's thought shows that the semantics of mutually dependent quantifiers cannot be dealt with by means of the "ranging over" idea only. Hence Hintikka's new logic forces us to think of the most basic logical constants, the standard quantifiers, in a new way. As a consequence, we must now distinguish from each other the dependence component and the "ranging over" component in the meaning of quantifiers. This need is seen especially poignantly when (as the case may be in suitable many-sorted first-order logics) two classes of quantifiers range over two different and even exclusive classes of values. These quantifiers can still be dependent and independent of each other without any restrictions. When one internalizes this idea, one sees that Hintikka's new logic sheds some new light, or perhaps one should say some first-time light, on the so-called (one awaits the proper apocryphal story on the origin of this term) *nonlocality* and related (in the familial or category theoretical sense) phenomena in quantum theory and beyond, also discussed in more detail below.

In spite of its many hidden surprises, Hintikka's independence-friendly logic might at first seem to be rather like ordinary ("dependence-handicapped") first-order logic. Indeed, as Hintikka himself points out, several of the "nice" metatheoretical theorems hold in it, including compactness (an infinite set of sentences is consistent if all its finite subsets are), upwards Löwenheim-Skolem theorem (a satisfiable sentence is always satisfiable already in some countable model), and the separation theorem (if the union $(\delta \cup \tau)$ of two consistent sets of formulas δ, τ is inconsistent, they are "separated" by a single formula of ordinary first-order logic using only the common vocabulary of δ and τ). So perhaps we ought to axiomatize the independence-friendly first-order logic and study it in the usual way, perhaps by comparing it with the ordinary first-order logic. But we cannot, for independence-friendly first-order logic does not admit of a semantically complete axiomatization. The set of its valid formulas is not recursively enumerable. In independence-friendly logic there is no way of capturing validity in general by means of provability in a suitable axiomatic system of logic nor capturing satisfiability by means of formal consistency, that is to say, of establishing in general that a set of propositions has a model in which they are all true by proving that no contradiction cannot be derived from them formally. Hence independence-friendly logic cannot be studied exhaustively by deductive-axiomatic means.

This semantical incompleteness of independence-friendly logic offers an example of the consequences of Hintikka's new logic for the foundations of logic and mathematics. For it was in fact the leading idea of Hilbert's metamathematical project to establish the concrete model-theoretical consistency (satisfiability) of mathematical theories by showing their proof-theoretical consistency (freedom from formal contradictions). Such a program is viable only if the logic that is being used is semantically complete. For if it is not, there is always the possibility that a hidden inconsistency might be brought to light by those axioms and rules of inference that we have not yet managed to capture into the net of our inevitably incomplete logic. Hence the unaxiomatizability of independence-friendly logic, together with its fundamental role as our basic operative logic, shows the impossibility of Hilbert's program, usually said to be shown to have been frustrated by Gödel's proof of the incompleteness of elementary arithmetic. In reality, it falters much earlier, not on the deductive incompleteness of elementary arithmetic, but on the semantical incompleteness of our "elementary" independence-friendly first-order logic.

This of course presupposes that something like independence-friendly logic is needed in mathematical theories. But this claim is easily argued, if we want to stay on the first-order level. For instance, the notions of equicardinality and infinity can be expressed in Hintikka's logic, but not in the ordinary first-order logic.

Hilbert and others have tried to get along using the flawed (handicapped) ordinary first-order logic. In this enterprise, they were given false hope by Gödel's completeness theorem for ordinary first-order logic. This is part of what Hintikka undoubtedly has in mind when he calls the effects of Gödel's completeness result on philosophers' way of thinking about logic a *disaster*.

The semantical incompleteness of independence-friendly logic has other remarkable consequences. The entire subject matter of "formal" logic is usually thought of as dealing with formal rules of inference, formal axiom systems for logical truths, and so on. Some philosophers even think that they can characterize the meaning of different logical constants by reference to the rules of inference they obey. All this is now seen to be a half-truth at best. We have to learn to look at logic in a new light, not as a study of formal systems but as a study of arbitrary structures that involves model-theoretical considerations and not only formal ones.

Hintikka has expressed the semantical incompleteness of his new logic in upbeat metaphoric terms by saying that logicians will henceforth need not fear unemployment: there will always be more logical principles to be discovered. What he means is that there is a creative component to the study of logic that had not been suspected before.

When well-informed contemporary logicians hear that independence-friendly logic cannot be exhaustively studied by means of syntactical (axiomatic and deductive) methods, they will not panic. Their first reaction is most likely: Be that as it may, we can study it by semantical means, for instance with the help of Tarski-type truth definitions. But if so, the logicians are in for a second shock. Hintikka's logic cannot be studied by means of Tarski-type truth definitions, because such definitions cannot be formulated for independence-friendly logic. And why not?

The reason, as diagnosed by Hintikka, is again clear-cut, and highly interesting theoretically as well as historically. Tarski's truth definitions are in his own terminology recursive. In our day and age, linguists and philosophers would express the same fact by saying that Tarski's truth definitions are *compositional*. This means that the semantical attributes (like truth) of a sentence (or some other kind of expression) are determined by the semantical attributes of its constituent expressions plus its form, that is, how it is put together from its syntactical constituents. In the case of truth, Tarski cold not formulate truth-conditions for a sentence in a formalized language in terms of the truth and falsity of its component expressions, for those component expressions are open formulas (expressions containing free variables) which are neither true nor false. Hence Tarski had to formulate his conditions in terms of *satisfaction* rather than *truth*. But this does not change the main point. Thus we can understand Tarski's procedure as an attempt to abide by the principle of compositionality.

Hintikka has argued, together with Gabriel Sandu, that the same commitment to compositionality is on what Tarski bases his claim that the concept of truth cannot be consistently used in natural language or "colloquial language," as Tarski calls it. (See their 1999 paper, "Tarski's Guilty Secret: Compositionality.") Contrary to popular belief, Tarski does not base this claim on his theorem about the impossibility of defining truth for an ordinary first-order language in the same language. He bases it on the alleged irregularities of natural language. Hintikka and Sandu argue that the only "irregularity" that really mattered to Tarski is the failure of compositionality in colloquial language.

Hintikka has pointed out that the principle of compositionality is tantamount to the assumption of semantical context-independence. If context-independence fails, quite obviously compositionality fails. Conversely, if context-independence holds, then there is nothing left to determine the semantical attributes of a sentence except its structure and the attributes of the components of this structure.

What Hintikka's new logic shows once and for all is that there can be a logic with fully determined meanings that does not obey the principle of compositionality. For the force of a slashed quantifier in Hintikka's new logic depends on the quantifier which it is independent of and which naturally occurs in the wider context of the quantifier, not within its syntactical scope. Hence the prima facie semantics of independence-friendly logic is not compositional.

Once again Hintikka's new logic destroys widely held opinions among linguists, logicians and philosophers. But if compositionality fails, where do we find a semantics that is capable of doing justice to independence-friendly logic? As it happens, Hintikka already had such a semantics on tap when he came upon the idea of independence-friendly logic, which calls *game-theoretical semantics*, and to which we shall now turn.

5. THE GAME OF THE NAME: FROM LOGIC TO SEMANTICS AND BACK AGAIN

There are many other fascinating and important implications of Hintikka's new logic. One question on many contemporary philosophers minds may run along the following sort of line. Because it is not compositional, Hintikka's independence-friendly logic does not have a Tarski-type semantics. Well, then what kind of semantics does it have? Here the news turns out, perhaps surprisingly, to be old news. For it turns out, Hintikka developed the semantics of IF logic before his discovery of independence-friendly logic itself, while he was working on refining Wittgenstein's views. In particular, one key (unit) idea in the later Wittgenstein is that crucial language-world links are not static relations of reference. Rather, language-world links are created by certain rule-governed human activities that Wittgenstein called *language-games* that, in fact, sustain them. Now, of course these are not games of speaking or writing language, for the simple and obvious reason that the moves are in this case neither speech acts nor any other sorts of language acts but, rather, games in which players belonging to a particular language community interact with the objects that their language allows them to speak about. There are exceptional cases, when a move can be made by saying something, but Hintikka here points out that the primary thing is not what is said—it's not the utterances or other language acts—but, rather, it's the complex of nonlinguistic activities in the context of which language is used. (Not *performative utterances* but *utterable performances*, borrowing Chisolm's pun).

The problem with Wittgenstein's view as formulated is that it fails to tell us what the language-games are like that give our logical words their meaning. Here Hintikka offers one of the most important refinements of Wittgenstein by formulating precise games (in the mathematical theory of games sense) that provides the foundations for the formal semantics of first-order languages. These "outdoor games," as Hintikka calls them, are like Wittgenstein's language-games, in that they are played among the objects of which the language speaks. They are not "parlor games" whose moves consist in verbal challenges and responses. Moreover, and perhaps most importantly, they are not the formal games of theorem proving.

In keeping with game theorists' frequent talk about "games against nature," Hintikka's tournament game consists of two players, the *defender* and the *opponent*, which Hintikka sometimes refers to as the *verifier* and the *falsifier*, sometimes as *the inquirer* and *nature*. With regard to the quantifiers in question, these aspects of the game have also an affinity, or a family resemblance, with Charles Peirce's notion of the *proponent* and the *responder*.[82]

These games define what Hintikka calls *game*-theoretical *semantics* (GTS).[83] The first and primary purpose of GTS is to define and characterize truth for an interpreted first-order language. Now of course such truth must be relative to a *possible world*, i.e., a *model*. The question then is what we have to know about the model to be able to speak of truth within that model. The answer is elementary. We have, first, to be given a domain D of individuals on which a number of predicates and functions are defined such that the truth-value of all atomic sentences and identities is thereby determined. The task for the semanticist, then, just as it is essentially in Tarski-type truth definitions, is that the notions of truth and falsity must be extended to cover all complex sentences. The way that this is done in Hintikka's treatment of GTS is as beautiful as it is simple. Hintikka associates with each sentence S of the interpreted language in question a two-person game G(S). Let us call the players the *verifier* and *falsifier*. Here then are the rules that define what Hintikka calls a *semantical game:*

(G.$\vee$) G(($S_1 \vee S_2$)) begins by a choice by the verifier of S_i (where i = 1 or 2). The game is then continued as in G(S_i).

(G.&) Likewise for G((S_1 & S_2)) except that the falsifier chooses S_i.

(G. E) The game G(($\exists$x)S[x]) begins with the choice by the verifier of a member of D. If the name of the individual chosen is "b," the game is then continued as in G(S[b]).

[82] This was first pointed out by Risto Hilpinen.

[83] Surveyed by Hintikka and Gabriel Sandu in their like-named article in the Elsevier *Handbook of Logic and Language*, by J. van Benthem and A. ter Meulen (1997).

(G. A) Likewise for $G((\forall x)\ S[x])$, except that the falsifier chooses b.

(G.~) G(~S) begins by an exchange of roles of the verifier and the falsifier, as defined by these rules. The game is then continued as in G(S).

(G.A) If A is an atomic sentence or an identity, the verifier has won and the falsifier has lost G(A) if and only if A is true. The falsifier has won and the verifier has lost if and only if A is false.

Thus consider, for instance, the following conjunction. "Snow is white and Hintikka can fly." Following (G.&) a falsifier of that sentence could choose "Hintikka can fly," and then proceed to (G.A) since "Hintikka can fly" is an atomic sentence. Applying (G.A) the falsifier has won since "Hintikka can fly" is not true. And *mutatis mutandis* for the other rules. How then do these rules help us to define truth and falsity? For clearly we cannot say that S is true if and only if the verifier wins, for this does not depend on S alone, but can depend also on the strategies chosen by the two players. (A win so to speak does not tell very much about one's prospects in a game in general if it is due merely to my opponent's stupidity.)

Here if we reflect but a moment what the appropriate definition is becomes crystal clear. In game-theoretical semantics for first-order logic, the truth of S is defined as the existence of a winning strategy for the verifier in the semantical game G(S). By the same token, the falsity of S means in GTS the existence of a winning strategy for the falsifier.

This crucial definition is the most characteristic feature of GTS. If it looks simple and straightforward, it is because, in many senses, it is simple and straightforward. However, as is so often the case in Hintikka's work, there is a great deal of hidden theoretical subtlety. Let us consider, for example, how the definition applies to IF logic. Well, in traditional (handicapped) first-order logic, the game-theoretical truth definition is equivalent to the usual Tarski-type ones. Can it be extended to Hintikka's new logic? Yes, and here game-theoretical concepts prove extremely useful. We can easily formulate a truth definition for IF first-order logic without making any changes in any of the rules formulated thus far. All we have to do is heed more closely the game-theoretical character of GTS, and ask what any game theorist worth his or her salt will ask about semantical games, namely, are they games of perfect information?

Hintikka here proposes to reproduce independence in the substantial sense within his semantics by *informational independence* in the sense of the general theory of games. It may help to think of this as an extension of the general idea, and in a sense clearly it is, but here there is an extremely important and easily overlooked point. For this idea is in fact already built into the received first-order logic, which allows for representing dependencies between variables by dependencies between the corresponding quantifiers. To complete the definition of GTS for IF logic, the only additional step we need make is to define the information set of a move in a semantical game, which then yields an elegant and powerful semantics for IF first-order logic. Such a move is a direct application of a game rule to a quantifier or connective. The moves connected with the quantifiers and connectives within whose scope the given one occurs comprise the information set.

That game-theoretical concepts actually do a great deal of work in GTS is illustrated by the use of the notion of informational independence. This is extremely important in the definition of truth with regard to the existence of a winning strategy, the most important notion in game theory. Hintikka has on several occasions quipped that game theory should be called *strategy theory*. But the notion of a winning strategy is not the only concept from game-theory for which we can find use in Hintikka's theory. Consider the concept of *determinacy*. Under what condition is a two-person zero-sum game determinate? If and only if the one or the other player has a winning strategy. Determinacy, we can thus see by looking at the game-theoretical characterization, is thus equivalent the law of excluded middle. Now, the question then is whether in IF first-order logic the law of excluded middle is valid. One of the most important lessons of game theory is that we cannot simply take determinacy for granted. Besides the fact that there are many indeterminate games, the assumption of determinacy is often equivalent to making strong set-theoretical assumptions. And because semantical games for IF logic are not determinate, in IF logic the law of excluded middle fails. Thus in the game rules above negation "~"is not an ordinary contradictory one but a strong (dual) negation. Another way to put this point would be to say, on the other hand, that the law of excluded middle does hold in IF first-order logic, but only within that fragment of it that we have been calling ordinary, i.e., handicapped, first-order logic.

Lest anyone think that IF logic is merely the latest addition to the list of nonclassical logics, let us point out that the laws, i.e., the rules for semantical games, in IF first-order logic and in ordinary first-order logic are the same. The only aspect that could be regarded as non-classical is that in IF first-order logic informational independence is allowed. But really the notion of informational independence is neither classical nor nonclassical. Part of the task of any first-order logic is to allow the representation of all possible patterns of dependence and independence between variables. It would be better to say that IF logic is the truly classical first-order logic and that the received classical first-order logic is "nonclassical." Another way to put this point would be to say that what Hintikka has done in IF logic is to free logic from unnecessary restrictions. The mistaken but widely held notion that "ordinary" (dependence-handicapped) first-order logic is our rock-bottom basic logic makes all the standard sorts of debates about classical vs. nonclassical logic moot.

Let us now see what happens when we extend Hintikka's independence-friendly first-order logic by adding to it contradictory negation, what Hintikka himself calls "*extended* IF logic." Because the semantical game rules for the strong negation are "classical" there is no way of using semantical rules to characterize contradictory negation. What we can say in a suitable metalanguage, however, is that the contradictory negation ¬S of S is true if and only if S is not true. We can thus characterize "¬" by means of contradictory negation in the metalanguage; moreover, if "¬" did not only sentence-initially, we would need game rules to handle it. This aspect of Hintikka's IF first order logic thus gives us a viable new perspective on the logic of negation both in formal and natural languages. That in a language in which there is contradictory negation present, there is inevitably also another, strong, negation present, explicitly or tacitly, applies as well to natural languages in which the contradictory negation is prominent. Hintikka shows this negation to be a derivative of another negation which alone is rule-governed. Although Hintikka has not elaborated the conclusions of this position, from his point of view there is quite a simple explanation of various natural-language phenomena, such as the fact that contradictory negation is a barrier to anaphora. The explanation results from Hintikka's theory according to which anaphoric pronouns are assigned values in the course of a semantical game. But since the meaning of ¬S does not depend on any play of a game, this means that the verifier does not have a winning strategy in G(S). As a result, there is no place for anaphora in ¬S. Take, as an example, the contrast between the following sentence pairs:

(A) Some soldiers survived the battle.
They must have fought hard.

(B) Not all soldiers were killed in the battle.
They must have fought hard.

Anaphora is possible in (A) but not (B).

Game-theoretical semantics is applicable also in the semantics of natural languages, as Hintikka already proved (with Jack Kulas, *The Game of*

Language [1983] and *Anaphora and Definite Descriptions* [1985]). Hintikka shows this for a variety of linguistic phenomena, from natural-language quantifiers to other anaphora and conditionals. Hintikka's theory of anaphora have much in common with Chomsky's government and binding theory, except that there are many theoretical advantages to Hintikka's approach. One reason why many linguists have failed to adopt Hintikka's superior approach is that he Hintikka has unfortunately not developed a full-fledged syntax to go together with his game-theoretical semantics. Here would be a wonderful opportunity here for an up and coming linguist with truly revolutionary aspirations. (In this regard, Hintikka has shown, as in "No scope for scope," that linguists' notion of scope is so deeply confused that it cannot bear any explanatory burden. His punch line: "In linguistics, once a day with scope does not do it.")

Nevertheless, the implications of Hintikka's theories on this topic are clear. For instance, although the rules for English quantifiers are naturally closely related to the rules for quantifiers in first-order languages, structurally they are quite different. The "logical forms" of English quantifier sentences are different from those of first-order logic. This shows that Chomsky-inspired linguists attempting to use formulas of ordinary first-order logic as representations of logical form of English sentences are literally barking up the wrong (logical) tree.

One of the most revolutionary ideas which Hintikka derives from his development of game-theoretical semantics for natural languages is the demolition of what he calls the Frege-Russell ambiguity thesis, what Russell himself called "the first serious advance in real logic since the time of the Greeks." One of the most common notions shared by analytic philosophers is that verbs for *being* in natural languages are multiply ambiguous between the *is* of identity, the *is* of predication, the *is* of existence, and the *is* of subsumption. This ambiguity thesis, built into the very notation of first-order logic, where the different *is*es are represented in different ways, is a staple of introductory logic courses taught to students everywhere. This goes well beyond the recognition that verbs for being like *is* are used in different ways on different occasions. It is an attempt to explain that difference in usage as being due to the ambiguity of a single word and not, for example, to differences in context. But Hintikka shows that in a game-theoretical treatment of ordinary English the Frege-Russell ambiguity thesis is unnecessary. The fact that in some cases it is impossible to differentiate between the different Frege-Russell meanings already throws doubt on the Frege-Russell thesis. Ordinarily, when a word is ambiguous, there are sentences containing it that are ambiguous because of the lexical ambiguity. But nobody has yet come up with an English sentence which is ambiguous because of the alleged ambiguity of *is*. This makes the Frege-Russell highly dubious. However, Hintikka is not asserting that *is* is or is not ambiguous in

English. What he is doing is showing that the possibility of constructing a semantics where *is* is not ambiguous reveals the relativity of ambiguity claims to their preferred semantical framework. That the majority of analytic philosophers should find this disturbing is part and parcel of Hintikka's point that ordinary first-order logic is not the logic of our language. Their response has been to try and sweep the problem under the rug, by ignoring Hintikka's observations (going as far back as his 1979 paper, "'Is,' semantical games and semantical relativity"[84]) and not applying them in the logical analysis of ordinary language.

All philosophers should heed Hintikka's discovery, if for no other reason than that before the nineteenth century philosophers did not rely on the Frege-Russell ambiguity, which makes the use of ordinary first-order logic highly suspect as an interpretational tool in the history of philosophy, since the Frege-Russell thesis is built into its notation. Some historians of philosophy have dispensed with the Frege-Russell ambiguity thesis and on a few occasions pointed out that this or that historical figure did not presuppose it.[85] The problem is that these insightful historians run the risk of being taken to accuse past fphilosophers of a logical howler, which is why anti-Fregean historians have by and large remained in the closet. Hintikka's result liberates them.

There are still other fruitful ways that Hintikka has related his game-theoretical semantics for natural language to the history of philosophy, which we have not the space to develop here, such as, for instance, showing striking similarities to Aristotle's theory of categories.

6. THE WHOLE TRUTH ABOUT TRUTH: THE SET-THEORETICAL IMPLICATIONS OF HINTIKKA'S NEW LOGIC

The possibility of dispensing with all quantification over sets and other higher-order entities throws a shadow over the familiar first-order axiomatizations of set theory, such as the ZF (Zermelo-Fraenkel) set theory in which quantifiers range over sets. Such axiomatic set theories have generally been thought of as the foundation of all mathematics. Hintikka has shown that they are nevertheless defective, reason being that in them one can

[84] "'Is,' Semantical Games and Semantical Relativity" in the *Journal of Philosophical Logic*, vol. 8. See also *The Logic of Being: Historical Studies,* edited by Simo Knuuttila and Jaakko Hintikka, D. Reidel 1986.

[85] Michael Frede argues in his brilliant *Habilitationsschrift* (1967) that Plato did not distinguish the predicative and the existential senses of *estin* in his *Sophist*.

prove theorems that are false according to our normal combinatorial sense of set-theoretical truth. As a consequence, provability or unprovability of a hypothesis in first-order axiomatic set theory does not automatically tell anything about the truth or falsity of that hypothesis. This applies in particular to Gödel's and Paul Cohen's well known results concerning the continuum hypothesis.

Are there better ways of approaching truth in set theory? Hintikka is convinced that suitable model-theoretical methods provide such a way. In particular, he is taking up Gödel‘s suggestion that appropriate maximality assumptions might produce new insights into the central problems of set theory, including the nature of the continuum. Maximality cannot of course mean maximal cardinality of the domain, but rather maximal richness as far as the different kinds of individuals are concerned. Hintikka is currently exploring ways of implementing this idea of a maximally rich model.

That the importance of Hintikka's refinement of the concept of truth—a central result of his IF logic and game-theoretical characterization—is not restricted to logical theory should be obvious, since truth and its properties, such as definability, figure predominantly in philosophy in general and epistemology in particular. With regard to explicit first-order languages, in the restricted received sense of ordinary "dependence-handicapped" first-order logic, it was already clear in the thirties from Tarski's results that the concept of truth for such a language can be defined only in a richer metalanguage. This is clearly laid out in his impossibility theorem. Tarksi not only showed how to formulate such a definition using a method that subsequent applications have hardly improved on, he argued persuasively against the possibility of using a clear notion of truth in our "colloquial language," as he called it, namely, our ordinary working language.

With regard to truth-definitions formulated in a suitable metalanguage, some have argued that the notion of truth is so fundamental that to try to define it in terms of anything more basic is unreasonable if not impossible. Thus, perhaps most famously, Donald Davidson has gone so far as to speak of "the folly of defining truth." It would therefore seem that there is not much left to be said about truth, truth-predicates, truth-definitions, and so on. The only remaining open question, it would seem, is to what degree our intuitive notion of truth might be captured by Tarski's truth-definitions.

Hintikka's work puts all this in a new light. If you look, for instance, beyond the technicalities of his highly illuminating 1998, "Truth-Definitions, Skolem Functions and Axiomatic Set Theory,"[86] you will see several shocking results. In asking about the grounds of Tarski's impossibility

[86] "Truth-Definitions, Skolem Functions and Axiomatic Set Theory," *Bulletin of Symbolic Logic* vol. 4, 1998, pp. 303-337.

theorem, Hintikka, like many other logicians, uses as a test case a first-order arithmetical language L in which a truth-definition for L would have to use something very much like Gödel numbering. In other words, there must be a way for us to use numbers to code the formulas of the language. However, when we actually do set up such a Gödel numbering, it is difficult if not impossible to see why a truth-predicate should be inexpressible. Given a sentence S, we can then compute its Gödel number g(S) using completely elementary operations that should be expressible in L. We can likewise arithmetically express the reverse construction of S from g(S). So then why can't we define a truth predicate simply by saying that it applies to g(S) if and only if S? Instead of S, we could use any sentence that is logically equivalent with S. This of course amounts to trying to turn Tarski's T-schema into a truth predicate. Tarski's T-schema is of the form "σ is true if and only if S" where "σ" is a placeholder for a quote or a structural description of S. While Tarski proves on certain assumptions that it is impossible to do so, his argument does not show an intuitive reason as to why. In the case of natural language Tarski seems to place the blame for the inexpressibility of truth on irregularities and perhaps even inconsistencies. According to Hintikka's diagnosis of this problem, in using Gödel numbering we are speaking of numbers in two different roles, either as numbers pure and simple or else as codifications of formulas. This is not any stranger than speaking of actors on the one hand as characters in a play and on the other hand as citizens of the world outside the play. No contradictions ensue, and no confusion needs to ensue, provided that we observe an important restraint when we introduce quantifiers.

Quantifier that range over characters in a play and quantifiers that range over real-life individuals must be independent of each other. So too with quantifiers which range over numbers as numbers and quantifiers that range over numbers that play the role of formulas; they too must be independent of each other. Quantifier dependencies serve to express actual dependencies between variables. For conceptual reasons there cannot be any such dependencies between numbers in their two different interpretations, just as in the analogy with stage actors. Quantifiers ranging over actors in their civilian life cannot be dependent on quantifiers ranging over the characters in a play, for if they were, the variables describing ordinary everyday life would depend on variables ranging over the imaginary play universe. In the transition from g(S) to S that was supposed to yield a truth predicate, we have to use independent quantifiers.

Now, what Hintikka's results here show, in a nutshell, is this: it is impossible to reduce the pattern of dependent and independent quantifiers needed in the transition to those of the ordinary Frege-Russell quantificational logic. This answers the heretofore unanswered question of why Tarski's impossibility theorem holds in the first place. The answer is that Tarski's theorem holds because Tarski relies on the received Frege-Russell logic, which is intrinsically flawed in that not all possible configurations of quantifiers can be expressed in it. This result extraordinary result shows once and for all that Tarski's impossibility theorem is philosophically vacuous. It holds because the logic Tarski is using, namely, received "independence-challenged" first-order logic, is too weak to be the general logic of all quantifier patterns. Received independence-challenged first-order logic does not speak to the possibilities of truth-definitions in general. Tarski's result therefore shows nothing about the prospects of using a truth predicate in natural languages. This explains, again in a nutshell, why Hintikka keeps insisting that informational independence is found among natural-language quantifiers. What Hintikka is showing us is that that we *can* use a truth predicate in natural language. It should force all philosophers to take another look at all the various theories of truth in a new light.

Hintikka's line of thought shows how as soon as we can represent independence in our first-order arithmetical language, we can formulate a truth predicate for it in the same language by carrying out a version of the argument that first led to the puzzle about the prima facie definability of truth. Hintikka's version says that the truth predicate applies to the Gödel number g(S) of S if and only if a full set of Skolem functions exists for S. We explain these functions below. What matters first is that the resulting existence claim is of the form

Σ^1_1

having a string of second-order existential quantifiers followed by a first-order formula, therefore expressible in IF first-order logic provided that the syntax of the first-order language in question can be expressed in the same language. This, then, is Hintikka's answer to the question of the definability of truth in first-order languages.

Although this has grave implications for all the so-called theories of truth, thus far Hintikka has not yet explored these implications in his writings or made them explicit. Nevertheless, what we can say, at this point, is fairly clear, namely, that if one cannot try to define truth, it is not because the task is difficult or impossible. Hintikka's reliance on Tarski's T-schema brings his views close to minimalist approaches to truth in which the centerpiece is Tarski's T-schema.

Consider now another fascinating feature of Hintikka's truth predicates, namely, their precise form. Such a predicate applies to the Gödel number g(S) of a sentence S just in case a full set of Skolem functions exists for S. Not only is this expressible on the first-order level by means of IF logic, but these functions have an intuitive meaning, in so far as these functions tell us how to find what Hintikka calls the *witness individuals* that vouchsafe the truth of a sentence S_o. As an example, if S_o is

(6.1) $(\forall x)(\exists y)\ S_1[x,y]$

where $S_1[x,y]$ does not contain quantifiers, the Skolem functions g are only those that satisfy

(6.2) $(\forall x)\ S_1[x,g(x)]$

Intuitively, what (6.1) says is this: given any individual x you can find an individual y which satisfies $S_1[x,y]$ as soon as there exists a function g satisfying (6.2). Similarly, what more complex quantificational sentences, including independence-friendly ones, show is that we can think of the existence of a full set of Skolem functions for a first-order sentence can as an implementation of our pre-theoretical conception of truth. Therefore, one leading line of objections leveled at Tarski-type truth definitions fails to apply to Hintikka's view.

Hintikka's truth predicate expresses the existence of the relevant Skolem functions. It can therefore be considered as an explication of our common-sense idea of truth. One line of objections to Tarski has insisted that while his truth definitions established a certain relation between closed sentences and certain facts, this relation need not be one of truth. But Hintikka's truth predicate, on the other hand, shows clearly by its very form that it is what we normally mean by truth. Moreover, Hintikka emphasizes the fact that his truth predicate is first-order. This means that Hintikka's truth predicate is independent of all questions of existence or nonexistence of higher-order entities, such as sets, meaning entities, intensions, and so on.

Hintikka's truth predicate, which he calls *combinatorial*, deals solely with the existence or nonexistence of structures of individuals. This has confounded some of his interpreters because *combinatorial* for the simple reason that it is a confusing term. From a practical standpoint it means, basically, "first-order," in the sense that all quantifiers range over individuals. But this should not obscure what Hintikka is thereby highlighting.

Consider, for instance, his "Truth-Definitions, Skolem Functions and Axiomatic Set Theory," where Hintikka pushes this line of thought to a remarkable conclusion. If you apply his novel treatment of truth to a first-order axiomatic set theory, what do you get? Although the logic in which such set theories are formulated is the received "dependence handicapped" first-order logic, the existence of functions, including Skolem functions, is nevertheless therein expressible. This means that it should then be possible to define truth for axiomatic set theory in the same theory in terms of the existence of Skolem functions. But such a truth definition is impossible by Tarski's theorem because the usual axiomatizations of set theory use ordinary first-order logic. Something must therefore be revised.

To see this point, and what needs revision, consider all pairs of set-theoretical sentences $<S, S^*>$ where S^* asserts the existence of all the Skolem functions of S. S^* logically implies S. S^* is expressible in the language of axiomatic set theory if S is expressible. Therefore, if each S^* were true as soon as S is in some model for axiomatic set theory, there would exist a truth definition for first-order axiomatic set theory. But Tarski's theorem shows that this is impossible.

This shows that we cannot add all the conditionals $(S \supset S^*)$ as axioms of set theory unless we make the resulting theory inconsistent. By compactness, some finite subset of the set of all these conditionals must be incompatible with the axioms of set theory. If this finite set is

(6.3) $\{(S_i \supset S_i^*)\} \quad i \subset I$

the disjunction of their negations is then

(6.4) $\bigvee_i \; (S_i \mathbin{\&} \sim S_i^*)$.

It is easy to see that (6.4) is provable in the first-order axiomatic set theory in question. But if this existence of functions is understood in the usual combinational sense, then (6.4) is false! Therefore, at least one false sentence is provable in any first-order axiomatic set theory.

Does this mean that axiomatic set theories are inconsistent? No! The reason is that (6.4) is *true in the set-theoretical sense*. This is the sense of truth when models of axiomatic set theory are viewed in the same way as any model of first-order theory, with sets considered as individuals of a certain kind and with the membership relation being considered like any ordinary two-place relation.

What we've just shown, then, is that what this line of Hintikka's thought reveals, expressed here in terms slightly different from Hintikka's, is not a contradiction in any axiomatic set theory. Rather, it reveals that set-theoretical truth is not a good guide to truth in the sense of the combinatorial truth (existence of Skolem functions) that codifies our intuitive idea of truth. To which a defender of the received axiomatizations of set theory might reply: So much the worse for our intuitive idea of truth! In set theory we've learned to distrust our intuitions, as demonstrated by sundry paradoxes. So then why not here? Because rejecting what Hintikka calls "the combinatorial notion of truth" would make nonsense of much of the foundations of set theory, that's why. For in fact many central notions and problems in set theory presuppose a combinatorial notion of truth.

Consider, for instance, the special continuum hypothesis CH. What it means is not that the set-theoretical statement that there are no cardinalities

α such that $à_0 < \alpha < 2^{à}{}_0$ (expressed as a first-order statement) is true in some first-order model of set theory, or true in every such model (true in the usual first-order sense when sets are thought of as individuals and membership as just another two-place relation among individuals). Rather, what CH means is that there is no subset C of reals purely extensionally considered such that no "function-in-extension," as Russell called them, can map C one-to-one either on the set of reals or on the set of natural numbers.

What this means, in a nutshell, is that CH makes sense only by reference to combinatorial truth. But then first-order axiomatic set theories do not offer any help in trying to decide whether CH is true. Since combinatorially false sentences are provable in them, the provability or unprovability of the continuum hypothesis in an axiomatic set theory does not by itself show anything about its truth. Either CH or its negation could be provable in axiomatic set theory and yet false. Gödel's and Paul Cohen's independence results do not by themselves show anything about the truth or falsity of the continuum hypothesis. Gödel's proved that CH is deductively compatible with the usual axioms of set theory, and Paul Cohen proved that its negation is deductively compatible with those same axioms. But since in axiomatic set theory some combinationally false sentences are provable, CH might be deductively provable from the axioms and yet false, and likewise for its negation.

These are some of the most shocking consequences of Hintikka's ideas, since these independence results are generally considered the culmination of logical research in the twentieth century. Indeed, the revolutionary character of Hintikka's results here are best seen in light of the historical development of axiomatic set theory, which after all was developed as a response to the paradoxes that threatened the foundations of mathematics in the form of contradictions. The strategy evoked by axiomatic set theory was itself predicated on the idea that we can stake out an unproblematic core area of set-theoretical truths which was

1) strong enough to satisfy the needs of most actual mathematical proofs,
2) was by itself insufficient for the derivation of all and sundry mathematical truths, and
3) could be extended gradually by the introduction of stronger new set-theoretical assumptions.

The idea was that we might even get lucky enough and justify the new axioms by means of consistency proofs. This overall strategy, endorsed particularly emphatically by Kurt Gödel, became the standard method. But now Hintikka's results force us to reject this entire strategy. For Hintikka shows that the allegedly unproblematic core theory already implies false theorems, i.e., theorems that are false in the intuitive combinatorial sense which set theory is calculated to capture, even in axiomatic systems of set theory that are formally consistent. Therefore, such a flaw is not ruled out by formal consistency proofs and cannot be eliminated by adding new axioms.

Hintikka thus questions the very *raison d'être* of first-order axiomatic set theory in any and all of its current variations. He shows not only that axiomatizations of set theory using ordinary first-order logic are not reliable guides to set-theoretical truth but that, even more generally, set theory is a fallible guide in the quest of mathematical truth.

Might IF logic help here? Perhaps, though no one, not even Hintikka, has yet attempted this. However, let us point out that should IF logic come to the rescue here, it means that we must give up one of the basic ideas of Cartesian set theory, namely, that a set is completely determined by the collection of its members. In our as yet to be constructed IF logic set theory there are objects that are neither members nor nonmembers of a given set.

7. THE REAL LOGIC OF QUANTUM THEORY?

The notions of dependence and independence of quantifiers as expressed by the dependence and independence of the corresponding variables have many other important ramifications and applications. One special case is an irreducibly mutual dependence of two variables. It turns out that this idea can be implemented only by allowing the values of the relevant variables to include probability distributions ("fuzzy objects"). This necessitates a generalization of the mathematical function, not by changing the notion of dependence it embodies but by allowing the argument values and function values to be sometimes probability distributions.

This line of thought has an unexpected application. Hintikka has in the past several years explored the conceptual foundations of quantum theory starting from the assumption that conjugate variables in the sense of classical mechanics are in quantum theory irreducibly mutually dependent. From the logic of mutual dependence it is then seen that whenever one of a pair of conjugate variables receives a definite value, the "value" of the other one is a probability distribution. Hintikka and some of his associates are in fact presently examining the precise connection between this dependence assumption and the received mathematical formulations of quantum theory. The connection is made somewhat confusing by the fact that the received mathematical techniques used in quantum theory do not rely on generalized functions but on matrices or operators. It is nevertheless already clear on a qualitative level that in this way a sharp new light can be thrown on the vexing conceptual problems in this area, such as the measurement problems (including the so-called collapse of the wave function), the locality problems (including the Einstein-Podolsky-Rosen argument), and the notion of the state of a system (including the superposition of states), a more complete list of which is simply enumerated for the purposes of elucidation below.

Indeed, one of the most intriguing problem complexes at the interface of mathematics, physics and logic concerns the algebraic, probabilistic and logical structures needed to understand quantum theory. John von Neumann as the topic of his address chose this problem complex to the International Congress of Mathematicians in 1954 on "Unsolved Problems in Mathematics." Even now this problem complex nevertheless remains largely to be cleared up, including some of the suggestions von Neumann made in his talk.

Some new light on this problem area is promised by Hintikka's ongoing revolution in the foundations of logic. If there is an obvious requirement on a satisfactory logical language, it is that all possible configurations of relations of dependence and independence between different variables must be expressible in it. Yet this requirement, as we explained above, is not satisfied by the received Frege-Russell logic. In such a logic, dependencies between variables are represented by dependencies between quantifiers. But as we have seen not all possible dependence structures among quantifiers are expressible in our usual logic.

Now that Hintikka has shown how the usual first-order logic can be extended into IF first-order logic that satisfies this expressibility requirement and which consequently is much stronger than the conventional Frege-Russell logic, mutually (symmetrically) dependent variables can more or less easily be handled. Although much more work remains to be done, it is quite apparent to Hintikka and his associates that this case of mutual dependence constitutes the true logic of quantum theory.

Indeed, such symmetrical dependence is existentially tantamount to noncommutativity in the sense used in quantum mechanics. Unsuprisingly the old von Neumann-Birkhoff quantum logic can be interpreted in a slight extension of IF logic. In the IF logic itself, unlike von Neumann-Birkhoff logic, which has the structure of an orthomodular lattice, distributive laws hold, which makes it possible to use the frequency interpretation of probability in it and hence to give it a firmer physical applicability. Furthermore, in the presence of mutually dependent variables, the possible value of generalized vectors (sequences of variables) is revealed by the eigen-vectors of certain generalized operators (totalities of Skolem functions for a proposition), just as in quantum theory.

When John von Neumann tried to apply his operation calculus to a thought experiment, with the help of Garett Birkhoff he ended up replacing the Boolean lattice logic of classical mechanics with an orthocomplemented modular lattice, where distributivity is replaced by the weak modular identity. It can easily be seen that the non-distributive modular lattice structure of subspaces represents relations of dependence and independence between measurements of different observables, in an eminently *independent friendly* way. It shows that these measuremenetts of different observables—and here if it helps to bridge, or should I say functor the relata of the relation, one may think once again of *witness individuals*—are apt to interfere with each other. And yet, at the same time, to disturb the concept of the distributive law of classical prepositional calculus seems to breach the correspondence principle. The highly provocative and influential Birkhoff-von-Neumann orthocomplemented modular lattice structure of subspaces was of course criticized—it should come as no surprise—by Karl Popper, who charged it as logically inconsistent. The Journal *Nature* where the debate ensued, received several letters sent as replies to Popper claiming that he too was inconsistent, but they were never published.

Another obvious resemblance, once one gets to know a few actual members of the family, is the probability logic of Suppes, in which non-classical logic arises from the probability assigned to every event and conjunction of event. And, we may be permitted for a moment to wax even more category theoretical, there is now more recently also the inexplicable nature of the holographic type nonlocal effective action in two-brane Randall-Sundrum models, involving phase transitions between the local and nonlocal phases of the theory, along with the impossibility of consistent descriptions of massive Kaluza-Klein modes; the nonminimal coupling of Einstein gravity to the Brans-Dicke type scalar describing the local distance between branes requiring "corrections" in terms of the squared Weyl tensor with locally *independent*, yet simultaneously scalar *dependent* coefficient. A plethora of field theories based on non-commutative spacetimes exhibit distinctive nonlocal effects that are acausal and inconsistent with conventional Hamiltonian evolution. But of course as everybody now knows even long before that, electrons were seen as behaving as self-organizing systems whose geometrical shape and linear dimensions are determined by the long-range Coulomb forces in their surroundings, in which the electron becomes an open system dependent upon but inseparably bound with its environment; that is the sense in which, it was realized quite early on in physics, that the whole universe must take part in the formation of the electron as a physical system.

In IF logic we thus have not an answer or a view but a new tool for exploring even the rich underlying and as yet unexplored logical signficance in the Feynman integral path parallelism, superstrings following the Einstein-Dirac equations, Mackey's axiomatic groups in Hilbert space, Mittelstedt's hidden variables, non Heyting logic, Bohm's notion of implicate order, the Bell inequalities, the Schrödinger quantization as eigen values, the Bohr complementarities, the Born state vector, the statistical interpretation and view of probability as an intermediate physical reality, the von Neumann operator calculus in Hilbert space and projection postulate, Heisenberg's uncertainty principle and reduction of the wave packet, Pauli's canonical conjugate variables, the de Broglie wave particle duality, probability waves, superposition, and the Planck correspondence principle. We may thus one day soon if not before come to see a new logical light at the proverbial end of the quantum tunnel, and a friendly one at that.

ANNOTATED BIBLIOGRAPHY OF JAAKKO HINTIKKA

1953

Books

(a) *Distributive Normal Forms in the Calculus of Predicates, Acta Philosophica Fennica* **6**, (1953), 71pp.

By distributing quantifiers deeper and deeper into the formulas of a finite first-order language, with or without identity, at the same time as truth-functions are transformed into a propositional normal form, we obtain a simple normal form for such formulas. This refutes the conjecture of Hilbert and Bernays that "a simply characterizable normal form is not obtainable in this way in general." The basic properties of these distributive normal forms are established, including a set of sufficient conditions of inconsistency.

Papers

(a) "A New Approach to Sentential Logic", *Societas Scientiarum Fennica*, *Commentationes Physico-Mathematicae* **17**, no. 3, (1953), 13pp.

1954

Papers

(a) "An Application of Logic to Algebra", *Mathematica Scandinavia* **2**, (1954), 243-246.

1955

Books

(a) *Two Papers on Symbolic Logic*, *Acta Philosophica Fennica* **8**, (1955), 115pp. (Includes "Form and Content in Quantification Theory", 11-55, and "Reductions in the Theory of Types", 57-115.)

First paper: For a finite first-order language, the concept of a model set (m.s.) of formulas is defined by means of certain simple closure and compatibility conditions. Intuitively, a m.s. is a partial description of a possible world. A set of formulas is accordingly consistent if it can be embedded in some m.s.. A simple proof procedure is obtained by considering rules for an attempt to embed the negation of a sentence *S* in a m.s.. If such an attempt fails in all directions, a proof is obtained for *S*. Second paper: For a formula *S* in the simple theory of types, a formula *r(S)* is constructed such that:

(i) *r(S)* is satisfiable if *S* is satisfiable;

(ii)*r(S)* is of the form (∃*X*) *F*(*X*) where *"X"* is a second-order formula which does not contain any second-order quantifiers.

Papers

"Notes on Quantification Theory", *Societas Scientiarum Fennicae, Commentationes Physico-Mathematicae* **17**, no. 12, (1955), 13pp.

1956

Papers

(a) "Identity, Variables, and Impredicative Definitions", *The Journal of Symbolic Logic* **21**, (1956), 225-245.

(b) "Loogisen kielentutkimuksen näköaloja", *Ajatus* **19**, (1956), 81-96. ("Perspectives on the Logical Study of Language".)

1957

Papers

(a) "Arvokäsitteistä sosiaalitieteiden metodiopissa", *Ajatus* **20**, (1957), 27-47. ("On Value-Concepts in the Methodology of the Social Sciences".)

(b) "Modality as Referential Multiplicity", *Ajatus* **20**, (1957), 49-64.

Problems of intensional context can be handled by assuming that they involve a multiplicity of different situations (worlds, models) in which the references of our singular terms can be different. (Note: This paper seems to be the first philosophical statement of the basic ideas of possible-worlds semantics in the literature. The problems of cross-identification are nevertheless overlooked in it. They make their appearance in *Knowledge and Belief: An Introduction to the Logic of the Two Notions,* (1962(a)).)

(c) "Necessity, Universality, and Time in Aristotle", *Ajatus* **20**, (1957), 65-90.

(d) "Quantifiers in Deontic Logic", *Societas Scientaiarum Fennicae, Commentationes Humanarum Litterarum* **23**, no. 4, (1957), 21 pp.

(e) "Vicious Circle Principle and the Paradoxes", *The Journal of Symbolic Logic* **22**, (1957), 245-249.

1958

Papers

(a) "On Wittgenstein's 'Solipsism'", *Mind* **67**, (1958), 88-91.

(b) "Remarks on a Paradox", *Archiv für Rechts- und Sozialphilosophie* **44**, (1958), 514-516.

(c) "Towards a Theory of Definite Descriptions", *Analysis* **19**, (1958), 79-85.

The methods used in the new theory of descriptions illustrate the virtues of our new quantification theory which dispenses with existential presuppositions. This system is superior to Russell's.

1959

Papers

(a) "Aristotle and the Ambiguity of Ambiguity", *Inquiry* **2**, (1959), 137-151.

(b) "An Aristotelian Dilemma", *Ajatus* **22**, (1959), 87-92. (An expanded version under the title "On Aristotle's Modal Syllogistic" appears in Jaakko Hintikka, *Time and Necessity: Studies in Aristotle's Theory of Modality*, Clarendon Press, Oxford, 1973, 135-146.)

(c) "Existential Presuppositions and Existential Commitments", *The Journal of Philosophy* **56**, (1959), 125-137.

The usual quantification theory can be modified in such a way as to allow individual constants to be empty. What is needed is a suitable change in the quantifier rules. (Note: Together with the paper by Leblanc and Hailperin in the *Philosophical Review* in 1959, this paper seems to have been the first explicit treatment of a logic without existence assumptions in the literature.)

(d) "Filosofian ajankohtaisista ja ajattomista tehtävistä", *Suomalainen Suomi* **27**, (1959), 538-542. ("On the Timely and Timeless Tasks of Philosophy".)

(e) "Kantin oppi matematiikasta: tutkimuksia sen peruskäsitteistä, rakenteesta ja esikuvista", *Ajatus* **22**, (1959), 5-85. ("Kant's Theory of Mathematics: Studies in its Basic Concepts, Structure, and Precedents".)

1960

Papers

(a) "Aristotle's Different Possibilities", *Inquiry* **3**, (1960), 17-28.

1961

Papers

(a) "Cogito, ergo sum, 1-11", *Nya Argus* **54**, (1961), 143-146 and 159-162. (An early Swedish version of "Cogito ergo sum: Inference or Performance?", *Philosophical Review* **71,** (1962), 3-32.)

(b) "Filosofia ja maailmankatsomukset", *Uusi Suomi*, (22 October 1961). ("Philosophy and *Weltanschauungen*".)

(c) "Käsitteilläkin on kohtalonsa", *Suomalainen Suomi* **29**, (1961), 459-464. ("Concepts Have Their Fates, Too".)

(d) "Modality and Quantification", *Theoria* **27**, (1961), 119-128. (An expanded version appears in Jaakko Hintikka, *Models for Modalities,* D. Reidel Publishing Co., Dordrecht, 1969, 57-70.)

Satisfiability may be defined for sets of formulas containing modal operators and quantifiers by means of the notion of model set (in short, m.s.; for a definition see JSL XX362): λ is satisfiable if and only if there is a model system Ω and a m.s. $\mu\in\Omega$ such that $\lambda\subseteq\mu$. Model system is defined as a set Ω of m.s.'s together with a reflexive dyadic relation (called the relation of alternativeness) on Ω which satisfies the following conditions: (1) if $\underline{Mf} \in\upsilon\in\Omega$, then there is in Ω at least one alternative υ^1 to υ such that $\underline{f}\in\upsilon^1$; (2) If $\underline{Nf} \in\upsilon\in\Omega$ and if $\upsilon^1\in \Omega$ is an alternative to υ, then $\underline{f}\in\upsilon^1$.

The resulting system corresponds to von Wright's M in that a quantifier-free f is provable in M if and only if {~f} is not satisfiable. If the relation of alternativeness is stipulated to be transitive, the resulting system will similarly correspond to Lewis's S4; if transitive and symmetric, to S5.

Most of the usual difficulties can be avoided by specifying whether and when free individual symbols (say a, b,… are transferable from a m.s. to another in a model system. If no transfer is permitted, (2) has to be replaced by a weaker condition obtained by adding: "provided that each free individual symbol a of f occurs in some of the other formulas of υ^1." (Condition (3).) However, we may (if we choose) permit transfer from a m.s. to its alternatives (in the old deductive systems the corresponding assumption is sometimes made unwittingly). Then (2) is acceptable, but one of the defining conditions of m.s.'s has to be strengthened to read: (4) If $(x)f \in \mu$, then $f(a/x) \in \mu$ for every a occurring in any $\lambda \in \Omega$ to which μ bears the ancestral of the alternativeness relation. ($f(a/x)$ results from f by replacing x by a.) However, if empty domains are disqualified (e.g. by the condition (5): if $(x)f \in \mu$ (μ being a m.s.), then $f(a/x) \in \mu$ for at least one a), then it suffices to strengthen (3) to (2).

If we also want to permit the converse transfer (thus in effect permitting arbitrary transfer), it can be done by adopting an additional condition (6) obrained from (4) by reversing the roles of λ and μ.

If identities are admitted to the formalism, the conditions of their transfer must likewise be carefully defined.

1962

Books

(a) *Knowledge and Belief: An Introduction to the Logic of the Two Notions,* Cornell University Press, Ithaca, 1962, pp. x+179.

An explicit logic is developed for knowledge and belief (knowing that and believing that), formulated by means of the author's model set method. The problem of "logical omniscience" (one apparently necessarily knows all the logical consequences of what one knows) is dealt with by reinterpreting the metalogical notion of provability. An analysis is presented of knowing wh-constructions in terms of "knowing that" plus quantifiers, and the interplay of knowledge and quantifiers is studied. The approach is applied to selected conceptual problems, especially to Moore's paradox of saying and disbelieving and to the notion of "knowing that one knows."

Papers

(a) "Cogito, ergo sum: Inference or Performance?", *Philosophical Review* **71,** (1962), 3-32.

Descartes' *cogito ergo sum* is not an inference from *cogito* to *sum.* Its special character is due to the self-defeating character of an attempt to think that I don't exist, analogous to the self-defeating character of the assertion "I don't exist." Hence "*cogito*" does not express a premise, but refers to the act through which the self-defeating or self-verifying character of certain thought-acts is manifested. This throws light on several aspects of the *cogito* in Descartes, e.g., its curiously momentary character.

(b) "Huomioita kreikkalaisten ajankäsityksestä", *Ajatus* **24**, (1962), 39-65. ("Observations on the Concept of Time in Ancient Greek Philosophy".)

(c) "Johdonmukaisen järkevyyden ihanteet: 90-vuotias Bertrand Russell", *Uusi Suomi*, (18 May, 1962). ("The Ideals of Consistent Reasonableness: Bertrand Russell at Ninety".)

(d) "Kaksi Spengleriä?" *Suomalainen Suomi* **30**, (1962), 86-92. ("Spengler Against Himself?")

(e) "Kepler ja Galilei", *Suomalainen Suomi* **30**, (1962), 278-281. ("Kepler and Galileo".)

(f) "Kieliopin uudet tiet", *Suomalainen Suomi* **30**, (1962), 106-107. ("New Paths in the Study of Grammar".)

(g) "Miksi hyve oli kreikkalaisten mielestä tietoa?", *Suomalainen Suomi* **30**, (1962), 341-349. ("Why Was Virtue Knowledge for the Ancient Greeks?")

(h) "On the Interpretation of 'De Interpretatione xii-xiii'", *Acta Philosophica Fennica* **14**, (1962), 5-22. (An expanded version appears in Jaakko Hintikka, *Time and Necessity: Studies in Aristotle's Theory of Modality*, Clarendon Press, Oxford, 1973, 41-61.)

1963

Papers

(a) "'Cogito ergo sum' as an Inference and a Performance: Reply to Comments by J.R. Weinberg and J.D. Carney", *The Philosophical Review* **72**, (1963), 487-496.

(b) "Filosofian tehtävästä", *Suomalainen Suomi* **30**, (1963), 379-382. ("On the Task of Philosophy".)

(c) "The Modes of Modality", in *Proceedings of a Colloquium on Modal and Many-Valued Logics*, *Helsinki, 23-26 August, 1962*, *Acta Philosophica Fennica* **16**, (1963), 65-82.

1964

Papers

(a) "Aristotle and the 'Master Argument' of Diodorous", *American Philosophical Quarterly* **1**, (1964), 101-114. (Appears with new material in Jaakko Hintikka, *Time and Necessity: Studies in Aristotle's Theory of Modality*, Clarendon Press, Oxford, 1973, 179-213.)

(b) "Definite Descriptions and Self-Identity", *Philosophical Studies* **15**, (1964), 5-7.

(c) "Distributive Normal Forms and Deductive Interpolation", *Zeitschrift fur mathematische Logik und Grundlagen der Mathematik* **10**, (1964), 185-191.

(d) "Galilein kohtalo 400 Vuotta sitten", *Uusi Suomi* (15 February 1964). ("The Fate of Galileo 400 Years Ago".)

(e) "The Once and Future Sea Fight: Aristotle's Discussion of Future Contingents in 'De Interpretatione'", *The Philosophical Review* **73**, (1964), 461-492.

Aristotle's formulation of his problem and his conclusion in *De Int.* 9 show that he was not concerned with the applicability of *tertium non datur* to future contingents, but with their necessity. An analysis of Aristotle's argument shows that his solution depended on a distinction between the necessity of *p* "when it is", i.e. the necessity of (i) "*p* at time *t*", and its necessity *haplos*, i.e., the necessity of (ii) "*p* (now)." His solution was in effect to declare the latter the only appropriate sense of necessity. This shows that Aristotle's discussion was based on the general Greek conceptual assumptions studied in 1967(j) including the hegemony of temporally indefinite propositions like (ii) and the identification of necessary if true.

(f) "Päämäärä, sattuma ja välttämättömyys: eräiden kreikkalaisten ajatustapojen tarkastelua", *Ajatus* **26**, (1964), 61-81. ("Purpose, Chance and Necessity: Observations on Certain Greek Ways of Thinking".)

(g) "Tieto on valtaa: Eräitä aatehistoriallisia näköaloja", *Valvoja,* (1964), 185-196. ("Knowledge is Power: Reflections on the History of an Idea".)

1965

Papers

(a) "Analyyttisyyden käsitteen eri merkityksistä", *Suomalainen Tiedeakatemia,* Esitelmät ja päytäkirjat 1964, Helsinki, 122-137. ("On the Different Senses of the Concept of Analyticity".)

(b) "Are Logical Truths Analytic?", *The Philosophical Review* **74**, (1965), 178-203.

Analytic truths are the ones that can be established by analytic inferences, and an analytic inference is one in which one does not go beyond what is given in the premises. This "what is given" can be taken to be information, but in an interesting sense of analyticity it is the configuration of individuals (entities) considered in the premises (or, in a variant notion, in the premises *or* in the conclusion). Then an inference is analytic if the number of individuals considered together in it does not increase. This can be taken to mean that the maximal length of nested sequence of quantifiers (including tacit quantifiers binding free variables does not increase. In this sense, valid logical inferences can be synthetic, and (on certain natural assumptions) some of them must be. This sense of analyticity is closely related to Kant's and vindicates his idea that mathematical inferences (which for us would be mostly logical ones) are typically synthetic. (Note: This analytic-synthetic distinction is the same as the trivial-nontrivial distinction in 1973(a).)

(c) "A Closure and Complement Result for Nested Topologies", *Fundamenta Mathematicae* **57**, (1965), 97-106.

(d) "Distributive Normal Forms in First-Order Logic", in *Formal Systems and Recursive Functions. Proceedings of the Eighth Logic Colloquium*, *Oxford, July 1963*, J.N. Crossley and M.A.E. Dummett, editors, North-Holland Publishing Co., Amsterdam, 1965, 47-90.

(e) "Kant's 'New Method of Thought' and his Theories of Mathematics", *Ajatus* **27**, (1965), 37-47.

(f) "On a Combined System of Inductive Logic", in *Studia Logico-mathematica et Philosophica in Honorem Rolf Nevanlinna,* Helsinki, 1965, 21-30.

(g) "Tieto, taito ja päämäärä: Kaksi tutkielmaa vanhojen kreikkalaisten tiedonkäsitteestä", *Ajatus* **27**, (1965), 49-67. ("Knowledge, Skill, and Purpose: Two Studies on the Ancient Greek Concept of Knowledge".)

(h) "Towards a Theory of Inductive Generalization", in *Proceedings of the 1964 International Congress for Logic, Methodology and Philosophy of Science,* Yehoshua Bar-Hillel, editor, North-Holland Publishing Co., Amsterdam, 1965, 274-288.

Carnap's inductive logic is incapable of handling inductive generalization; in particular, in it all non-trivial generalizations have a zero probability in an infinite universe. This shortcoming can be eliminated by defining prior probabilities (in a monadic first-order language) by first dividing probabilities evenly between different constituents and then dividing the probability-mass of each constituent evenly between all the structure-descriptions satisfying it. Then the posterior probability (degree of confirmation) of the simplest constituent compatible with evidence grows with this evidence and eventually converges to one. The same idea can in principle be extended to any finite first-order language. Hence inductive generalization can be handled by means of a Carnap-type inductive logic.

1966

Books

(a) (edited with Patrick Suppes) *Aspects of Inductive Logic*, North-Holland Publishing Co., Amsterdam, 1966, viii+320 pp.

Papers

(a) "An Analysis of Analyticity", in *Deskription, Analytizität und Existenz: 3-4 Forschungsgepräch des Internationalen Forschungszentrums für Grundfragen der Wissenschaften Salzburg,* Paul Weingartner, editor, Pustet, Salzburg und München, 1966, 193-214.

(b) "Are Logical Truths Tautologies?", in *Deskription, Analytizität und Existenz: 3-4 Forschungsgepräch des Internationalen Forschungszentrums für*

Grundfragen der Wissenschaften Salzburg, Paul Weingartner, editor, Pustet, Salzburg und Müchen, 1966, 215-233.

(c) "Aristotelian Infinity", *The Philosophical Review* **75**, (1966), 197-219.

The following main points are argued and discussed: (1) In his theory of infinity, Aristotle did *not* give up the principle that every genuine possibility is sometimes realized; for (2) his infinity exists in a special sense of existence, and in this sense it is actualizable. (3) Aristotle denied even the potential existence of arbitrarily large spatial magnitudes; (4) this led into difficulties in his philosophy of mathematics. (5) Aristotle did *not* give up the principle that conceivability implies realizability; he only denied its applicability to the special case of absolute spatial magnitudes. (6) The distinction between conceivability and realizability is different from Aristotle's distinction between absolute and relative possibility.

(d) "Individen och statens ändamål", *Ajatus* **28**, (1966), 23-37. (In Swedish: "The Individual and the Aims of the State".)

(e) "Kant Vindicated", in *Deskription, Analytizität und Existenz: 3-4 Forschungsgepräch des Internationalen Forschungszentrums für Grundfragen der Wissenschaften Salzburg,* Paul Weingartner, editor, Pustet, Salzburg und München, 1966, 234-253.

(f) "Kant and the Tradition of Analysis", in *Deskription, Analytizität und Existenz: 3-4 Forschungsgepräch des Internationalen Forschungszentrums für Grundfragen der Wissenschaften Salzburg,* Paul Weingartner, editor, Pustet, Salzburg und München, 1966, 254-272.

(g) "Knowing Oneself and Other Problems in Epistemic Logic", *Theoria* **32**, (1966), 1-13.

(h) "Kommunikaatiovälineet ja yleinen kulttuurikehitys", *Parnasso* **16**, (1966), 21-27. ("Methods of Communication and General Cultural Development".)

(i)"Parmenideen peruslause ja kreikkalaisten tiedonkäsitys", *Valvoja* **86**, (1966), 138-146. ("The Axiom of Paramenides and the Ancient Greek Concept of Knowledge".)

(j)"Semanttisen informaation teoriasta", *Arkhimedes* **18**, (1966), 12-22. ("On the Theory of Semantic Information".)

(k) "Studies in the Logic of Existence and Necessity: Existence", *The Monist* **50,** (1966), 55-76. (A revised version appears as "Existential Presuppositions and their Elimination" in Jaakko Hintikka, *Models for Modalities*, D. Reidel Publishing Co., Dordrecht, 1969, 23-44.)

Suppose we want to change our quantification theory so as to admit empty singular terms. It is argued that the *only* way of doing so in accordance with our normal logical assumptions is to identify *B exists* with *(Ex)(B=x)*. This suggests the following consequences: (a) existence cannot be an *unanalyzable* predicate; (b) existence can be a *complex* predicate; (c) the prime vehicle of existential assertions is the existential quantifier; (d) to be *is* to be identical with one of the values of a bound variable. For this is what *(Ex)(B=x)* says of B; (e) we may make existential commitments in many ways, but all of them could be made equally well by means of the existential quantifier.

(l)"A Two-Dimensional Continuum of Inductive Methods", in *Aspects of Inductive Logic,* Jaakko Hintikka and Patrick Suppes, editors, North-Holland Publishing Co., Amsterdam, 1966, 113-132.

A two-dimensional continuum of inductive methods in Carnap's sense is defined for monadic inductive logic. One of the operative parameters is Carnap's lambda. It governs the prior improbability of inferences between particular cases, and is hence an index of caution for singular inductive inference. A similar index of caution alpha is defined for inductive generalization. When alpha goes to infinity, we obtain Carnap's lambda-continuum. This is the only case in which inductive generalization is impossible in an infinite universe. When alpha = 0, the result is the system proposed in 1965(h). When in addition lambda equals the width of a constituent, we obtain Hintikka's combined system expounded in 1965(f). Thus inductive generalization can be dealt with essentially the same way as singular inductive inference in Carnap's lambda-continuum.

(m) "Yhteiskunta-ja käyttäytymistieteet", *Luotain* **5**, (1966), 16-23. ("Social and Behavioral Sciences".)

(n) "Yksilö ja valtion päämäät ruotsiksi otsikolla: Individen och statens ändamål", *Ajatus* **28**, (1966), 23-37.

(o) (with Risto Hilpinen) "Knowledge, Acceptance, and Inductive Logic", in *Aspects of Inductive Logic*, Jaakko Hintikka and Patrick Suppes, editors, North-Holland Publishing Co., Amsterdam, 1966, 1-20.

(with Juhani Pietarinen) "Semantic Information and Inductive Logic", in *Aspects of Inductive Logic,* Jaakko Hintikka and Patrick Suppes, editors, North-Holland Publishing Co., Amsterdam 1966, 96-112.

1967

Papers

(a) "A.O. Lovejoy on Plenitude in Aristotle", *Ajatus* **29**, (1967), 5-11.

(b) "Existence and Identity in Epistemic Contexts. A Comment on Føllesdal's paper", *Theoria* **33**, (1967), 138-147.

Føllesdal defends the substitutivity of identity as a prerequisite for making sense of quantification into opaque contexts. Two versions of the principle must be distinguished, however, a free-variable form and a bound-variable form. The latter version can be incorporated in the author's treatment. Føllesdal's reasons for trying to uphold also the former version are criticized as restricting (or complicating) the applicability of modal logic, and as hiding interesting analogies between existential and uniqueness assumptions. No essential difference is found in the underlying semantical apparatus between Føllesdal and the author.

(c) "Individuals, Possible Worlds, and Epistemic Logic", *Nous* **1**, (1967), 33-62.

This paper is partly an answer to Castañeda's, Chisholm's and Sleigh's papers in the same number. Against Castañeda, it is argued that my treatment of quantification into modal contexts is demonstrably unavoidable given certain very natural semantical assumptions. Castaneda's handling of himself – locutions (logic of self-knowledge) can be greatly simplified. Against Chisholm, it is argued that his contrast between essentialism and the total failure of cross-identification between possible worlds is oversimplified. Sleigh's objections can be removed by adding to my earlier treatment an interesting new principle (essentially, Leibniz's law for bound variables).

(d) "Kant on the Mathematical Method", *The Monist* **51**, (1967), 352-375.

According to Kant, the mathematical method is characterized by the use of constructions, which he defines as exhibitions of an intuition corresponding to a general concept. Since by intuition Kant simply means a representation of a particular, Kant's characterization amounts to saying that the characteristic feature of the mathematical method is the use of instantiation rules. For this, there is a historical precedent in Euclidean propositions where a general theorem is always dealt with by first subjecting it to instantiation (*ekthesis*), by augmenting the instantiated configuration by means of auxiliary constructions, and then by carrying a proof proper (*apodeixis)* in terms of the augmented configuration. It is the use of auxiliary constructions that makes a mathematical argument synthetic for Kant, whereas *apodeixis* is for him analytic. Different aspects of Kant's theory are explained on the basis of this Kant-Euclid similarity.

(e) "Luovat mahdollisuudet ja päätöksenteko", in *Aikamme kaksi kulttuuria,* Eero Saarenheimo, editor, Werner Söderström Oy, Helsinki, 1967, 96-103. ("Creative Possibilities and Decision Making".)

(f) "New Essays on Old Philosophers", *Inquiry* **10**, (1967), 101-113.

(g) "A Program and a Set of Concepts for Philosophical Logic", *The Monist* **51**, (1967), 69-92.

The philosophical perspectives opened by (or associated with) certain developments in first-order logic are surveyed. Carnap's state-descriptions are made awkward as a tool in semantics by their infinitude. They can be modified in two different ways so as to make them more flexible tools in semantics. Thus we obtain either partial descriptions of possible worlds with unlimited vocabulary (model sets) or else maximally full descriptions with limited resources of expression. If this limitation restricts the number of individuals considered in relation to each other, we obtain the author's constituents. Some philosophical uses of these concepts are outlined.

(h) "Some Conceptual Presuppositions of Greek Political Theory", *Scandinavian Political*
Studies **2** (1967), 11-25.

(i)"Suomen filosofisen tutkimuksen tila ja tavoitteet", *Ajatus* **29**, (1967), 11-25. ("The State and the Aims of Philosophical Research in Finland".)

(j)"Time, Truth, and Knowledge in Ancient Greek Thought", *American Philosophical Quarterly* **4**, (1967), 1-14.

The Greeks had a tendency to think of temporally indefinite (e.g., now-sentences as typical vehicles of cognitive communication. This partly explains why for several of them all genuine knowledge was about the immutable. There was also a matching tendency to think of thoughts (opinions, propositions, lekta) as being similarly "temporally indefinite." As a consequence, Aristotle (among others) assumed that the truth-value of a sentence and of an opinion can change. These observations are tentatively related to the primacy of the spoken word over the written in Greek culture.

1968

Papers

(a) "Are Mathematical Truths Synthetic a Priori?", *The Journal of Philosophy* **65**, (1968), 640-651.

A sense of information (surface information) is distinguished from the usual (semantic) information (depth information) in which non-trivial quantificational reasoning adds to one's information, i.e., is (in a sense) synthetic. The latter may be thought of as the limit of everything that can be done to increase the former by drawing logical inferences. Is surface information about the reality or about our concepts? It is argued that it is inevitably and inextricably about both. The reason for this is the undecidability of first-order logic. This fact is seen as a partial rational reconstruction of the Kantian idea of the inaccessibility of things in themselves.

(b) "Behavioral Criteria of Radical Translation", *Synthese* **19**, (1968), 69-81.

This paper criticizes Quine's identification of that part of language which "can be made sense of in terms of its stimulus conditions" with what in a language is empirically conditioned and empirically ascertainable (e.g. for purposes of radical translation). As a counter-example the "language-games for quantifiers" are mentioned which are discussed by the author elsewhere (APQ, monograph series no. 2, 1968, 64-72). They enable in principle the empirical recognition of quantifiers without being characterizable in terms of stimulus meaning. The translation of propositional connectives in terms of these games is also discussed.

(c) "Conditionalization and Information", *Synthese* **19**, (1968), 303-306.

(d) "Epistemic Logic and the Methods of Philosophical Analysis", *Australian Journal of Philosophy* **46**, (1968), 37-51.

When applied to natural language, logic (e.g., epistemic logic) is not a regimentation or a translation of ordinary usage. Rather, it is an explanatory model (theory) by means of which such usage can be explained when certain other factors (mostly pragmatic) are also taken into account. Often, this explanatory model codifies one paramount purpose which a concept can serve. Often, its applications are much more indirect than ordinary language philosophers assume. These facts show that the notion of “having different senses” has different senses. The situation is illustrated by remarks on the problem of “knowing that one knows.”

(e) “Filosofinen ja mietekirjallisuus”, in *Suomen Kirjallisuus VII: Kirjallisuuden kenttä*, Matti Kuusi, editor, Suomalaisen Kirjallisuuden Seura ja Otava, Helsinki, 1968, 399-433. (“Philosophical and meditative literature”, in ‘Literature in Finland’.)

(f) “Induction by Enumeration and Induction by Elimination”, in *The Problem of Inductive Logic. Proceedings of the International Colloquium in the Philosophy of Science*, *London, 1965,* Imre Lakatos, editor, North-Holland Publishing Co., Amsterdam, 1968, 191-216.

The author’s inductive logic (for monadic first-order logic) combines ingredients from induction by elimination and induction by enumeration. The main way to raise the degree of confirmation of a constituent is to show that all cells (Q-predicates) it allows are actually instantiated, which means eliminating competing simpler hypotheses (constituents). After they have all been eliminated, the degree of confirmation grows with the number of positive instances. The alleged contrast between induction by elimination and induction by enumeration is thus spurious.

(g) “Language-Games for Quantifiers”, in *Studies in Logical Theory,* Nicholas Rescher, editor, American Philosophical Quarterly Monograph Series **2,** Basil Blackwell, Oxford, 1968, 46-72.

According to Wittgenstein, the words of our language have their meanings in virtue of the role they play in certain rule-governed activities (“language-games”) which link language and the world with each other. These activities are often described by verbs, to which the corresponding words thus have an especially close logical relation. For quantifier words like “some” and “every”, these “games” are the activities of seeking and finding. In order for quantifier words to have a meaning, the field of search, or at last its relevant part, must be given and criteria must be assumed for the end-points of search. Moore’s “proof of the external world” can be thought of as a dramatization of the latter need. Language-games of seeking and finding can be thought of as games in the strict sense of the mathematical theory of games.

(h) “Logic and Philosophy”, in *Contemporary Philosophy – La philosophie contemporaine*, vol. 1, R. Klibansky, editor, La Nuova Italia Editrice, Firenze, 1968, 3-30.

(i)“Meaning as Multiple Reference”, in *Proceedings of the Fourteenth International Congress of Philosophy* **1***,* Hölder-Verlag, Vienna, 1968, 340-345. (Appears also under the title “Semantics for Propositional Attitudes” in *Philosophical Logic*, W. Davis et al., editors, D. Reidel Publishing Co., Dordrecht, 1968, 21-45.)

(j)“On Semantic Information”, in *Physics, Logic and History. Proceedings of the International Colloquium on Logic, Physical Reality, and History, University of Denver*, W. Yourgrau, editor, Plenum Press, New York, 1968, 147-168.

(k) "The Possibility of Acceptance Rules", in *The Problem of Inductive Logic. Proceedings of the International Colloquium in the Philosophy of Science*, *London 1965* **2**, Imre Lakatos, editor, North-Holland Publishing Co., Amsterdam, 1968, 98-119.

(l)"Reply", in *The Problem of Inductive Logic. Proceedings of the International Colloquium in the Philosophy of Science, London, 1965* **2**, Imre Lakatos, editor, North-Holland Publishing Co., Amsterdam, 1968, 223-231.

(m) "Review: *The Encyclopedia of Philosophy*, Paul Edwards, Editor-in-Chief, Collier-Macmillan/ The Macmillan Co., London/ New York, 1967, vols. 1-8", *Synthese* **19**, (1968-69), 466-469.

(m) "The Varieties of Information and Scientific Explanation", in *Logic, Methodology, and Philosophy of Science III. Proceedings of the 1967 International Congress*, B. van Rootselaar and J.F. Staal, editors, North-Holland Publishing Co., Amsterdam, 1968, 151-171.

Several different senses and varieties of information can be distinguished from each other, among them substantial information vs. surprise value, incremental information vs. conditional information, information concerning the subject matter of a certain proposition, expected information, etc. Scientific inquiry can on different occasions aim at the maximization of different types of information depending on the character of the inquiry. E.g., the aim may be the explanation of the given date (local theorizing) or it may aim at a general theory using the data as a stepping-stone (global theorizing). Since different aims lead to different methods, there is no unique "scientific method." E.g., local theorizing leads naturally to the maximal likelihood principle, whereas philosophers' different methods of explanatory power are related to the idea of global theorizing.

1969

Books

(a) *Models for Modalities,* D. Reidel Publishing Co., Dordrecht, 1969, x+220pp.

The following theses are argued for: (i) The logic and logical semantics of a concept should be thought of as codifying its "depth logic" which is not a generalization from surface data. Rather, those data should be explained as a joint result of depth logic plus sundry pragmatic and other contextual factors (chapter 1). (ii) The logical rules of the model set technique can be modified so as to allow empty singular terms (names) by using (Ex) (a=x) as the explication of "a exist." It can be shown (on extremely plausible assumptions) that any "predicate of existence" must be logically equivalent with (Ex) (a=x). This can be taken as a vindication of Quine's dictum that "to be is to be a value of a bound variable" (chapter 2). (iii) The ontological argument can be thought of as a dramatization of certain features of first-order logic with possibly empty singular terms (chapter 3). (iv) An explicit modal logic is formulated by means of the model set technique by associating to each model set considered a set of alternatives. Different Lewis-type systems are captured by imposing different conditions (transitivity, symmetry etc.) on the alternativeness relation. Different transfer assumptions for identity and existence are studied and their manifestations in the validity of different formulas (e.g., the Barcan formula) are studied (chapters 4 and 5). (v) Different types of resulting logics are seen to be the appropriate logics of different concepts (epistemic, temporal, etc.) (chapter 5). (vi) As a model-theoretic counterpart of such logics, a possible-worlds semantics is developed for different propositional attitudes. A crucial role is then played by the "individuating functions" defining identities of individuals in different worlds (chapter 6). (vii) Formal counterparts (uniqueness presuppositions) of such functions are studied in chapter 7. (viii) The logic and semantics developed in earlier chapters is applied to the analysis of perception. A characteristic feature of the semantics of perception is the presence of two irreducibly different methods of cross-identification, the perceptual and the physical. Perceptually identified individuals are shown to be like philosophers' sense-data in certain respects. Philosophers' views of sense-data can be discussed and evaluated by means of this similarity. The contrast between physical and perceptual cross-identification is manifested in the contrast between wh-constructions (e.g., "seeing who someone is") and the direct-object construction (e.g., "seeing someone") (chapter 8). (ix) Deontic logic is dealt with by the same model set (and model system) technique. Several conceptual problems are analysed by means of this treatment: (a) certain mistakes of earlier deontic logicians are traced to a confusion between logical validity (truth in all logically possible worlds) and deontic validity (truth in all deontically perfect worlds); (b) the ideas "ought implies can" and Kantian "Kingdom of ends" are analyzed; (c) Searle's "derivation of ought from is" is shown to depend on an equivocation between two logically different kinds of obligation (chapter 9).

(b) *Tieto on valtaa: ja muita aatehistoriallisia esseitä*, Werner Söderström Osakeyhtiö, Helsinki, 1969, 298pp. (*"Knowledge is Power" and Other Essays in the History of Ideas*.)

(c) (edited) *Philosophy of Mathematics*, Clarendon Press, Oxford, 1969, 186 pp.

(d) (edited with Donald Davidson) *Words and Objections: Essays on the Work of W.V. Quine*, D. Reidel Publishing Co., Dordrecht, 1969, vii+366 pp.

Papers

(a) "Deontic Logic and its Philosophical Morals" in *Models for Modalities,* D. Reidel Publishing Co., Dordrecht, 1969, 184-214.

(b) "Inductive Independence and the Paradoxes of Confirmation", in *Essays in Honor of Carl G. Hempel,* Nicholas Rescher et al., editors, D. Reidel Publishing Co., Dordrecht, 1969, 24-46.

(c) "Leibniz, Plentitude, Relations and the 'Reign of Law'", *Ajatus* **31**, (1969), 117-144.

(i) Leibniz did not try to reduce relations to primitive properties, but relational propositions to non-relational propositions, whose subject and predicate can still contain relational concepts. (ii) Leibniz's ideas of infinite analysis, of a monad's reflecting the entire world, etc. can be explicated in terms of recently developed infinitely deep logics. (iii) Much of Leibniz's philosophy, especially his criticism of Descartes, stems from his rejection of the so-called principle of plentitude (realization of all possibilities in time). Leibniz assumed it to be applicable in such a way that the principle would imply the denial of all unrestricted natural laws.

(d) "On Kant's Notion of Intuition (Anschauung)", in *The First Critique: Reflections on Kant's Critique of Pure Reason,* Terrence Penelhum and J.J. Macintosh, editors, Wadsworth, Belmont, CA., 1969, 38-53.

By intuition (*Anschauung*), Kant by definition meant a representation of a particular. This is connected with the prevalent earlier use of the term, in which intuitiveness meant immediacy, by the idea that a general term represents the object or objects to which it applies only through the mediation of the general characteristics they have when the term applies to it or to them. This is the force of the term in Kants theory of the mathematical method as involving the use of intuitions, by which Kant only means instantiations of general concepts. This is the force of the terms also in Kant's transcendental aesthetic. This sense implies no "intuitiveness", in our sense, e.g., no particular relation to sense-perception or imagination. (It was in this sense that Kant called algebraic symbols intuitive.) As a result of Kant's arguments in his "transcendental aesthetic", he thinks that he has established such a connection, however, and goes on to assume the intuitiveness of intuitions in the rest of his *Critique of Pure Reason.*

(e) "On the Logic of Perception", in *Perception and Personal Identity,* Norman S. Care and Robert H. Grimm, editors, Case Western Reserve University Press, Cleveland, OH, 1969, 140-175.

(f) "On the Logic of the Ontological Argument: Some Elementary Remarks", in *The Logical Way of Doing Things*, Karel Lambert, editor, Yale University Press, New Haven, 1969, 185-197.

(g) "Partially Transparent Senses of Knowing", *Philosophical Studies* **20**, (1969), 4-8.

If knowledge is construed transparently, 'it is known that A = A' entails '(Ex) it is known that A = x, i.e., entails that it is known who A is. Sleigh has claimed in *Philosophical Studies* **18** (1967), 12-14, that this absurdity affects not only Quine but also my reconstruction of the transparent sense in knowledge and belief. His charge is rebutted by pointing out that the transparent-opaque distinction is to be applied to each occurrence of a singular term separately. This yields a variety of 'partially transparent' constructions. All of them are formalizable in my approach.

(h) "Quantification Theory and the Picture Theory of Language", *The Monist* **55**, (1969) 204-230.

A version of the picture theory of Wittgenstein's *Tractatus* is extended to all first-order (quantificational) languages, by interpreting the author's model sets as partial pictures (isomorphs) of possible worlds in which their members are true, each quantificational proof (say of S) thus becomes a frustrated attempt to depict a world in which S were false. The picture theory cannot any longer serve as a model of the use of language, however, for the main emphasis is shifted on stepwise comparisons between (possibly infinite) model sets and the world. Their "logic" is more fundamental than the picture theory.

(i)"Semantics for Propositional Attitudes", in *Philosophical Logic,* W. Davis et al., editors, D. Reidel Publishing Co., Dordrecht, 1969, 21-45.

(j)"Statistics, Induction and Lawlikeness: Comments on Dr. Vetter's paper", *Synthese* **20**, (1969), 72-85.

Two related subjects are discussed: the relation of mathematical statistics and inductive generalization. It is argued that Veiter's reliance on current statistics is uncritical. Traditional statistical methods are insufficient to understand inductive generalization, and hence cannot be invoked to criticize inductive logicians' work on this subject, which ought to present a challenge to statisticians, too. My alpha-continuum of inductive methods is defended against charges of arbitrariness by relating it to the notion of lawlikeness. Sundry corrections to Veiter's paper are made.

(k) "Tieteen metodi analyyttisena toimituksena", in *Societas Scientiarum Fennica, Yearbook*, **95B**, no. 2, Helsinki, 1969. ("Scientific Method as an Analytical Procedure".)

(l)"Wittgenstein on Private Language: Some Sources of Misunderstanding", *Mind* **78**, (1969), 423-425.

Two passages from Wittgenstein's *Philosophical Investigations* are discussed: Part 1, sections 245 and 265. It is suggested that the force of the former is to pose an objection to be disposed of, not to offer a reason for Wittgenstein's own position. The operative words of the latter ("die richtige Erinnerung") can be read. It is suggested, not as referring to the actual correctness of memory, but rather to the right choice among several available memory-images. In both cases, a more indirect connection between sensations and their external manifestations seems to be intended than the standard interpretation allows for.

1970

Books

(edited with Lauri Routila) *Filosofian tila ja tulevaisuus*, Weilin+Göös Ab:n kirjapaino, Tapiola/Helsinki, 1970, 229 pp. (*The State and Future of Philosophy.*)

(edited with Patrick Suppes) *Information and Inference*, D. Reidel Publishing Co., Dordrecht, 1970, vii+336pp.

The main problems discussed are: theory of definitions, problems of experiment, problem of entropy, probability theories, problems of measurement and structure, learning models, etc. The book shows that the concept of information may render valuable services to various philosophical and scientific investigations, (Bp, edited)

Papers

(a) "Creative Process, Crystallization and Cumulation?", in *Scientists at Work: Festschrift in Honour of Herman Wold,* Tore Dalenius et al., editors, Almqvist och Wiksell, Stockholm, 1970, 62-65.

(b) "Existential Presuppositions and Uniqueness Presuppositions", in *Philosophical Problems in Logic: Some Recent Developments. Proceedings of the Irvine Colloquium*, *May 1968*, K. Lambert, editor, D. Reidel Publishing Co., Dordrecht, 1970, 20-55.

"Filosofian looginen välineistö" ("Logical Tools of Philosophy") in *Filosofian tila ja tulevaisuus*, Jaakko Hintikka and Lauri Routila, editors, Weilin-Göös Ab:n kirjapaino, Tapiola/ Helsinki, 1970, 195-220. (*The State and Future of Philosophy.*)

(d) "Inductive Generalization and Its Problems: A Comment on Kronthaler's Comment", *Theory and Decision* **1**, (1970), 393-398.

In reply to Kronthaler's note in the same number, the following main points are made: (1) Kronthaler notwithstanding, universal statements obviously can be corroborated by means of limited evidence. (2) The purpose of the alpha-continuum was to enable us to conceptualize assumptions of lawlikeness, not to solve "the problem of induction." (3) Parameters like alpha are useful, not because their values can be fixed for good, but because they enable us to take into account useful background information by choosing their values differently on different occasions. (4) Kronthaler misunderstands the relation between priors and "anpassungsgeschwindigretten."

(e) "Information, Deduction, and the A Priori", *Nous* **4**, (1970), 131-152.

A sense of information (surface) is defined (for first-order languages) in which deduction can yield new information. It differs from usual semantic information (depth) in that inconsistent sentences receive non-zero informational weights unless they are (in a specifiable sense) trivially inconsistent. Is surface information about the reality or about our conceptual system? It is inextricably both, we argue. This answer is related to some classical problems concerning a priori knowledge.

(f) "'Knowing That One Knows' Reviewed", *Synthese* **21**, (1970), 141-162.

The semantical basis of the KK-thesis (implication from knowing to knowing that one knows) is explained and defended against objections by Chisholm and others. It presupposes a very (perhaps unrealistically) strong sense of knowing. This is compared with Malcolm's strong sense. A qualification if needed because of the de re character of the thesis, which essentially restricts it if logical omniscience is not presupposed. A comparison with Sarire's pre-reflective cogito is made.

(g) "Knowledge, Belief, and Logical Consequence", *Ajatus* **32**, (1970), 32-47.

(h) "Kolme itävaltalaista rautatieasemaa ja kiinalainen onnenpeli: Huomioita Grazin filosofikokouksesta 1-4.10.1970", *Parnasso* **20**, (1970), 512-515. ("Three Austrian Railway-stations and a Chinese Game of Chance: Observations From the Meinong-Colloquium in Graz, 1-4 October 1970".)

(i)"Kontinuumiongelma ja joukko-opin aksiomatiikan probleemat", *Arkhimedes* **22**, (1970), 1-7. ("The Continuum Problem and the Problems of Axiomatic Set Theory".)

(j)"Kybernetiikka ja yhteishuntatieteiden metodologia", *Sosiologia* **7**, (1970), 217-225. ("Cybernetics and the Methodology of Social Sciences".)

(k) "Nykyinen logiikka filosofian apuvälineenä", in *Logiikka ja matematiikka-Studia Logica et Mathematica,* Werner Söderström Oy, Porvoo-Helsinki, 1970, 41-60. ("Modern Logic as a Tool in Philosophy".)

(l)"Objects of Knowledge and Belief: Acquaintances and Public Figures", *The Journal of Philosophy* **67**, (1970), 869-883.

Individuals can be identified between different possible worlds in two ways, descriptively and by acquaintance. The latter depends on someone's first-hand cognitive relationships to individuals, and identifies them if they play the same role in these relations. If "B" picks out the same individuals from all worlds compatible with A's memories, A remembers who or what B is; if the same individual by acquaintance, A remembers B. This yields a semantical analysis of the direct-object construction with "remembers" (and likewise with "knows", "perceives", etc.).

(m) "On Attributions of 'Self-Knowledge'", *The Journal of Philosophy* **67**, (1970), 73-87.

This is a reply to a paper by H.N. Castañeda in the same journal. The following points are made: (1) ordinary language should be interpreted by means of logical semantics, not vice versa. (2) Castañeda misinterprets the relation of my condition to instantiation. (3) Castañeda's example of the amnesiac hero is solved by a distinction between different methods of individuation (descriptive and by acquaintance). (4) Castaneda's back-reference notation is compared with a semantical theory of the subject.

(n) "On Kant's Background", *Ajatus* **31**, (1970), 164-170.

(o) "On Semantic Information", in *Information and Inference,* Jaakko Hintikka and Patrick Suppes, editors, D. Reidel Publishing Co., Dordrecht, 1970, 3-27.

(p) "Philosophy of Science (Wissenschaftstheorie) in Finland", *Zeitschrift für allgemeine Wissenschaftstheorie* **1**, (1970), 119-132.

This is a brief survey of recent work in the philosophy of science in Finland. The main sources of influence emphasized are Eino Kaila (1890-1958) and G.H. von Wright (b. 1916). The main topics covered are: induction and probability; information and explanation; the role of auxiliary (theoretical) terms; measurement; general methodology of social and behavioral sciences; finalistic explanation; methodology of sociology and history.

(q) "'Prima Facie' Obligations and Iterated Modalities", *Theoria* **36**, (1970), 232-240.

A reply to Trandy's paper in the same issue pp. 221-231. Main points: (1) A semantic system of doentic logic stands on its own feet, and (2) does not have to rely on any straightforward translation into ordinary language. (3) My semantics deals with 'seinsullen' rather then 'tunsullen'. (4) It associates automatically a clear sense to the iteration of deontic operators. (5) The notion of permission it relies on is essentially that of Hohfeld's liberty. (6) Trandy notwithstanding, it does distinguish between obligation and necessity.

(r) "The Semantics of Modal Notions and the Indeterminacy of Ontology", *Synthese* **21**, (1970), 408-424.

Quantification into modal contexts depends on cross-identifications of individuals between possible worlds, which in turn depends on the structure and interrelations of these worlds. There is hence no guarantee that cross-identification always succeeds. It will fail for the worlds needed for realistic applications of logical modalities, partly vindicating Quine's criticism of them. In general, world lines of individuals cannot always be extended from a world to others.

(s) "Statistics, Induction and Lawlikeness: Comments on Dr. Vetter's Paper", in *Induction, Physics, and Ethics*. *Proceedings and Discussions of the 1968 Salzburg Colloquium in the Philosophy of Science*, P. Weingartner and G. Zech, editors, D. Reidel Publishing Co., Dordrecht, 1970, 91-102.

(t)"Surface Information and Depth Information", *Information and Inference,* Jaakko Hintikka and Patrick Suppes, editors, D. Reidel Publishing Co., Dordrecht, 1970, 263-297.

Measures of probability and information are defined for the sentences of a finite first-order (quantificational) language in a natural way such that logically equivalent sentences don't always have the same probability of information. The resulting "surface probability" satisfies all the other axioms of probability calculus. A valid logical argument from p to q yields more surface information if and only it it is nontrivial in the sense of 1973(a). The usual "depth" probability and "depth", information are related to surface notions. There is a sense in which the depth information of p is the limit of its surface information when more and more of the logical consequences of p are uncovered, e.g., by turning it into deeper and deeper distributive normal form. (Note: A model theory for surface notions can be obtained by means of Rantala's urn models along the lines sketched in (1975(b).)

(u) "Two Studies in Probability", in *Reports from the Institute of Philosophy, University of Helsinki*, (1970), 58 pp.

(v) (with Raimo Tuomela) "Towards a General Theory of Auxiliary Concepts and Definability in First-order Theories", in *Information and Inference,* Jaakko Hintikka and Patrick Suppes, editors, D. Reidel Publishing Co., Dordrecht, 1970, 298-330.

1971

Papers

(a) "Different Kinds of Equivocation in Aristotle", *Journal of the History of Philosophy* **9**, (1971), 368-372.

The interrelations of (1) synonymy, (2) homonymy, and (3) the intermediate class of "pollakhos legetai" in Aristotle are studied here. The independence of (3) "vis-à-vis" (2) is defended against G.E.L. Owen. The role of development of (3) is emphasized. In Aristotle, (3) "owes its genesis as much to the breakdown of the homonymy-amphiboly distinction as to the breakdown of the synonymy-homonymy dichotomy."

(b) "Inductive Generalization and Its Problem: A Comment on Kronthaler's Comment", *Theory*

and Decision **1** (1971), 393-398.

(c) "Knowledge and its Objects in Plato", *Ajatus* **33**, (1971), 168-200.

Plato thought of the relation of such 'dynameis' as knowledge, belief, saying, etc., to their objects as something like 'aiming at' or 'trying to realize themselves in' these objects. Several consequences of this observation are studied: (1) "Euthydemus" 284 B-C represents a serious temptation; (2) difference between objects implies difference between the 'dynameis'; (3) hence Plato's argument for the forms in "Rep" 475-480; (4) meaningful but false saying becomes a problem, for meaningful (successful) saying seems to involve realization of the 'dynamis' in its objects, i.e., truth; (5) by the same token, true belief and knowledge seem to be inseparable.

(d) "The 'Lottery Paradox' and the Concept of Shared Information", *Ajatus* **33**, (1971), 266-270.

The lottery paradox always involves a situation in which P(S) is greater than or equal to T – E, P(T) is greater than or equal to T – E. P(S & T) is greater than or equal to T – E, where T – E is the acceptance level. A condition for this not to happen is that T – P(S or T) is greater than or equal to E. This can be interpreted as saying that S and T must have a certain minimum amount of information in common (in the sense of transmitted content). (Edited)

(e) "On Defining Information", *Ajatus* **33**, (1971), 271-273.

Suppose that we first know S and then come to know T. The increase of information at the second step is cont(S & T) – cont(S). It must be assumed that this equals the information of the weakest additional premise which jointly with S implies T, i.e., equals cont(CST). It is shown that this equation justifies, together with certain normalizing assumptions, the usual difinition cont(R) = T – P(R). P(R) being the probability of R.

(f) "On the Ingredients of an Aristotelian Science", *Reports from the Institute of Philosophy: University of Helsinki***,** no. 3, (1971). (An early version of "On the Ingrediants of an Aristotelian Science" *Nous* **6**, (1972), 55-69.)

(g) "Semantics for the Prepositional Attitudes", in *Reference and Modality*, Leonard Linsky, editor, London/New York, Routledge & Kegan Paul, 1971, 145-167.

(h) "Some Main Problems of Deontic Logic", in *Deontic Logic: Introductory and Systematic Readings,* Risto Hilpinen, editor, D. Reidel Publishing Co., Dordrecht, 1971, 59-104.

(i)"Sosa on Propositional Attitudes de dicto and de re", *The Journal of Philosophy* **68**, (1971), 489-497.

Comments on Sosa's paper in the same issue, pp. 883-896. Sosa's main question concerning 'exportability' (validity of existential generalization) is an apt one, but he neglects the resources of semantics. As a consequence, Sosa is led to exaggerate the role of pragmatic and contextual factors and to neglect e.g. the connection between exportation and who- and what-constructions. Several examples of Sosa's are analyzed and shown not to support his conclusions. His main statements of the conditions of exportability can nevertheless be accepted with relatively small changes.

(j)(with Risto Hilpinen) "Rules of Acceptance, Indices of Lawlikeness, and Singular Inductive

Inference: Reply to a Critical Discussion", *Philosophy of Science* **38**, (1971), 303-307.

1972

Papers

(a) "Concept as vision. Todellisuuden esittämisen ongelmasta modernissa kuvataiteessa ja modernissa filosofiassa", *Aika* **66**, (1972), 133-146. ("On the Problem of Representation in Modern Art and Modern Philosophy".)

(b) "Constituents and Finite Identifiability", *Journal of Philosophical Logic* **1**, (1972), 45-52.

Syntactic criteria are formulated in terms of the author's constituents for (i) the finite definability (identifiability) and (ii) the restricted identifiability (i. Up to the cardinality of the domain) of a given predicate in a first-order theory T. They are manifested by the disappearance of certain uncertainty sets in the members of an expansion of T. But (unlike the case of explicit or piecewise definability) these sets may occur inside constituents, not in the outmost layer of quantifiers.

(c) "Different Constructions in Terms of the Basic Epistemological Concepts: A Survey of Some Problems and Proposals", in *Contemporary Philosophy in Scandinavia,* Raymond Olsen and Anthony M. Paul, editors, The Johns Hopkins Press, Baltimore and London, 1972, 105-122.

"Die Intentionen der Intentionalität", *Neue Hefte für Philosophie,* Martinus Nijhoff Publishers, Netherlands, 1972.

"Kantian Intuitions", *Inquiry* **15**, (1972), 341-345.

By way of a reply to Charles Parson's paper in the Nagel Festschrift, Kant's notion of intuition (Anschauung) is examined. It is argued that for Kant the immediate relation which an intuition has to its object is a mere corollary to its singularity. It does not presuppose (as Parsons suggests) any presence of the object to the mind. This is shown, e.g., by the prolegomena section 8, where the objects of intuitions a priori are denied by Kant to be present, they yield knowledge, not in virtue of their immediacy but in virtue of their ideality.

(f) "Knowledge by Acquaintance — Individuation by Acquaintance" in *Bertrand Russell: A Collection of Critical Essays*, David Pears, editor, Anchor Books/ Doubleday and Co.,Garden City, NJ, 1972, 52-79.

Russell's distinction between knowledge by description and knowledge by acquaintance is not only epistemological in nature, but also semantical, concerning the entities whose existence is presupposed in the semantics of our language. The distinction is here traced back to a contrast which in a possible-worlds treatment can be found between two kinds of methods of cross-identification, by means of the criteria we use in deciding whether it is known who or what someone or something is or by means of the criteria we use in deciding whether one knows someone or something. This distinction is parallel to a distinction between perceiving who or what and perceiving someone or something. This rational reconstruction of Russell's distinction is related to his use of sense-data as objects of acquaintance, his notion of "logically proper name", and his attempted "reduction to acquaintance."

(g) "Leibniz on Plenitude, Relations, and the Reign of Law", in *Leibniz: A Collection of Critical Essays,* Harry Frankfurt, editor, Anchor Books, Doubleday and Co., Garden City, N.J., 1972, 155-190.

(h) "Mitä on kybernetiikka? in *Mitä-Missä-Milloin: Kansalaisen vuosikirja 1972*, Paul Kojo et al., editors, vol. 23, Otava, Helsinki, 1972, 294-295. (What is cybernetics?")

(i)"On the Ingredients of an Aristotelian Science", *Nous* **6**, (1972), 55-69.

Because an Aristotelian science relied on syllogistic logic, Aristotelian scientific inferences use sequences of nested terms. The primitive assumptions of any one science were hence threefold: (i) common axioms, assumptions shared by all sciences; (ii) atomic premises, dealing with the connections between two adjacent terms; (iii) generic premises, postulating the existence of the genus studied in the science in question. Assumptions (iii), but not (ii), carry existential force. Aristotle sometimes calls assumptions (ii) and even (iii) definitions. This is argued for among other things by reference to the text of *Post An.* 1, 2 and 10 and to the force of the terms *deixis* and *apodeixis* in Aristotle.

(j)"Transcendental Arguments – Genuine and Spurious", *Nous* **6**, (1972), 274-281.

Kant called knowledge transcendental when it deals with 'the mode of our knowledge of objects … a priori' (A11=B25) – knowledge which is possible only of 'what we ourselves put into' objects (B xvii). The suggestion is made that transcendental arguments for Kant dealt essentially with this contribution of the human mind to our knowledge. The views of Strawson, Gram (*Nous* **5** (1971), 135-172) and others of so-called transcendental arguments are criticized from this point of view.

(k) "Valtasuhteet, määräenemmistösuhteet ja parlamentarismin luonne", *Aika* **66**, (1972) 406-411. (Power Relations, Qualified Majority Rules, and the Nature of Parliamentary Democracy".)

(l)"Some Main Problems in Epistemic Logic: Two comments", *Ajatus* **34**, (1972), 144-148.

This note is a reply to the notes by Tomberlin and Wu in the same volume. It is suggested that Tomberlin's desiderata might be achieved by not assuming the epistemic law "K (P) and K(Q) implies K(P and Q)", the semantic interpretation of the resulting system is outlined.

1973

Books

(a) *Logic, Language-games, and Information*, Clarendon Press, Oxford, 1973, pp. x+291pp.

The two leading ideas of this book are: (i) the logic, semantics and pragmatics of quantifiers are all based on their role in certain language-games of seeking and finding; (ii) a distinction can be made between nontrivial and trivial logical inferences depending on whether new individuals have to be brought to bear on the inference. As a background and a set of tools, two different logical techniques are expounded: the model set method (chapter 1) and the theory of constituents and distributive normal forms (chapter 11). The following more specific theses are argued for: (a) Model sets can serve as "pictures", in the sense of Wittgenstein's so-called picture theory (chapter 2). (b) Language-games of seeking and finding can be used for the translation of quantifier words (chapter 4). (c) These language-games are knowledge-seeking activities in the sense in which Kant in his transcendental philosophy emphasizes such activities (chapter 5). (d) Trivial logical inferences are analytic in a natural sense of the term while nontrivial ones are not (chapter 6). (e) Nontrivial logical arguments are not tautological in any natural sense (chapter 7). (f) Kant's analytic-synthetic distinction is in the area of mathematical reasoning tantamount to the trivial-nontrivial one. In particular, Kant's idea that interindividual existential inferences are synthetic can be vindicated (chapter 8). (g) Kant's notion of analyticity is a natural sequel to earlier ideas of analysis and analyticity (chapter 9). (h) Senses of information (measures of information) can be defined in which logical inferences increase one's information if they are nontrivial. Such information can be said to be conceptual in a sense and about the world in another (chapter 10).

(b) *Time and Necessity: Studies in Aristotle's Theory of Modality,* Clarendon Press, Oxford, 1973, vi + 225 pp.

Several interrelated interpretational results are argued for: (i) Aristotle distinguishes between outright ambiguity (homonymy) of a word and its being used in many ways (*pollakhos legetai*). (ii) Aristotle distinguishes substantially but not terminologically "contingency" and "possibility proper", only the latter of which includes cases of necessity. (iii) In the light of (i)-(iii), we can understand *De Int.* 12-13 if we realize that *akolouthein* does not there mean logical following but literally "going together." (iv) The paradigmatic vehicle of communication for Greek philosophers was a temporally indefinite ("now") sentence depending on the moment of its utterance. This explains several features of the views of Greek philosophers, e.g., their ideas that the truth-value of a belief or proposition could change and that we can have knowledge only of what is unchangeable. (v) Aristotle assumed that no possibility can remain unfulfilled through an infinity of time. (vi) Aristotle did not think that the infinite is "potential but never actual"; rather, he thought that the infinite *is* potentially *and* actually in an unusual sense of existence. (vii) Aristotle did not deny the applicability of *tertium non datur* to contingent future events. Rather, he distinguished what can be said of temporally definite statements (which according to (v) must be necessary if true) and temporally indefinite statements about them. (viii) Aristotle uses an argument apparently parallel with the Master Argument of Diodorus to defend (v) instead of arguing for determinism. This enables us to conjecture the line of thought used in the Master Argument.

(c) (edited with Julius M.E. Moravcsik and Patrick Suppes) *Approaches to Natural Languages*. *Proceedings of the 1970 Stanford Workshop on Grammar and Semantics*, D Reidel

Publishing Co., Dordrecht, 1973, viii+526 pp.

The papers are arranged under three sub-titles, viz. grammar, semantics, and special topics. Part I on grammar contains papers of stress, phonology, syntax, transformational grammars, and the like. Part II on semantics contains, among other things, papers on the relations between grammar and on belief sentences. Special topics treated in Part III are topicalization and self-reference. Some papers in Parts II and III are followed by special comments. The volume contains 27 essays.

Papers

(a) "Aristotle on the Realization of Possibilities in Time", in Jaakko Hintikka, *Time and Necessity: Studies in Aristotle's Theory of Modality*, Clarendon Press, Oxford, 1973, 93-113.) (A fuller version of "Necessity, Universality and Time in Aristotle", *Ajatus* **20**, 1957), 65-90.)

(b) "Carnap's Semantics in Retrospect", *Synthese* **25**, (1973), 372-397.

Carnap's work in logical semantics is a culmination of the Fregean tradition based on the contrast extension vs. intension. In characterizing his intensions Carnap came close to the subsequent 'possible-worlds semantics' of Montaque, Kripke, and others. Carnap was prevented from developing a full-fledged possible-world as distinguished from an arbitrarily defined logical model. Only the former makes it possible to analyze such propositional attitudes as belief semantically, thus improving essentially on Carnap's discussion. Possible-worlds semantics vindicates the use of intensions in dealing with the problems of identity, but it shows that traditional intensions fail to cope with quantification into nonextensional contexts. Possible-worlds semantics also opens new avenues for behaviouristic interpretations of intensional concepts and of notions like belief. In this respect, too, possible-worlds semantics carries further certain Carnapian ideas.

(c) "Grammar and Logic: Some Borderline Problems", in *Approaches to Natural Languages,* Jaakko Hintikka, Julius M. E. Moravcik and Patrick Suppes, editors, D. Reidel Publishing Co., Dordrecht, 1973, 197-214.

(d) "Knowing How, Knowing That, and Knowing What: Observations on Their Relation in Plato and Other Greek Philosophers", in *Modality, Morality, and Other Problems of Sense and Nonsense: Essays Dedicated to Sören Hallden*, C.W.K. Gleerup, Lund, 1973, 1-12.

(e) "On the Different Ingredients of an Empirical Theory", in *Logic, Methodology, and the Philosophy of Science,* Patrick Suppes et al., editors, North-Holland Publishing Co., Amsterdam, 1973, 313-322.

(f) "Quantifiers, Language-games, and Transcendental Arguments", in *Logic and Ontology,* Milton K. Munitz, editor, New York University Press, New York, 1973, 37-57.

(g) "Quantifiers vs. Quantification Theory", *Dialectica* **27**, (1973), 329-358.

Quantifiers can be dealt with in game-theoretical semantics by allowing the verifier ("myself") to choose the values of existentially bound variables and the falsifier ("nature") to choose the values of universally bound quantifiers. By allowing imperfect information we can obtain a semantics for partially ordered (e.g., branching) quantifiers. It is argued that such quantifier prefixes are exemplified by semantics of English sentences.

(h) "Remarks on *Poiesis, Praxis* and *Ergon* in Plato and Aristotle", in *Studia Philosophica in Honorem Sven Krohn,* Timo Airaksinen and Risto Hilpinen, editors, Turun Yliopisto, Turku, 1973, 53-62.

(i)"Surface Semantics: Definition and its Motivation", in *Truth, Syntax, and Modality. Proceedings of the Temple University Conference on Alternative Semantics*, Hughes Leblanc, editor, North-Holland Publishing Co., Amsterdam, 1973, 128-147.

A kind of model ("possible world") can be described by specifying what ramified sequences of individuals (to a given length d) an observer might come upon in such a world. Labelled tree diagrams specifying them are called surface models, on one further condition: that its different parts match as closely as possible, reflecting the requirement that the set of individuals the observer may come upon is the same at each stage of the investigation of the world. Such surface models are described by the constituents of 1953(a). Hence each proposition in a first-order language admits some surface models of its own depth and excludes the others. Not every surface model corresponds to possible world (model). It corresponds to one if it can be extended indefinitely by increasing d to infinity.

(j)"Theoretical Terms from Ramsey's Reductions: Outline of Scientific Logic", translated from Russian, V.A. Markov, translator, in *Scientific Lectures for Higher Education: Philosophy of Science,* (1973), 49-61.

(k) (with Ilkka Niiniluoto) "On the Surface Semantics of Proof Procedures", *Ajatus* **35**, (1973), 197-215.

The concepts of Hintikka's surface semantics (see H. Leblanc, editor, *Truth, Syntax, and Modality"* 1973) are applied to the evaluation of the efficiency of different first-order proof techniques. By means of an example it is suggested that they cannot be more effective than the use of distributive normal forms, when efficiency is measured in terms of the nonextendable surface models the proof techniques rule out.

1974

Books

(a) *Knowledge and the Known: Historical Perspectives in Epistemology,* D. Reidel Publishing Co., Dordrecht, 1974, xii+ 243 pp.

The following historical or systematic points are argued: (i) Plato dealt with the concepts of knowledge, thinking and saying on a goal-directed model. This explains several peculiarities of his philosophy, including his preoccupation with the problem of meaningful falsehood and his emphasis on knowledge-based skills as unerring skills (chapter 1). It also explains Plato's identification of knowing what X is (what its definition is) with ability to bring about X and Plato's use of a craftsman as a conceptual model (chapter 2). (ii) See chapter 4 of 1973(a) (chapter 3). (iii) There is a historical tradition maintaining the superiority of knowledge of objects which one has brought about or has in one's power ("maker's knowledge"). This tradition includes Vico and Kant. The superiority claim nevertheless leads into systematic problems (chapter 4). (iv) See 1961(d) (chapter 5). (v) Kant's transcendental method led him to explain the peculiarity of mathematical method, which he took to be the use of constructions (by which he in effect meant instantiations) by considering the way we come to know the objects of applied mathematics, which are for him particular objects. This way he (mistakenly) identified with sense-perception. Hence he concluded that mathematics is based on the structure of our faculty of sense-perception. This way of looking at Kant's theory of mathematics is supported by a close analogy of his description of the mathematical method and Euclid's expositional practice (chapter 6, 8). (vi) See 1965(b), (chapter 7). (vii) If Kant's "things in themselves" are interpreted as objects as they are independently of the effects of our knowledge-seeking activities and the conceptual framework they use, then first-order logic offers an example of their unknowability. For because of its undecidability, we cannot ever eliminate all merely apparent possibilities which a proposition admits concerning the world (chapter 10). (vii) See 1969(d), (chapter 11).

(b) *Induzione, accettazione, informazione*, ed. e trad. da Marco Mondadori e Paolo Parlavecchia, Societa editrice il Mulino, Bologna, 1974.

(c) (with Unto Remes) *The Method of Analysis: Its Geometrical Origin and Its General Significance*, D. Reidel Publishing Co., Dordrecht 1974, xviii+ 144 pp.

The nature of the Greek geometrical method of analysis and synthesis is analyzed, especially by reference to Pappus' description of the method. It is argued that Pappus does not assume that the analytic passage from the desired consequence to known premises is deductive. The inevitable role of auxiliary constructions is noted and the difficulties it presented to the theoretical understanding of the method of analysis and synthesis (e.g., the two processes cannot be mirror images of each other) are discussed. What is analyzed in geometrical analysis is a geometrical configuration, not a geometrical proof. This feature explains the use of analysis as a paradigm of the experimental methodology of early modern scientists, who were analyzing physical configurations by essentially the same method.

Papers

(a) "Concept as Vision", *Iyyun* **25,** (1974), 139-157.

(b) "'Dinge an Sich' Revisited", in *Akten des 4. Internationalen Kant-Kongresses,* Mainz 6-10 April 1974, Teil 1, hg. Gerhard Funke und Joachim Kopper, Walter de Gruyter, Berlin, 86-96.

(c) "Logic, Philosophy of", in *Encyclopedia Britannica* **11**, Helen Hemingway Benton, editor, Chicago, 1974, 72-77.

(d) "Logiikka ja kielitieteen vallankumous", in *Suomalainen Tiedeakatemia, esitelmät ja pöytäkirjat 1973,* The Finnish Academy of Science and Letters, 1974. ("Logic and the Revolution in Linguistics".)

(e) "On the Proper Treatment of Quantifiers in Montague Semantics", in *Logical Theory and Semantic Analysis: Essays dedicated to Stig Kanger on his Fiftieth Birthday*, Sören Stenlund, editor, D. Reidel Publishing Co., Dordrecht, 1974, 45-60.

(f) "Practical vs. Theoretical Reason: An Ambiguous Legacy", in *Proceedings of the 1972 Bristol Colloquium on Practical Reason,* Stephan Körner, editor, Basil Blackwell, Oxford, 1974, 83-102.

The inseparability of practical and theoretical reason is illustrated by discussing the history of the assumption that we can have genuine knowledge only of what we can bring about and by discussing Aristotle's concepts of deliberation and practical syllogism.

(g) "Questions on Questions", in *Semantics and Philosophy,* Milton K. Munitz and Peter Unger, editors, New York University Press, New Delhi, 1974, 103-158.

(h) "Reply to Dorothea Frege", *Synthese* **28,** (1974), 91-96.

The thesis (put forward earlier in *Nous* **6** (1972)) is defended against Dorothea Frede that an Aristotalian science has three different kinds of starting points: (1) assumptions common to all sciences; (2) generic premisses; (3) atomic premisses. Frede's doubts about (3) are rebutted. It is maintained that for any particular science (2) carry all the existential assumptions. As a new point, it is emphasized here that Aristotle's failure to distinguish the existential and the predicative force of syllogistic premisses clearly from each other confuses the situation somewhat. Sundry objections by Mrs. Frede are also met.

(i) "Transparent Knowledge Once Again", *Philosophical Studies* **24**, (1974), 125-127.

(j)(with Lauri Carlson) "Conditionals, Generic Quantifiers and Other Applications of Subgames", in *Meaning and and Use*, A. Margalit, editor, D. Reidel Publishing, Dordrecht, 1974, 179-214.

1975

Books

(a) *The Intentions of Intentionality and Other New Models for Modalities*, D. Reidel Publishing Co., Dordrecht, 1975, x + 262 pp.

After a survey of the different constructions with epistemic verbs and of the problems concerning their interrelations (chapter 1), the following main points are argued: (i) in possible-worlds semantics, the identity of individuals in different worlds cannot be taken for granted, but presents an important conceptual problem. The class of individuals that can be cross-identified depends on the class of worlds being considered, which in the case of propositional attitudes like belief depends on the believer, i.e., on the person or community in question. This vindicates Quine's emphasis on the indeterminacy of ontology, but not his pessimistic conclusions concerning its inscrutability (chapter 2). (ii) In most contexts there are two principal methods of cross-identification, by reference to a public impersonal framework or by reference to a person's direct cognitive relations to the objects of knowledge or belief. The former may be called descriptive identification, the latter identification by acquaintance. This contrast is the same as the contrast between wh-constructions (knowing who, seeing who, etc.). The contrast underlies Russell's distinction between knowledge by description and knowledge by acquaintance (chapter 3). (iii) In cross-identification by acquaintance, world lines connecting the actual world with one's epistemic alternatives are drawn by means of causal chains (chapter 4). (iv) Carnap anticipated possible-worlds semantics technically, but he never considered models as realistic alternatives to the real world (or the actual situation) which preserve the normal meanings of non-logical words and by means of which those meanings can be studied. Once this step is taken, however, all the usual conceptual problems about intensional contexts disappear. In their stead, the problem of cross-identification becomes paramount (chapter 5). (v) This development partially answers Quine's critical questions concerning the viability of quantified model logic but partly throws them into a sharper focus, viz. by turning them into questions concerning cross-identification (chapter 6). (vi) On the basis of possible-worlds semantics for epistemic logic, a simple and illuminating analysis can be given on the relation of a question to its (conclusive) answers. This analysis relies on the questioner's state of knowledge after the reply is given, and illuminates the pragmatic nature of replies (chapter 7). (vii) The author's earlier distinction between trivial and nontrivial logical inferences (cf. item 1973 *Books* (a) – *Papers* (b) above) is brought to bear to distinguish cases where knowing (fully) the premises entails knowing their consequences and cases where this does not hold (chapter 9). (viii) The intentionality of a concept in the sense of phenomenologists should not be construed as directedness. Rather, it means that the semantics of the concept involves a variety of possible words (situations, scenarios) (chapter 10). (ix) There are analogies between the problem of representation in modern art and in recent meaning theory, e.g., cubists did not seek to capture the (appearances of) objects (references), but the noemata (Husserl) or the senses (Frege) by means of which we refer to them. Also, in both fields the choice of the method of representation has been freed from the hegemony of one preferred mode. Indeed, the very interplay of objects and their representations has become a problem (chapter 11).

(b) *Logica giochi linguistici e informazione: Temi kantiani nella filosofia della logica,* tranduzione di Marco Mondadori e Paolo Parlavecchia, in the series Biblioteca di filosofia e metodo scientifico **40**, il Saggiatore, Milano, 1975, 334pp.

(c) (edited) *Rudolf Carnap, Logical Empiricist: Materials and Perspectives*, D. Reidel Publishing Co., Dordrecht, 1975, lxviii+ 400pp.

The contributions to this volume contain criticisms of Carnap's works as well as articles developing further his ideas.

Papers

(a) "Answers to Questions" in *The Intentions of Intentionality and Other New Models for Modalities*, D. Reidel Publishing Co., Dordrecht, 1975, 137-158.

(b) "Carnap and Essler versus Inductive Generalization", *Erkenntnis* **9**, (1995), 235-244.

Essler criticizes me for assigning nonzero prior probabilities to generalizations in infinite domains because the estimate (expected value) of the frequencey of different kinds of individuals may nevertheless remain nonzero on finite evidence. This trades on the misleading connotations of terms like "estimate" and "expected value" which are happy only with a sequence of probabilistically independent events. Another criticism alleges that the nonzero probabilities of generalizations have no practical consequences. Yet De Finetti's representation theorem shows that the probabilities one associates with generalizations are determined by the bets one is willing to make on singular events on finite evidence.

(c) "Comment on Professor Bergstrom", *Theoria* **41**, (1975), 35-38.

(d) "Concept as Vision: On the Problem of Representation in Modern Art and in Modern Philosophy", in *The Intentions of Intentionality and Other New Models for Modalities*, D. Reidel Publishing Company, Dordrecht, 1975, 223-251.

The most important cubist painters conceived of their art as representational and even realistic. This surprising claim is interpreted here by means of the contrast "sinn-bedeutung" (Frege), sense-reference, or noema-object (Husserl): cubists were representing noemata, not objects. This is in keeping with their rejection of perspective and lighting. The general theoretical problems connected with the concept of sense ("Sinn", "noema") likewise have partial counter-parts in cubist theory and practice, in particular, the giving up of the idea of "logic as language" in logic and in philosophy of language, i.e., the idea of one inescapable medium of communication which can neither be viewed from the outside nor arbitrarily re-interpreted, parallels the cubist rejection of one preferred method of pictorial representation, i.e., the naturalistic and illusionistic one. Cubists were aware of the relativity of methods of pictorial representation. This technique has a partial analogue in logical model theory in the possibility of interpreting certain sets of expressions as speaking of themselves. (edited)

(e) "A Counterexample to Tarski-type Truth-definitions as Applied to Natural Languages", *Philosophia* **5**, (1975), 207-212.

Sentences of the following type are counter-examples to Davidson's use of Tarski's T-schema as applied to natural languages: the sentence "any corporal can become a general" is true if any corporal can become a general. This counter-example is not subject to easy refutations, for it really is a counter-example to the principle that the meaning of a sentence is a function of the meanings of its constituent parts.

(f) "G.H. von Wright on Logical Truth" in *The Philosophy of G.H. von Wright*, P. Schilpp, editor, Open Court, La Salle, Illinois, 1975, 25-39.

(f) "Impossible Possible Worlds Vindicated", *Journal of Philosophical Logic* **4**, (1975), 475-484.

By means of Rantala's notion of urn model, a model theory can be developed for the distinction between trivial vs. nontrivial logical truths developed earlier. It is noted that certain urn models cannot in a natural sense be told apart from classical (invariant) models. Then a logical truth is trivial if it is true in all such "almost invariant" models. This yields the same class of logical truths as the earlier formal (syntactical) characterizations.

(g) "Quine on Quantifying: A Dialogue", in *The Intentions of Intentionality and Other New Models for Modalities*, D. Reidel Publishing Co., Dordrecht, 1975, 102-136.

(i)(with Ilkka Niiniluoto) "An Axiomatic Foundation of the Logic of Inductive Generalization", in *Formal Methods in the Methodology of Empirical Sciences,* M. Przelecki et al., editors, D. Reidel Publishing Co., Dordrecht, 1975, 57-81.

(k) (with Veikko Rantala) "Systematizing Definability Theory", in *Proceedings of the Third Scandinavian Logic Symposium*, *Uppsala, April 1973*, Stig Kanger, editor, North-Holland Publishing Co., Amsterdam, 1975, 40-62.

(h) (with Unto Remes) "Ancient Geometrical Analysis and Modern Logic", in *Essays in Memory of Imre Lakatos,* R.S. Cohen et al., editors, D. Reidel Publishing Co., Dordrecht, 1975, 253-276.

(j)(with Esa Saarinen) "Semantical Games and the Bach-Peters Paradox", *Theoretical Linguistics* **2**, (1975), 1-20.

1976

Books

(a) *The Semantics of Questions and the Questions of Semantics: Case Studies in the Interrelations of Logic, Semantics and Syntax, Acta Philosophica Fennica* **28**, no. 4, (1976), 200 pp.

A direct question Q is analyzed as a request to bring out the state of knowledge specified by the desideratum of Q. Desiderata are studied by means of epistemic logic and used to define the presupposition of Q and a criterion which (conclusive) answers to Q have to satisfy. Existential and universal readings of desiderata are distinguished from each other. This approach is applied to multiple questions in English. It is argued that the set of acceptable readings cannot be explained either by the epistemic logic treatment or by any reasonable variety of the generative treatment, but can be explained by treating (subordinate) questions by means of game-theoretical semantics. Other problems concerning questions are also dealt with, including the relationship of subordinate questions to relative clauses, especially relative clauses without antecedents. Among the general theoretical suggestions of the results achieved there are specific limitations of generative-syntactical methods and the relativity of several central semantical concepts (e.g., ambiguity) to the underlying framework of semantical representation.

(b) (edited with others) *Essays on Wittgenstein in Honour of G.H. von Wright*, *Acta Philosophical Fennica* **28**, 1-3, (1976), 516 pp.

This volume is a Festschrift on the occasion of G.H. von Wright's sixtieth birthday. The contributions to it can be devoted to the following aspects of Wittgenstein'' life and philosophy: personal reminiscences, editing Wittgenstein, logic and philosophy of language, philosophy of mathematics, epistemology, philosophy of mind, philosophy of action, ethics, aesthetics and wider perspectives.

Papers

(a) "Back to Frege? A reply to Dr. Potts", in *Proceedings of the Bristol Colloquium on Philosophical Logic*, Stephan Körner, editor, Basil Blackwell, Oxford, 1976.

(b) "Gaps in the Great Chain of Being: An Exercise in the Methodology of the History of Ideas", *Proceedings and Addresses of the American Philosophical Association* **49**, (1976), 22-38.

There are no "unit ideas", in Lovejoy's sense to serve as the ultimate subject matter of the history of ideas. The implications of any idea depend on its context and on the background of its use. Their central role in the history of ideas is due instead to the fact that they define conceptual issues involved in the history of ideas. These points are illustrated and argued for by reference to the idea Lovejoy calls "the Principle of Plenitude" (the realization of all possibilities in time).

(c) "Information, Causality, and the Logic of Perception", *Ajatus* **36**, (1976), 76-94.

(d) "Language-games", in *Essays on Wittgenstein in Honour of G.H. von Wright,* Jaakko Hintikka, et al., editors, *Acta Philosophica Fennica* **28**, nos. 1-3, (1976), 105-125.

The basic idea of Wittgenstein's later philosophy of language is that the basic representative relations between language and the world are mediated by certain rule-governed human activities, language-games. This does not affect the basic idea of his picture theory, which is that combinations of symbols ("names") represent analogous combinations of entities F ("objects") in the world, because this idea is independent of the way basic semantical relations operate. The reason why Wittgenstein does not emphasize the semantical role of language-games is his general assumption that semantics is ineffable.

(e) "Partially ordered quantifiers vs. partially ordered ideas", *Dialectica* **30**, (1976), 89-99.

In response to Stenius (see the same number of *Dialectica*) the following main points are made: (1) Stenius' claim that my semantical games are inevitably games with perfect information is based on a failure to understand the concept of strategy; (2) his translation of one of my examples into a linear-quantifier notation is based on ad hoc assumptions that are not available in general; (3) my semantical games can be considered language-games in Wittgenstein's sense.

(f) "Possible Worlds Semantics as a Framework for Critical and Comparative Philosophy", in *Contemporary Aspects of Philosophy,* Gilbert Ryle, editor, Routledge & Kegan Paul Ltd., London, 1976, 57-69.

(g) "The Prospects of Convention T", *Dialectica* **30**, (1976), 61-66.

Davidson's reliance on the T-schema (see the same number of *Dialectica*) is criticized by pointing out a counter-example to its unlimited application to natural languages. There is no obvious way of getting around this counter-example, which casts serious doubts on the whole idea of recursive truth-conditions as the main tool of semantics.

(h) "Quantifiers in Logic and Quantifiers in Natural Language", in *Philosophy of Logic. Proceedings of the 1974 Bristol Colloquium*, Stephan Körner, editor, Basil Blackwell, Oxford, 1976, 208-232.

Quantifier phrases behave in natural languages rather like other denoting noun phrases. This fact is not accounted for by using the usual first-order logic as one's canonical notation. It is shown how a game-theoretical treatment explains the similarity: each quantifier phrase will denote one particular individual, but only relative to a play of a semantical game. Moreover, the values (denotations) of existential and universal quantifiers are selected by a different player.

(i)"The Question of Question Mark: A Comment on Urs Egli", *Dialectica* **30** (1976), 101-103.

Certain difficulties in Egli's treatment of questions in the same number of *Dialectica* are pointed out, especially, the difficulty of treating direct and subordinate questions in the same way.

(j)"Quine vs. Peirce?", *Dialectica* **30**, (1976), 7-8.

Supplementing Føllesdal's paper in the same number of "Dialectica", it is pointed out the way in which Quine's sense of "possible observation" differs from other uses of the same expression in being extremely narrow. The course of natural events is thought of by him as being fixed and only the movements of observers as being variable.

(k) "Who is afraid of Ludwig Wittgenstein? Reply to Professor Fogelin", in *Proceedings of the*

Bristol Colloquium on Philosophical Logic, Stephan Körner, editor, Basil Blackwell, Oxford, 1976.

(l) (with Heikki Kannisto) "Kant on 'The Great Chain of Being' or the Eventual Realization of all Possibilities: A Comparative Study", *Philosophic Exchange* **2**, (1976), 69-85.

In his early pre-critical writings, Kant accepted the principle that each possibility is eventually realized. In 1770, he rejected it, but turned back to a qualified acceptance in his mature period. The reason for the last change is that Kant limited the relevant possibilities to experiential ones. But since this limitation is due to ourselves (to the mode of functioning of human sensibility, understanding, and reason), there is no independently given range of possibilities for the principle to apply to. This explains Kan'ts ambivalence towards the principle.

(m) (with Ilkka Niiniluoto) "An Axiomatic Foundation for the Logic of Inductive

Generalization" in *Formal Methods in the Methodology of Empirical Sciences*, M.

Przelecki, et al., editors, D. Reidel Publishing Co., Dordrecht, 1976, 57-81.

(n) (with Veikko Rantala) "A New Approach to Infinitary Languages", *Annals of Mathematical Logic* **10**, (1976), 95-115.

In most of the infinitary languages currently studied, formulas are thought of as having been built by recursion from atomic ones. This assumption imposes various finitistic features on the formulas. In this paper, new languages are defined, whose formulas are not constructed step by step from atomic ones but are defined directly as certain tree-like structures. The notion of satisfaction is defined for them game-theoretically. In these languages, infinitary counterparts to Hintikka's constituents can be defined. Some fundamental properties of these generalized constituents are briefly studied.

(o) (with Unto Remes) "Ancient Geometrical Analysis and Modern Logic" in *Essays in*

Memory of Imre Lakatos, R. Cohen and M. Wartofsky, editors, D. Reidel Publishing Co.,

Dordrecht, 1976, 253-276.

1977

Books

(a) (with Unto Remes and Simo Knuuttila) *Aristotle on Modality and Determinism*, *Acta*

Philosophica Fennica 29, no. 1, (1977), 124pp.

(b) (edited with Robert Butts) *Proceedings of the Fifth International Congress of Logic,*

Methodology and Philosophy of Science, London, Ontario, Canada, 1975, 4 vols., D. Reidel Publishing Co., Dordrecht, 1977. (Includes **1** *Logic, Foundations of Mathematics and Computability Theory,* x+406pp.; **2** *Foundational Problems in the Special Sciences*, x+427pp.; **3** *Basic Problems in Methodology and Linguistics*, x+420pp.; **4** *Historical and Philosophical Dimensions of Logic, Methodology and Philosophy of Science*, x+336pp.)

Papers

(a) "Quantifiers in Natural Languages: Some Logical Problems II", *Linguistics and Philosophy* **1**, (1977), 153-172.

Two logical problems concerning natural-language quantifiers are discussed: branching quantifiers and the behavior "any." These exemplify two major explanatory strategies made possible by a game-theoretical treatement, viz. the possibility of informationally independent moves and the use of ordering principles governing the applications of different game rules. As to branching quantifiers, more evidence of their presence in natural language is uncovered. As to "any", the ordering principles governing it are registered and a condition, the *any*-thesis, for its acceptability in a given context is formulated. Given certain further assumptions, it is shown that, if the *any*-thesis is correct, the class of acceptable sentences of English is not recursively enumerable. This would show that no generative methods can fully account for the acceptability of English sentences.

(b) "The Ross Paradox as Evidence for Reality of Semantical Games", *The Monist* **60**, (1977), 370-379.

The Ross paradox is exemplifed by the intuitive invalidity of "John ought to P; therefore Jouhn ought to P or Q" in spite of the validity of "OP implies O(P $\vee$ Q)" in deontic logic. The explanation here offered turns on the fact that in the semantical game on "John ought to P or Q" my move in choosing R or Q is easily confused with John's decision to P or to Q. This explanation supports the reality of semantical games, for otherwise their moves could not be confused with real life decisions. Supplementary evidence is adduced for this resolution of the paradox, which also applies to the paradox of free choice permission.

(c) (with Lauri Carlson) "Pronouns of Laziness in Game-theoretical Semantics", *Theoretical Linguistics* **4**, (1977), 1-29.

1978

Papers

(a) "Answers to Questions", in *Questions,* Henry Hiż, editor, D. Reidel Publishing Co., Dordrecht, 1978, 279-300.

(b) "Aristotle's Incontinent Logician", *Ajatus* **37**, (1978), 48-65.

Aristotle's actual syllogistic theory is based on a distinction between perfect syllogisms, which are self-explanatory, and imperfect ones, which are not self-explanatory and which therefore are to be reduced to perfect ones. It is argued that Aristotle's ideas about the psychology of reasoning nevertheless committed him to holding that all syllogisms are automatic and self-explanatory. All thinking involves *ekthesis*-like instantiations, and such instantiations will automatically implement all syllogisms. Yet Aristotle avoided the use of *ekthesis* in his syllogistic theory as much as possible. This theory is hence a shaky compromise between different theoretical ideas. Aristotle could never explain fully how the knowledge of syllogistic premises does not automatically mean knowledge of the conclusion. This problem is the theoretical counterpart to the corresponding problem of accommodating failures to draw a practical inference, which is Aristotle's problem of incontinence (*Akrasia)*.

(c) "Degrees and Dimensions of Intentionality", *Versus: Quaderni di studi semiotici* **19**, (1978), 73-76.

(d) "A Discourse on Descartes' Method", in *Descartes: Critical and Interpretative Essays,* Michael Hooker, editor, The Johns Hopkins University Press, Baltimore, 1978, 74-88.

Descartes' philosophical and scientific method was a variant of the method of analysis originating from ancient Greek geometry, and a generalization of his analytic (algebraic) method in geometry. The problems connected with it are the same as the problems of understanding the method of analysis in general. The three different types of analysis Buchdahl distinguishes are combined in Descartes' idea. A difference between Descartes and Newton was that for the former analysis was essentially conceptual analysis whereas for Newton analysis "consists in making experiments and observations."

(e) (with Merrill Provence (Hintikka)) "Wittgenstein on Privacy and Publicity", in *Wittgenstein and His Impact on Contemporary Thought,* Elisabeth Leinfellner et al., editors, Hölder-Pichler-Tempsky, Wien 1978, 353-362.

1979

Books

(a) (edited with Ilkka Niiniluoto and Esa Saarinen) *Essays in Mathematical and Philosophical Logic. Proceedings of the 4th Scandinavian Logic Symposium and of the 1st Soviet-Finnish Logic Conference, Jyäskylä, Finland, June 29-July 6, 1976*, D. Reidel Publishing Co., Dordrecht, 1979, 462 pp.

Papers

(a) "Frege's Hidden Semantics" *Revue Internationale de Philosophie* **33**, (1979), 716-722.

Frege's main achievement as a semanticist is not his theory of sense and reference, but the creation of semantics for first-order logic. His awareness of the problems of intensional contexts was but a corollary to his insights into the power (and limitations) of the extensional semantics which has since become almost universally adopted. Frege was prevented from formulating this semantics explicitly by his belief in "logic as language" (van Heijenoort) or, more generally, in "language as the universal medium." Only now are the main limitations of Frege's first-order semantics under criticism.

(b) "'Is', Semantical Games, and Semantical Relativity", *Journal of Philosophical Logic* **8**, (1979), 433-468.

Frege and Russell (followed by most subsequent logicians, philosophers, and linguists) claimed that "is" is ambiguous between identity, existence, predication, and general implication. It is shown that no such ambiguity can be present in Hintikka'' game-theoretical semantics. This shows that central semantical notions (e.g., ambiguity) can be relative to the underlying semantical (logical) framework. The resulting ""semantical relativity" has several important implications for the methodology of linguistics, telling, e.g., against all reliance on semantical intuitions or on "the language of thought." It is also anachronistic to project the Frege-Russell ambiguity to most pre-Fregean logicians and philosophers.

(c) "Quantifiers in Natural Languages: Some Logical Problems", in *Game-Theoretical Semantics,* Esa Saarinen, editor, D. Reidel Publishing Co., Dordrecht, 1979, 81-117.

(d) "Quantifiers in Natural Language: Some Logical Problems I", in *Essays in Mathematical and Philosophical Logic,* Jaakko Hintikka, Ilkka Niiniluoto and Esa Saarinen, editors, D. Reidel Publishing Co., Dordrecht, 1979, 295-314.

(e) "Rejoinder to Peacocke", in *Game-theoretical Semantics*, Esa Saarinen, editor, D. Reidel Publishing Co., Dordrecht, 1979, 135-151.

(f) "Virginia Woolf and our Knowledge of the External World", *Journal of Aesthetics and Art Criticism* **38**, (1979), 5-14.

As a case study of the interrelations of literature and philosophy in the Bloomsbury group, a comparison is made between Virginia Woolf's fictional technique and Russell's central construction in his book *Our Knowledge of the*

External World. Virginia Woolf lets her readers construct her fictional universe. Especially her main characters, out of the impressions of other characters or even from those of merely potential observers. Russell constructs one common world out of the perspectives of individual observers, including merely potential ones. The construction is in neither case a reduction but calculated to enhance the reality of our common everyday world.

(g) (with Lauri Carlson) "Conditionals, Generic Quantifiers, and other Applications of Subgames", in *Meaning and Use*, Avishai Margalit, editor, D. Reidel Publishing Co., Dordrecht 1979, 179-214.

The natural treatment of a conditional "if *X*, then *Y*", in game-theoretical semantics is to divide the game on it into two subgames: First, a game with roles reversed is played on *X*. Only if the verifier wins this subgame is a game played on *Y* (with normal roles) in which the verifiers (nature's) strategy in the game on *X* is remembered. If the syntactical order of *X* and *Y* is reversed, a different arrangement of subgames is more natural. In this way, several problems in the semantics of conditionals, including the behavior of anaphoric pronouns in them, can be understood.

(h) (with Esa Saarinen) "Information-seeking Dialogues: Some of their Logical Properties", *Studia Logica* **38**, (1979), 355-363.

1980

Books

Logical-Epistemological Studies, V.N. Sadovski and V.A. Smirnova*,* editors, translated into Russian by V.I. Bryushinkina, et al., in the series Logic and Methodology of Science , V.M. Leontyev, editor, Publishing House 'Progress', Moscow, 1980, 448 pp.

Papers

(a) "Aristotelian Induction", *Revue Internationale de Philosophie* **34**, (1980), 422-439.

(b) "C.S. Peirce's 'First Real Discovery' and its Contemporary Relevance", *The Monist* **63**, (1980), 304-315.

C.S. Peirce made the same distinction between trivial and nontrivial logical truths as was made in 1973(a) calling the two "corollarial" and "theorematic" and attaching a great significance to the distinction. Peirce's distinction was a generalization from elementary geometry where some arguments do not need auxiliary constructions whereas others do. His insight was that this need of "auxiliary objects" is not obviated by a formalization of the geometrical arguments.

(c) "Degrees and Dimensions of Intentionality", in *Language, Logic, and Philosophy. Proceedings of the Fourth International Wittgenstein Symposium, 28th August to 2nd September, 1979*, Rudolf Haller and Wolfgang Grassl, editors, Holder-Pichler-Tempsky, Wien, 1980, 283-296.

In the title essay of *The Intentions of Intentionality* it was argued that the intentionality of a concept means that its semantics involves a multiplicity of possible scenarios ("worlds"). This idea is developed further by suggesting that a concept is the more intentional the more different the alternative worlds are from the actual one that are used in its semantics. Since these differences are themselves unlike each other, we obtain a variety of different dimensions of intentionality. The most important is the one in which the alternative worlds need not even be logically possible, only epistemically possible. It turns out that this dimension is found to be especially significant both in our conceptual practice (it distinguishes, e.g., subjective conceptions of probability from objective ones" and in the formal criteria of intentionality that philosophers have proposed.

(d) "In What Sense can Values be Absolute?", in *Proceedings of the Eighth International Conference on the Unity of Sciences,* New York, 1980, 35-39.

(e) "On the *Any*-thesis and the Methodology of Linguistics", *Linguistics and Philosophy* **4**, (1980), 101-122.

New evidence is adduced for the author's "any"-thesis ("any" is acceptable if and only if "every" is acceptable in its place and yields a nonequivalent string). Chomsky's attempted explanation of the same data (in *Rules and Representations*) is refuted. "Any"-thesis implies (jointly with other assumptions) that the set of acceptable English sentence is not generable in any generative grammar. Chomsky's claim that this entails no major methodological consequences is criticized. E.g., Chomsky's present position differs radically from his early one; "any"-thesis implies a greater violation of the autonomy of syntax than he acknowledges.

(f) "On Sense, Reference, and the Objects of Knowledge", *Epistemologia* **3**, (1980), 143-164.

(g) "Parmenides' Cogito Argument", *Ancient Philosophy* **1**, (1980), 5-16.

(h) "Philosophy in Finland since 1945", in *Handbook of World Philosophy,* John R. Burr, editor, Greenwood Press, Westport, Conn., 1980, 15-32.

(i)"Standard vs. Nonstandard Logic: Higher-order, Modal, and First Order Logics", in *Modern Logic,* Evandro Agazzi, editor, D. Reidel Publishing Co., Dordrecht, 1980, 283-296.

In the usual Kripke-semantics for modal logic, the set of alternatives to the actual world can be any set of worlds. This does not capture the notion of *logical* necessity, for logical necessity obviously means truth in all logically possible alternatives, not just in some arbitrary set of alternatives. Hence Kripke models for alethic modal logic must be modified and a further requirement imposed on them. This yields a new kind of models for modal logics, which are related to the old ones in the same way standard models of higher-order logics in Henkin's sense are related to a kind of (weak) nonstandard models. The "standard" modal logic that they define is not well defined, however, before assumptions are made as to what individuals may exist in the alternative worlds, and on suitable further assumptions it is equivalent in power to standard second order logic and hence unaxiomatizable.

The contrast between standard and nonstandard models can be extended to first-order logic.

(j)“Theories of Truth and Learnable Languages”, in *Philosophy and Grammar: Papers on the Occasion of the Quincentennial of Uppsala University*, Stig Kanger and Sven Ohman, editors, D. Reidel Publishing Co., Dordrecht, 1980, 37-57.

Compositionality (Frege principle) says that the meaning of a xomplex expression depends functionally on the meanings of its parts. It is shown to amount to semantical context-independence, which fails in e.g., English, as exemplified by branching quantifiers, backwards-looking operators, “any”, etc., Davidson notwithstanding, compositionality is also unnecessary for learnability. Moreover, T-schema (Tarski, Davidson) fails (witness “ ‘anyone can become a millionaire’ is true if anybody can become a millionaire”).

(k) (with Merrill B. Hintikka) “Different Language-games in Wittgenstein”, in *Language, Logic, and Philosophy. Proceedings of the Fourth International Wittgenstein Symposium,* Rudolf Haller and Wolfgang Grassl, editors, Hölder-Pichler-Tempsky, Wien, 1980, 417-422.

1981

Books

(a) (edited with David Gruender and Evandro Agazzi) *Theory Change, Ancient Axiomatics, and Galileo’s Methodology: Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science. Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science***,** Synthese Library, 2 vols., D. Reidel Publishing Co., Dordrecht, 1981. (Includes Synthese Library **145,** xvi+354pp. and Synthese Library **146**, xiv+324pp.)

Papers

(a) “Aristotelian Axiomatics and Geometrical Axiomatics”, in *Theory Change, Ancient Axiomatics, and Galileo’s Methodology; Probabilistic Thinking, Thermodynamics, and the Interaction of the History and Philosophy of Science. Proceedings of the 1978 Pisa Conference on the History and Philosophy of Science* **I**, Synthese Library **145**, Jaakko Hintikka, David Gruender and Evandro Agazzi, editors, D. Reidel Publishing Co., Dordrecht, 1981, 133-144.

(b) “Intuitions and Philosophical Method”, *Revue Internationale de Philosophie* **35**, (1981), 74-90.

(c) “Kant on Existence, Predication, and the Ontological Argument”, *Dialectica* **35**, (1981), 127-146.

The ontological argument fails because of an operator order switch between (1) “necessarily there is an (existentially) perfect being” and (2) “there is a being which necessarily is (existentially) perfect”. Here (1) is trivially true logically but (2) is problematic. Since Kant’s criticisms were directed at the notion of existence, not at the step from (1) to (2), they are misplaced. They are also wrong, because existence can be a predicate. Moreover, Kant did not anticipate Frege's claim that "is" ("isi") is ambiguous between existence, predication, identity, nd class-inclusion. To restore the ontologicl argument, an extra premise is needed to the effect (roughly) that it is known who the existentially perfect being is. The question is raised whether Kant could have meant the failure of this extra premise by his thesis that existence is not a "real" predicate.

(d) "The Logic of Information-seeking Dialogues: A Model", in *Konzepte der Dialektik,* Wilhelm Essler und Werner Becker, editors, Vittorio Klostermann, Frankfurt A.M., 212-231.

(e) "On Common Factors of Dialectics", in *Konzepte der Dialektik*, Wilhelm Esler and Werner Becker, editors Vittorio Klostermann, Frankfurt A.M., 1981, 109-110.

(f) "On Denoting What?", *Synthese* **46**, (1981), 167-183.

(g) "On the Logic of an Interrogative Model of Scientific Inquiry", *Synthese* **47**, (1981), 60-84.

(h) "Phenomenology vs. Possible-worlds Semantics: Apparent and Real Differences", *Revue Internationale de Philosophie* **35**, (1981), 113-119.

(i)Russell, Kant, and Coffa", *Synthese* **46**, (1981), 265-270.

(j)"Semantical Games and Transcendental Arguments", in *Theory of Argumentation,* E.M. Barth and J. Martens, editors, John Benjamins, Amsterdam, 1981.

(k) "Semantics: A Revolt Against Frege", in *Contemporary Philosophy: A New Survey,* **1**, in the series *Philosophy of Language/Philosophical Logic*, G. Floistad and G.H. von Wright, editors, Martinus Nijhoff, The Hague, 1981, 57-82.

(l)"Theories of Truth and Learnable Languages" in *Philosophy and Grammar: Papers on the Occasion of the Quincentennial of Uppsala University*, Stig Kanger and Sven Ohman, editors, D. Reidel Publishing Co., Dordrecht, 1981, 37-57.

(m) "Tieteen prosessiluonne ja sen seuraukset tiedesuunnittelulle", *Tieteen tila, KTTS:n monistesarja* **9**, Foundation for Research in Higher Education and Science Policy, Helsinki, 1981, 58-80. ("The Process Character of Science and its Consequences for Science Policy".)

(n) "What is an Answer?, and Other Questions in the Theory of Questions and Answers", in *Philosophy As Science and Philosophy of Science,* Edgar Morscher, et al., editors, Comes Verlag, Bad Reichenhall, 1981, 261-277.

(o) "Wittgenstein's Semantical Kantianism", in *Ethics: Foundations, Problems and Applications*. *Proceedings of the Fifth International Wittgenstein Symposium,* Edgar Morscher and R. Stranzinger, editors, Hölder-Pichler-Tempsky, Wien, 1981, 375-390.

(p) (with Merrill B. Hintikka) "Wittgenstein: Some Perspectives on the Development of his Thought", in *Essays in Philosophical Analysis*: *Dedicated to Erik Stenius on the Occasion of his 70th birthday*, Ingmar Pörn, editor, *Acta Philosophica Fennica* **32**, (1981), 79-95.

The concept of ostensive definition plays an important role in Wittgenstein's early middle period. It is a descendant of the idea of showing in "Tractatus." But the allegedly direct naming relations which ostensive definitions can establish were in Wittgenstein's later philosophy constituted by complex language-games, which can be learned only by training, not by ostension. This change led Wittgenstein to deemphasize rules and criteria in his mature philosophy.

(q) (with Merrill B. Hintikka) "Wittgenstein and the 'Universal Language' of Painting", in *Ethics, Foundations, Problems and Applications. Proceedings of the Fifth International Wittgenstein Symposium,* E. Morscher and R. Stranzinger, editors, Hölder-Pichler-Tempsky, Wien, 1981, 492-497.

1982

Books

(a) *Kieli ja mieli: Katsauksia kielifilosofiaan ja merkityksen teoriaan,* Otava, Helsinki 1982. 255pp. (*Language and Meaning. Surveys of the Philosophy of Language and the Theory of Meaning*.)

Papers

(a) "A Dialogical Model of Teaching", *Synthese* **51**, no.1, (1982), 39-59.

A simple model of teacher-student interaction is set up using a game theoretical framework. This enables us to study instructional strategies, i.e., dependencies of the teacher's and the student's several "moves" on each other. Since some of these moves are question-answer pairs, the model relied on my theory of questions. We can e.g. distinguish different purposes questions can serve in instruction. The relation of instructional strategies to problem solving is briefly discussed.

(b) "Game-Theoretical Semantics: Insights and Prospects", *Notre Dame Journal of Formal Logic* **23**, (1982), 219-241.

The basic ideas of game-theoretical semantics are implicit in logicians' and mathematicians' folklore but used only sporadically (e.g., game quantifiers, back-and-forth methods. Partly ordered quantifiers). The general suggestions of this approach for natural languages are emphasized: the univocity of "is", the failure of compositionality, a reconstruction of Aristotelian categories, limitations of generative grammars, unity of sentence and discourse semantics, an new treatment of senses and other temporal notions, etc.

(c) "Is Alethic Modal Logic Possible?", in *Intensional Logic: Theory and Applications,* Ilkka Niiniluoto and Esa Saarinen, editors, *Acta Philosophica Fennica* **35**, (1982), 89-105.

The correct semantics for logical modalities is not Kripke's. It is insufficient for the logical truth of P that P be true in each alternative, unless we (unlike Kripke) require every set of alternatives to contain all the relevant logically possible structures. Even then we have a problem about the domains of individuals of the alternatives. If they are not restricted, paradoxes threaten. If they are restricted to actual individuals, we obtain an unaxiomatizable logic.

(d) "Kant's Theory of Mathematics Revisited", in *Essays on Kant's Critique of Pure Reason,* J.N. Mohanty and Robert W. Shehan, editors, University of Oklahoma Press, Norman, Oklahoma, 1982, 201-215.

The author's interpretation of Kant's theory of mathematics is defended by pointing out the precise meaning of Kant's statements in the light of the mathematical practice of his day (Euclid, Descartes). Among the most salient points there are the following: the "datum" of Kant's theory is the need of auxiliary constructions (instantiations), not appeal to intuitions; the analyticity of mathematical inferences in B 14 refers only to the "apodeixis" part of Euclidean arguments, and hence does not support attempts to trace the syntheticity of mathematics apud Kant back to the axioms.

(e) "Questions with Outside Quantifiers", in *Papers from the Parasession on Nondeclaratives,* Robinson Schneider, Kevin Tute, and Robert Chametzky, editors, Chicago, 1982, 83-92.

(f) "Semantical Games and Transcendental Arguments", in *Argumentation: Approaches to Theory Formation,* E.M. Barth and J.L. Martens, editors, John Benjamins, Amsterdam, 1982, 77-91.

(g) "Tag-questions and Grammatical Acceptability", *Nordic Journal of Linguistics* **5**, (1982), 129-132.

(h) "Temporal Discourse and Semantical Games", *Linguistics and Philosophy* **5**, (1982), 3-22.

A game-theoretical semantics for temporaal discourse offers several advantages, largely because of its semantical context sensitivity: the choices of time-moments correlated with different words can be co-ordinated so as to instantiate a general principle; important regularities (e.g., the any-thesis) can be extended to temporal contexts; Reichenbach's notion of reference-time is avoided, especially in the contrast between simple past and past perfect.

(i)"Transcendental Arguments Revived", in *Philosophers on Their Own Work – Philosophers critiques d'eux-mêmes* **9**, édité par Andre Mercier and Maja Svilar, Peter Lang, Bern, 1983, 116-133.

(j) (with Merrill B. Hintikka) "Sherlock Holmes Confronts Modern Logic: Toward a Theory of Information-seeking Through Questioning", in *Argumentation: Approaches to Theory Formation,* E.M. Barth and J.L. Martens, editors, John Benjamins, Amsterdam, 1982, 55-76.

In ordinary discourse (e.g., detective novels) logical deductions (inferences) are assumed to yield new factual information. Such "deductions" cannot be logical inferences in philosophers' technical sense, for in this sense all inferences are tautological. It is proposed that they be construed as a series of questions addressed to some source of information, which can be the inquirer's tacit background knowledge, interspersed by logical inferences in the narrow technical sense. Such "deductions" cannot be construed as enthymemes, for the questions needed to elicit this knowledge may depend on earlier inferences and earlier questions. Such question-answer sequences can be construed as "games against nature."

(k) (with Merrill B. Hintikka) "Towards a General Theory of Individuation and Identification", in *Language and Ontology. Proceedings of the Sixth International Wittgenstein Symposium*, Werner Leinfellner, Eric Kraemer and Jeffrey Schank, editors, Hölder-Pichler-Tempsky, Wien, 1982, 137-150.

The crucial conceptual problem in possible-worlds semantics is cross-identification. Since in typical cases, the different "possible worlds" share a part, cross-identification is possible if re-identification is possible, for then we can try to compare individuals in different worlds by tracing them in space-time to the common part. But how do we re-identify, say, physical objects? It is shown that if the basic data of the re-identification of propertyless mass points include their instantaneous velocities at different times, their world lines are obtained as solutions to certain systems of differential equations. Physical objects proper are then defined by their surfaces, which are smooth, stable sets of singularities of solutions of those equations. Conceptually, re-identification problem is thus a problem in the stability theory of differential equations. This has philosophical implications, e.g., because it suggests that the concepts of space and time are more basic that the concept of discreet object, for they have to be relied on in the conceptual constitution of physical objects.

(l)(with Jack Kulas) "Russell Vindicated: Towards a General Theory of Definite Descriptions", *Journal of Semantics* **1**, (1982), 387-397.

1983

Books

(a) (with Jack Kulas) *The Game of Language: Studies in Game-Theoretical Semantics and Its Applications,* Synthese Language Library **22**, D. Reidel Publishing Co., Dordrecht, 1983, xii + 344 pp.. (Second, corrected ed. 1985.)

The first chapter presents a survey of the basic ideas and results of game-theoretical semantics (GTS). The following theses are argued in the other chapters: (i) GTS is the correct realization of Kant's theory of mathematical (for us, logical) reasoning when Kant's mistaken reliance on perception as the only source of our knowledge of particulars is eliminated (chapter 2). (ii) The true logic of GTS is given by Gödel-type functional interpretations. By their means, certain tricky problems, including the treatment of Geach's "donkey sentences", can be solved (chapter 3). (iii) The semantical behavior of negation, "any", temporal notions, and definite descriptions in natural languages can be fruitfully studied by means of GTS (chapters 4-6). (iv) In GTS, words like "is" are not, and cannot be, assumed to be ambiguous in the Frege-Russell sense between the "is" of identity, predication, existence, and general implication. This ambiguity claim is mistaken, and differences in use between different "is" can be explained contextually. Even if GTS is not accepted as the sole account of the semantics of natural languages, it shows that many of the basic semantical concepts (e.g., ambiguity) are relative to a semantical theory. Since the Frege-Russell ambiguity idea is built into the usual first-order logic, this logic is not the only nor the best framework of semantical representation for natural languages (chapter 7). (v) By spelling out the range of player's choices in semantical games on quantified sentences in natural languages, we obtain a close approximation to Aristotle's theory of categories, which thus was not a theory of logical types but a theory of largest classes of particular entities (values ranges of quantifiers). This approximation runs into problems closely related to Aristotle's problems (chapter 8). (vi) Further evidence is presented for the thesis that the class of acceptable sentences of English is not recursively enumerable. Chomsky's alternative proposal is criticized and the methodological implications of the result examined (chapter 9). (vii) The principle of compositionality presupposes a kind of semantical context-independence. It is not assumed in GTS, which therefore can handle linguistic phenomena due to failures of the principle. Several such apparent counter-examples to the principle are examined, including counter-examples to Tarski's T-schema. (It is not true that "anybody can become a millionaire" is true if anybody can become a millionaire.) Such counter-examples can in principle be explained away, but only at too high a cost in psycholinguistic and theoretical simplicity.

Papers

(a) "*Any* Problems — *No* Problems" in Jaakko Hintika, with Jack Kulas, *The Game of Language: Studies in Game-Theoretical Semantics and Its Applications,* Synthese Language Library **22**, D. Reidel Publishing Co., Dordrecht, 1983, 77-112.

(b) "New Foundations for a Theory of Questions and Answers", in *Questions and Answers,* F. Kiefer and Hans Karlgren, editors, KVAL, Stockholm, 1983, 159-190.

(c) "Paras teoria", in *Huippuluokan tutkielmia,* Lilli Alanen et al., editors, Reports from the Department of Philosophy, University of Helsinki, no. 2, (1983), 8-14. ("The Best Theory".)

(d) "Semantical Games, the Alleged Ambiguity of 'is', and Aristotelian Categories", *Synthese* **54**, (1983), 443-467.

(e) "Semantical Games, Subgames, and Functional Interpretations" in Jaakko Hintikka, with Jack Kulas, *The Game of Language: Studies in Game-Theoretical Semantics and Its Applications,* Synthese Language Library **22**, D. Reidel Publishing Co., Dordrecht, 1983, (Second, corrected ed. 1985.), 47-76.

(f) "Sherlock Holmes Formalized", in *The Sign of Three: Dupin, Holmes, Peirce,* Umberto Eco and Thomas Sebeok, editors, Indiana University Press, Bloomington, Indiana, 1983, 170-178.

(g) "Situations, Possible Worlds, and Attitudes", *Synthese* **54**, (1983), 154-162.

(h) "Super Models", in *Vexing Questions: An Urnful of Essays in Honour of Veikko Rantala,* Ilkka Patoluoto et al., editors, Reports from the Department of Philosophy, University of Helsinki, no. 3, (1983), 12-18.

(i)"Transsendentaalitiedon paradoksi", *Ajatus* **40**, (1983), 20-48. ("The Paradox of Transcendental Knowledge".)

(j)(with Merrill B. Hintikka) "The Development of Ludwig Wittgenstein's Philosophy: The Hidden Unity", in *Epistemology and Philosophy of Science. Proceedings of the Seventh International Wittgenstein Symposium*, Paul Weingartner and Hans Czermak, editors, Hölder-Pichler-Tempsky, Wien, 1983, 425-437.

(k) (with Merrill B. Hintikka) "How Can Language be Sexist?", in *Discovering Reality: Feminist Perspectives on Epistemology, Metaphysics, Methodology, and Philosophy of Science,* Sandra Harding and Merrill B. Hintikka, editors, D. Reidel Publishing Co., Dordrecht, 1983, 139-148.

(l)(with Merrill B. Hintikka) "Some Remarks on (Wittgensteinian) Logical Form", *Synthese* **56**, (1983), 155-170.

(m) (with Merrill B. Hintikka) "Wittgensteinin Tractatus-teoksen salaisuus", in *Suomalainen, Tiedeakatemia – Academia Scientiarum Fennica — Vuosikirja – Yearbook 1982*, Lauri A. Vuorela, editor, Helsinki, 1983, 121-133. ("The Enigma of Wittgenstein's *Tractatus*")

1984

Books

(a) (edited with Lucia Vaina) *Cognitive Constraints on Communication*, D. Reidel Publishing Co., Dordrecht, 1984, xiv+428pp.

Papers

(a) "Are There Nonexistent Objects? Why Not? But Where are They?", *Synthese* **60,** (1984), 451-458.

"Das Paradox transzendentaler Erkenntnis", in *Bedingungen der Möglichkeit: 'Transcendental Arguments' und Transzendentales Denken,* Hrsg. Eva Schaper und W. Vossenkuhl, Klett-Cotta, Stuttgart, 1984, 123-149.

(c) "Hundred Years Later: The Rise and Fall of Frege's Influence in Language Theory", *Synthese* **59**, (1984), 27-49.

Frege established a paradigm which has dominated philosophical language theory for the last 100 years. Some of his main ingredients are identified and criticized: (1) the idea of quantifiers as simply ranging over a set of entities, which leads to the problems of (i) atomism and (ii) cross-identification; (2) distinction between allegedly different meanings of "is" (identity, prediction, existence, and class-inclusion); (3) compositionality (semantical context-independence).

(e) "The Logic of Science as a Model-Oriented Logic", in *Philosophy of Science Association 1984* **1**, Peter Asquith and Philip Kitcher, editors, Philosophy of Science Association, East Lansing, Michigan, 177-185.

(f) "Luovuus ja ihmiskäsitykset", *Ajatus* **41**, (1984), 83-88. ("Creativity and Conceptions of Man".)

(g) "Kant's Transcendental Method and His Theory of Mathematics", *Topoi* **3**, (1984), 99-108.

Following his transcendental method, Kant tried to explain mathematical knowledge as reflecting the way we humans come to know particulars. This way Kant mistakenly identified with sense-perception, concluding that mathematical knowledge reflects the forms of our sense-perception. Here is the reason why things-considered-in-themselves are transcendent causes of perceptions. A true Kantian should identify it with the "language-games" of seeking and finding, which would lead him to my game-theoretical semantics.

(h) "Questioning as a Philosophical Method", in *Principles of Philosophical Reasoning,* James H. Fetzer, editor, Rowman and Allanheld, Totowa, N.J., 1984, 25-43.

(i)"Rules, Utilities, and Strategies in Dialogical Games", in *Cognitive Constraints on Communication,* Lucia Vaina and Jaakko Hintikka, editors, D. Reidel Publishing Co., Dordrecht, 1984, 277-294.

(j)(with Charles Harvey) "Review Article on David W. Smith and Ronald McIntyre, *Husserl and Intentionality*", *Husserl Studies* **2**, (1984), 201-212.

(k) (with Lucia Vaina) "Introduction", in *Cognitive Constraints on Communication,* Lucia Vaina and Jaakko Hintikka, editors, D. Reidel Publishing Co., Dordrecht, 1984, vii-xvii.

1985

Books

(a) (with Jack Kulas) *Anaphora and Definite Descriptions: Two Applications of Game-Theoretical Semantics,* Synthese Language Library **26**, D. Reidel Publishing Co., Dordrecht, 1985, xiv + 250 pp.

Definite descriptions (the-phrases) are treated in game-theoretical semantics as involving two interdependent choices of individuals from a certain choice set *I* which essentially is the set of individuals so far chosen by the players of a semantical game or otherwise made available to them. Anaphoric pronouns are treated similarly. Hence they are not like variables of quantification, "bound" to their grammatical antecedents, but independently evaluated choice terms rather like Hilbert's *epsilon*-terms except that the choice is limited to *I*. Restrictions on co-reference for such pronouns are consequences of the ordering principles which in general govern semantical games, including the introduction of individuals into *I*. (If a non-anaphoric expression has not been dealt with so as to introduce its value into *I* before a rule is applied to an anaphoric pronoun, this expression cannot be "co-referential" with the pronoun.)

(m) (edited with Fernand Vandamme) *Logic of Discovery and Logic of Discourse,* Plenum Press, NY, 1985, xv+271pp.

Papers

(a) "Legal Reasoning and Legal Systems", in *Man, Law and Modern Forms of Life,* E. Bulygin et al., editors, D. Reidel Publishing Co., 1985, 209-220.

(b) "Philosophical Logic", (in Hebrew) *Modern Trends in Philosophy* **2**, A. Kasher and Shalom Lappin, editors, Yachdav United Publishers, Tel Aviv, 1985, 71-93.

(c) "A Spectrum of Logics of Questioning", *Philosophica* **35**, (1985), 135-150.

In an interrogative game, the inquirer tries to proce C (or not –C) from a theory 1 plus nature's answers to his/her questions. The character of the game depends on structural restrictions on available answers, with no restrictions. The strategy selection is virtually identical with the purely deductive case. Traditionally, nature's only answers are assumed to be (possibly negated) atomic propositions. In the logic of experimental inquiry, AE answers nevertheless also occur, changing radically the situation.

(d) "True and False Logics of Scientific Discovery", *Communication and Cognition* **18**, (1985), 3-14.

Several conceptions of the logic of scientific discovery are criticized, including its impossibility, its interpretation as the logical syntax of the language of science and the structuralist view. The logic of science is construed as a logic of questioning. Different varieties of this idea (Kant, Laudan) are compatible because questions play two different rules in interrogative inquiry: The inquirer is trying to answer a "big" initial question by pulling a number of "small" questions to nature and using her answers as additional premises.

(e) (with Merrill B. Hintikka) "Ludwig Looks at the Necker Cube: The Problem of 'Seeing as' as a Clue to Wittgenstein's Philosophy", *Acta Philosophica Fennica* **38**, (1985), 36-48.

Wittgenstein's different comments on ambiguous figures help to confirm the interpretation offered in *Investigating Wittgenstein* (1986(a)). In the *Tractatus* seeing the same configuration of physical objects in two different ways was supposed to show that the objects we have to assume are phenomenological, not physical. Conversely, "seeing as" became a problem for Wittgenstein when he gave up phenomenological languages. He had to explain, not only the possibility of seeing a figure in different ways, but also the spontaneity (non-interpretation character) of "seeing as." That Wittgenstein comments conform to our interpretation is seen also from his use of the term "aspect."

(f) (with Merrill B. Hintikka) "Wittgenstein über private Erfahrung", in *Sprachspiel und Methode: Zum Stand der Wittgenstein-Diskussion,* Dieter Birnbacher and Armin Burkhardt, editors, Walter de Gruyter, Berlin, 1985, 1-26.

(g) (with Merrill B. Hintikka) "Wittgenstein's 'annus mirabilis':1929", in *The Tasks of Contemporary Philosophy. Proceedings of the Tenth International Wittgenstein Symposium*, Hölder-Pichler-Tempsky, Vienna, 1985, 437-447.

Wittgenstein's notebooks show that on October 22, 1929, he rejected phenomenological languages in favor of physicalistic ones. This forced him to face new problems, including (i) how to speak of internal (phenomenological) objects and events in a physicalistic language, and (ii) how language-world links are constituted now that simple objects are no longer presented to us in direct experience. Wittgenstein's eventual answer to (i) is the ill-named 'private language argument', and to (ii), his concept of language-game.

(h) (with Jack Kulas) "Different Uses of the Definite Article", *Communication and Cognition* **18**, (1985), 69-80.

Anaphoric definite descriptions are like Russellian ones except that the quantifiers they involve range over values available in a semantical game at the time. Russellian, Generic and Platonic uses of Definite descriptions are pragmatic variants of the anaphoric one. The generic use arises when the uniqueness presupposed by definite descriptions can only be satistifed by assuming a "museum scenario" where one representative of each homogeneous kind is being considered. Hence the generic sense expresses what is species-characteristic, not what is lawlike.

(i)(with Simo Knuuttila) "Introduction", in *The Logic of Being: Historical Studies*, Synthese Historical Library **28,** Simo Knuuttila and Jaakko Hintikka, editors, D. Reidel Publishing Co., Dordrecht, 1985, ix-xvi.

1986

Books

(a) (with Merrill B. Hintikka) *Investigating Wittgenstein,* Basil Blackwell, Oxford, 1986, xx + 326 pp.

Wittgenstein's early philosophy can be seen as a further development of Russell's theory of acquaintance. In his *Theory of Knowledge* (1913, published only in 1984), Russell tried to account for logic in terms of his theory by postulating logical forms as objects of acquaintance. Wittgenstein modified this by rejecting logical forms as independent objects of acquaintance. All logical forms can be built out of the logical forms of basic objects. These objects are given in direct experience, and the language to be used of them is a phenomenological one. The meanings of simple names are such phenomenological objects; they cannot be expressed in language, but have to be presented in immediate experience, i.e., "shown." Wittgenstein's development out of his early position began in October 1929 when he rejected phenomenological languages in favor of everyday physicalistic languages as philosophically and logically basic ones. That implied that meanings cannot any longer be taught by ostensive confrontations with direct experience, but have to be mediated somehow. Wittgenstein experimented with rules and criteria as such mediators, but rejected them in favor of language-games which are conceptually prior to their rules. The ultimate basis for rejecting private phenomenological languages is the need of language-games as mediators of meaning, for games qua games cannot be private. Thus Wittgenstein is not denying the reality, knowability or privacy of private experiences, only the possibility of speaking of them without recourse to public language-games. This does not hold of such notions as expecting or hoping, for they do not deal with particular experiences at all. Only of such propositional attitudes is it true that an inner process is in need of external criteria. In general, a distinction is needed between primary and secondary language-games.

(b) (edited with Leila Haaparanta) *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege*, D. Reidel Publishing Co., Dordrecht, 1986, vi+395pp.

(c) (edited with Simo Knuuttila) *The Logic of Being: Historical Studies*, Synthese Historical Library **28,** D. Reidel Publishing Co., Dordrecht, 1986, xvi+300pp.

Papers

(a) Comments and replies", *Philosophia* **2,** Part 1, nos. 1-2, (1986) 105-119 and Part 2, nos. 3-4, (1986), 277-287.

To Geach, more attention to the semantics of deontic logic is recommended. Harrah misses the discourse character of my analysis of questions. P the analogy between denials of "akrasia" and logical omni-science is emphasized. The applicability of L.J. Cohen's comparison between chess and language is limited. Lehrer prompts a further criticism of Bayesianism because of the need of experiential revision of one's indices of caution. Mellema misunderstands my logic of perceptions because he follows natural language too closely.

(b) "Filosofian tulevaisuus", in *Tulevaisuus* (a *Festschrift* for G.H. von Wright), Ilkka Niniluoto and Heikki Nyman, editors, Otava, Helsinki, 1986, 265-275.

(c) "The Languages of Human Thought and the Languages of AI" (résumé), in *AI and Philosophy. STEP-86 Invited Papers*, **1**, M. Karjalainen, J. Seppänen and M. Tamminen, editors, Finnish Society of Information Processing Science, Espoo, 1986, 1-3.

(d) "Logic of Conversation as a Logic of Dialogue", in *Philosophical Grounds of Rationality, Intentions, Categories, and Ends*, Oxford: Clarendon Press, 1986, 259-276.

Grice and Strawson proposed to study the logic of discourse, but their crucial concept (presupposition, conversational maxims, etc.) nevertheless apply only to individual utterances. Moreover, they apply to different kinds of utterances differently, as is shown by references to question-answer dialogues. In a dialogical 'game', rationality can only be attributed to entire strategies, not to individual 'moves' (utterances). Hence concepts like coherence (relevance) are essentially attributes of strategies, not of utterances.

(e) “Quine on Who’s Who”, in *The Philosophy of W.V. Quine,* L.E. Hahn and P. A. Schilpp, editors, Library of Living Philosophers, Open Court, La Salle, Illinois, 1986.

The general problem of “quantifying in” reduces to the model-theoretical problem of cross-identification. In particular, Hintikka’s condition for quantifying into an epistemic context, e.g., the truth of “ “A” knows who “B” is”, is basically model-theoretical, independent of how it is approximately expressed in English. Once this is understood, Quine’s doubts about Hintikka’s condition are resolved. Its apparent vagaries illustrate, in fact, important semantical phenomena, e.g., the duality of methods of cross identification.

(f) “Reasoning about Knowledge in Philosophy: The Paradigm of Epistemic Logic”, in *Reasoning About Knowledge,* Joseph Halpern, editor, Morgan Kaufmann Publishers, Los Altos, CA, 1986, 63-80.

(g) “The Semantics of ‘a certain’”, *Linguistic Inquiry* **17**, no. 2, (1986), 331-336.

(h) “The Varieties of Being in Aristotle”, in *The Logic of Being: Historical Studies,* Synthese Historical Library **28,** Simo Knuuttila and Jaakko Hintikka, editors, D. Reidel Publishing Co., Dordrecht 1986, 81-114.

(i)(with Leila Haaparanta) “General Introduction” in *Frege Synthesized: Essays on the Philosophical and Foundational Work of Gottlob Frege*, Jaakko Hintikka and Leila Haaparanta, editors, D. Reidel Publishing Co., Dordrecht, 1986, 3-8.

(j)(with Merrill B. Hintikka) “Wittgenstein and Language as the Universal Medium”, in *Investigating Wittgenstein,* Basil Blackwell, Oxford, 1986, 1-29.

(k) (with Simo Knuuttila) “Introduction” in *The Logic of Being: Historical Studies*, Synthese Historical Library **28,** Jaakko Hintikka and Simo Knuuttila, editors, D. Reidel Publishing Co., Dordrecht, 1986, ix-xvi.

1987

Papers

(a) “Comment je vois la philosophie”, in French translation, in *Encyclopédie Philosophique,* A. Jacob, editor, Presses Universitaires de France, Paris, 1987.

(b) Comment on Jeffrey’s “Alias Smith and Jones: The Testimony of the Senses”, *Erkenntnis* **26***,* (1987), 407.

(c) Comments on Kamlah’s “What can Methodologists Learn From the History of Probability”, *Erkenntnis* **26**, (1987), 327.

(d) “Extremality Conditions in the Foundations of Mathematics”, in *Philosophy of Science Association 1986* **2,** A.Fine and M. Forbes, editors, Philosophy of Science Association, East Lansing, MI, 1987, 5 pp.

(e) “The Fallacy of Fallacies”, *Argumentation* **1**, (1987), 221-238.

Several of the so-called “fallacies”, in Aristotle were not mistaken inference-types, but mistakes or breaches of rules in the questioning games practiced in the Academy and in the Lyceum. The entire Aristotelian theory of fallacies should therefore be studied by reference to the author's interrogative model on inquiry, not as a part of a theory of inference. Many Aristotelian fallacies can be diagnosed in this way, including *petitio principii*, multiple questions, "babbling", etc., and also his alleged anticipation of *argumentum ad hominem*. Indeed, Aristotle's initial conception of inquiry is an interrogative one. Deductive conclusions caught Aristotle's attention as those answers that every rational inquirer must give, assuming his prior admissions. Several features of Aristotle's methodology can now appreciated, e.g. the role of *endoxa* in it and the typical organization of Aristotle's philosophical discussion of a given problem.

(f) "Game-theoretical Semantics as a Synthesis of Truth-conditional and Verificationist Meaning Theories", in *New Directions in Semantics*, E. LePore, editor, Academic Press, London and Orlando, Florida, 1987, 235-258.

(g) "The Interrogative Approach to Inquiry and Probabilistic Inference", *Erkenntnis* **26**, (1987), 429-442.

(h) "Is Scope a Viable Concept in Semantics?", in *ESCOL '86. Proceedings of the Third Eastern States Conference on Linguistics,* Ann Miller and Zheng-Shen Zhang, editors, Ohio State University, Columbus, Ohio, 1987, 259-270.

(i)"Language Understanding and Strategic Meaning", *Synthese* **73**, (1987), 497-529.

(j)"Logic Translation: An Impossible Dream?", *LMPS 87,**5**, Abstracts*, 1987, 30-32.

(k) "Mental Models, Semantical Games and Varieties of Intelligence", in *Matters of Intelligence: Conceptual Structures in Cognitive Neuroscience,* Lucia Vaina, editor, D. Reidel Publishing Co., Dordrecht, 1987, 197-215.

(l)"Model Minimization: An Alternative to Circumscription", *Journal of Automated Reasoning* **3**, (1987), 1-13.

(m) "A Note on Anaphoric Pronouns and Information Processing by Humans", *Linguistic Inquiry* **18**, (1987), 111-119.

(n) "Replies and Comments", in *Jaakko Hintikka: A Profile*, Radu Bogdan, editor, D. Reidel Publishers, Dordrecht, 1987, 227-344.

(o) "Self-profile", in *Jaakko Hintikka: A Profile,* Radu Bogdan, editor, D. Reidel Publishing Co., Dordrecht, 1987, 3-40.

1988

Papers

(a) "Advice to Prospective Philosophers", in *Proceedings and Addresses of The American Philosophical Association*, Supplement to vol. 62, no.1, (September, 1988), 272-273.

(b) "'Die Wende der Philosophie': Wittgenstein's New Logic of 1928", in *Philosophy of Law, Politics and Society. Proceedings of the 12th International Wittgenstein Symposium*, Hölder-Pichler-Tempsky, Vienna, 1988, 380-396.

Around 1928 Wittgenstein abandoned his belief that truth-function theory is complete. Accordingly, language-world comparisons could not be immediate, but required human operations on the propositions in question. Initially , this meant that atomic propositions contained numerical parameters and that the operations needed were arithmetical calculations. Hence mathematics was now more fundamental for Wittgenstein than logic, a view he also found in Brouwer. Since language belongs to the physical world, those calculations involve temporarlly persistent physical objects. In October 1929 these ideas led Wittgenstein to think that language can directly represent only the world of physical objects.

(c) "Oikeustieteellinen päättely ja oikeusjärjestelmät", *Lakimies,* no. 3, (1988), 219-231.

(d) "On the Development of the Model-theoretical Viewpoint in Logical Theory", *Synthese* **77**, (1988), 1-36.

All model theory presupposes a modicum of belief in what I have called language as calculus, as contrasted to belief in language as the universal medium. A functional interpretation of quantifiers (in Gödel's sense) is argued to be an important aspect of the model-theoretic way of thinking and hence of the entire calculus view. This idea is traced back from Gödel to Hilbert, Löwenheim, Schröder and Peirce, who formulated it most clearly in his semiotic theory. What initally prevented the full development of thie idea was the absence of the concept of strategy in von Neumann's sense.

(e) "On the Incommensurability of Theories", *Philosophy of Science* **55**, (1988), 25-38.

The commensurability of two theories can be defined (relative to a given set of questions) as the ratio of the total information of their shared answers to the total information of the answers yielded by the two theories combined. Answers should be understood here as model consequences (in the sense of the author's earlier papers), not deductive concequences. This definition is relative to a given model of the joint language of the theories, but can be generalized to sets of models. It turns out to capture also the idea of incommensurability as conceptual alienation. Imcommensurability so defined does not imply incomparability.

(f) "Todistiko Gödel matematiikan epätäydelliseksi?", in *Suomalainen Tiedeakatemia vuoikirjassa 1988-89, Esitelmät ja pöytäkirjat*, 1988, 117-126.

(g) "Was Leibniz's Deity an *Akrates*?", in *Modern Modalities: Studies of the History of Modal Theories from Medieval Nominalism to Logical Positivism*, Simo Knuuttila, editor, Synthese Historical Library **33**, Kluwer Academic Publishers, Dordrecht, 1988, 85-108. Jaakko Hintikka, Toim., Simo Knuuttila, Dordrecht, 1988, 85-108.

Leibniz tried to reconcile contingency with lawlikeness by envisaging God choosing the most lawlike world from all possible ones. Since the laws of our actual world do not hold in others, they are metaphysically contingent. But was God's choice really free? God had both the major premise (knowledge of the best world) and the minor premise (power to create it) of a practical syllogism; hence he would be an "akrates" unless he created this particular world. Leibniz's subile response was to reject syllogistic models of rational agency and instead conceptualize decision as a rule-governed resultant of competing force-like "appetites."

(h) "What is the Logic of Experimental Inquiry?", *Synthese* **74**, (1988), Dordrecht, 1988, 173-190.

(i)(with Stephen Harris) "On the Logic of Interrogative Inquiry", in *Philosophy of Science Association 1988* **2**, A.Fine and J. Lepkin, editors, Philosophy of Science Association, East Lansing, Michigan, 1988, 233-240.

1989

Books

(a) *L'intentionnalité et les mondes possibles,* traduit et présenté par Nadine Lavand, in the series Opuscule **6,** dirigée par André Laks et Jean Quillien, Presses Universitaires de Lille, Paris, 1989, 228 pp.

(b) (with Merrill Hintikka) *The Logic of Epistemology and the Epistemology of Logic: Selected Essays*, Synthese Library **200**, Kluwer Academic Publishers, 1989, ix + 243 pp.

In each of the central essays collected here, the authors put forward a general idea apparently capable of sustaining an extensive logico-philosophical theory. They include a new type of semantics for logical modalities; a diagnosis of Frege's and Russell's central problems; a solution to the problem of logical omniscience; a general theory of the individuation and the identification of physical objects; a connection between two different modes of identification and two anatomically distinguishable actual cognitive systems; a possibly sex-linked difference in individuation; the different dimensions of intentionality; and a logical theory of questions and answers.

Papers

(a) "The Cartesian *cogito*, Epistemic Logic, and Neuroscience: Some Surprising Interrelations" in Jaakko Hintikka, with Merrill Hintikka, *The Logic of Epistemology and the Epistemology of Logic: Selected Essays*, Synthese Library **200**, Kluwer Academic Publishers, Dordrecht, 1989, 113-136. (Appears also in *Synthese* **83**, no. 1, (1990), 133-157.)

(b) "Concepts of Scientific Method from Aristotle to Newton", in *Knowledge and the Sciences in Medieval Philosophy. Proceedings of the Eighth Congress of Medieval Philosophy*, *Helsinki, 24-29 August 1987*, Monica Asztalos, John Murdoch and Ilkka Niiniluoto, editors, *Acta Philosophica Fennica* **48**, (1989), 72-84.

The strength of interrogatively construed methodology depends on (1) its initial premises; (2) available answers. Nontrivial conclusions require that (1) or (2) include general propositions. Otherwise inquirers face Hume's problem of induction as an inference from particulars to generalizations. Neither Aristotle, medieval nominalists nor Newton faced this problem; Aristotle because he thought one can perceive the properties and interrelations of general forms in their instantiations in the soul; nominalists (who rejected Aristotle) because they assumed strong initial premises; and Newton because controlled experiments and systematic observations yield general truths about dependencies between variables. Induction meant for them a non-Humean process of extrapolating and interpolating partial generalizations.

(c) "Exploring Possible Worlds", in *Possible Worlds in Humanities, Arts and Sciences. Proceedings of Nobel Symposium 65*, Sture Allén, editor, Walter de Gruyter, Berlin, 1989, 52-73.

(d) "G. H. von Wright on Logical Truth and Distributive Normal Forms", in *The Philosophy of G.H. von Wright*, P. A. Schilpp and L. Hahn, editors, The Library of Living Philosophers **19**, Open Court, La Salle, Illinois, 1989, 517-537.

(e) "Is There Completeness in Mathematics after Gödel?", *Philosophical Topics* **17**, no. 2 (1989), 69-90.

(f) "Is Truth Ineffable?", in *Les formes actuelles du vrai. Entretiens de Palermo 1985*, (no editor indicated) Endichiridion, Palermo, 1989, 89-120.

(g) "Knowledge Representation and the Interrogative Model of Inquiry", in *Knowledge and Skepticism*, Marjorie Clay and Keith Lehrer, editors, Westview Press, Boulder, Colorado, 1989, 155-183.

(h) "Logical Form and Linguistic Theory", in *Reflections on Chomsky,* Alex George, editor, Basil Blackwell, Oxford, 1989, 41-57.

(i) "Ludwig's Apple Tree: Evidence Concerning the Philosophical Relations between Wittgenstein and the Vienna Circle, in *Traditionen und Perspektiven der Analytischen Philosophie: Festschrift für Rudolf Haller*, Wolfgang L. Gombocz, Heiner Rutte und Werner Sauer, editors, Hölder-Pichler-Tempsky, Wien, 1989, 187-202.

Contrary to a widespread view, Wittgenstein shared many philosophical problems and ideas around 1930 with the Vienna Circle. In May 1932 Wittgenstein complained that Carnap's paper "Die physikalisch Sprache als Universalsprache der Wissenschaft" amounted to plagiarism of his new ideas, developed since 1928, in which the primacy of physicalistic languages did play a crucial role. Offended by Carnap's reaction, Wittgenstein later claimed that Carnap had also appropriated ideas from the *Tractatus*. Among such alledged borrowings, Wittgenstein listed the idea of a formal mode of speech and his peculiar conceptions of ostensive definition, hypotheses, and the nature of philosophy.

(j)"On the Limitations of Generative Grammar", in *Proceedings of the Scandinavian Seminar on Philosophy of Language,* Filosofiska Förening and Filosofiska Institutionen vid Uppsala Universitet, Uppsala **26**, no. 1, (1989), 1-92.

(k) "On the Role of Modality in Aristotle's Metaphysics", in *Of Scholars, Savants and Their Texts,* Ruth Link-Salinger, editor, Peter Lang, New York, 1989, 123-134.

(l)"The Paradox of Transcendental Knowledge" in *An Intimate Relation*, J. R. Brown and J. Mittelstrass, editors, Kluwer Academic Publishers, Dordrecht, 1989, 243-57.

(m) "The Role of Logic in Argumentation", *The Monist* **72**, (1989), 3-24.

The author's "interrogative model of inquiry" (MI), which can also be used as a model of argumentation, is used here to diagnose the role of logic in argumentation. In the simplest form of IMI an "inquirer" is arguing for a conclusion "C" from an initial premise "I". Using deductive inferences as new premises. Thus the role of deductive logic in argumentation appears to be that of one component of a larger enterprise. If "strategic" rules (i.e., rules telling how to play "well") are considered, it turns out that the strategic principles of question choice are essentially the same as the strategic principles of deductive logic. (edited)

(n) "Rules, Games and Experiences: Wittgenstein's Discussion of Rule-following in the Light of His Development", *Revue Internationale de Philosophie* **43**, (1989), 279-297.

When Wittgenstein rejected phenomenological languages in 1929, he first could not accommodate rules (especially rules of language) in his new outlook. He had believed (he confessed) that a rule could be gathered from one single experience of using it correctly. This phenomenological conception is what he later criticized in denying that rule-following is a matter of having certain experiences. But a rule as a physicalistic entity (e.g., as a symbolic formula) cannot explain rule-following, either. It has to play a role in some language-game, Wittgenstein eventually argued. Thus language-games are conceptually primary with respect to their rules.

(o) "Todistiko Gödel matematiikan epätäydelliseksi?", in *Finnish Academy of Science and Letters, Year Book 1988-89,* Helsinki, 1989, 117-126. ("Did Gödel Show that Mathematics is Incomplete?")

(p) (with Gabriel Sandu) "Informational Independence as a Semantical Phenomenon", in *Logic, Methodology and Philosophy of Science VIII*, J.E. Fenstad et al., editors, Elsevier, Amsterdam, 1989, 571-589.

Insofar as a formal or natural language can be treated game-theoretically, the notion of informational independence (II), in the sense of game theory, applies to its different ingredients. A notation is proposed for II and the most salient facts about it are noted. Even though II is not indicated syntactically in English, it is the gist of such varied phenomena as the *de dicto* vs. *de re* distinction, complex questions, negation-raising, branching quantifiers, actuality operators, etc. It is therefore an extremely important component of the overall semantics of natural languages, both for philosophical and for linguistic purposes.

1990

Papers

(a) "The Languages of Human Thought and the Languages of Artificial Intelligence", *Acta Philosophica Fennica* **49**, (1990), 307-330.

John von Neumann argued that in computational tasks humans as distinguished from computers are hardwired to prefer low functional depth (low degree of the nesting of functions). The same contrast is here extended to logical reasoning by relating functional depth to quantificational depth via the concept of Skolem function. This has implications to the languages favored by human reasoners as distinguished from the automata of AI.

(b) "Nonstandard Models and the Completeness of Mathematical Theories" in Russian translation, from *The Joint Soviet Finnish Colloquium on Logic and New Tendencies in Logical Semantics, July, 1989* in *Humanism, Science, Technology* **I**, V.S. Stepin, editor-in-chief, Academy of Science, Moscow, 1990, 96-110.

(c) "Obstacles to Understanding" (on the fate of Wittgenstein's *Nachlass*), *Times Literary Supplement,* September 28 October 4 1990, 1030.

(d) "Paradigms for Language Theory" in *Language, Knowledge and Intentionality: Perspectives on the Philosophy of Jaakko Hintikka*, Leila Haaparanta, Martin Kusch and Ilkka Niiniluoto, editors, *Acta Philosophica Fennica* **49**, 181-209.

Language can be viewed either as a rule-governed process (the recursive paradigm) or as a goal-directed process (the strategic paradigm). The former has dominated recent approaches to language, such as generative grammar, reliance on compositionality, formalizations of logical and mathematical reasoning, etc. The strategic paradigm is instantiated by game-theoretical semantics and certain aspects of Wittgenstein's development. Test cases are constituted by semantical phenomena which are not marked syntactically. They are instantiated by transcategorial phenomena like informational independence in the sense of game-theoretical semantics, including the so-called *de dicto* vs. *de re* distinction.

(e) "Quine as a Member of the Tradition of the Universality of Language", in *Perspectives on Quine,* Robert Barrett and Roger Gibson, editors, Basil Blackwell, Oxford, 1990, 159-175.

Quine is considered here as a member of a largely tacit tradition of believers in the universality of (one's home) language and in the ineffability of semantics. This unacknowledged membership is consistent *inter alia* with Quine' disinterest in model theory, his criticisms of modal logic and his belief in the indeterminacy of radical translation. It leaves Quine with the behavior of native speakers as the sole guide for the semantics of their jargon. In this direction, Quine is seriously handicapped by his disregard of strategic behavior as a clue to meaning.

(f) "Wittgenstein as a Philosopher of Immediate Experience", in *Wittgenstein: Towards a Re-evaluation. Proceedings of the 14th International Wittgenstein Symposium,* **1**, Rudolf Haller et al., editors, Hölder-Pichler-Tempsky, Wien, 1990, 155-67.

(g) "Wittgenstein and the Problem of Phenomenology" in *Language, Knowledge and Intentionality: Perspectives on the Philosophy of Jaakko Hintikka,* Leilla Haaparanta, Martin Kusch and Ilkka Niiniluoto, editors, *Acta Philosophica Fennica* **49**, (1990), 15-46.

In early Wittgenstein, simple objects are phenomenological objects. In October 1929 Wittgenstein gave up phenomenological objects as references of names, but still maintained a phenomenological ontology. Later, he envisaged phenomenological languages as an alternative "notation." The contrast between physicalistic and phenomenological discourse then turns on the principle of identification relied on. This helps to understand the meaning of "phenomenology" (as distinguished from "phenomenalism") in Wittgenstein as well as his views on time, memory, solipsism, private language, and identity.

(h) (with Gabriel Sandu) "Metaphor and the Varieties of Lexical Meaning", *Dialectica*, (1990), 55-77.

The "meaning lines" connecting the references of an expression in different situations or scenarios ("worlds") are usually "drawn" with the help of both similarity and continuity. In metaphoric use, emphasis shifts predominantly on suitable similarity considerations; in metonymic use, it shifts on continuity considerations. Even though metaphoric meaning lines are nonstandard, they have to be "anchored" to a literal reference of the expression in some situation or world (not necessarily in the actual one). Hence metaphor is not a matter of truth or of a special kind of language act (use of sentences).

1991

Books

(a) (with James Bachman) *What If…? Toward Excellence in Reasoning*, Mountain View, Mayfield, 1991, v + 465 pp.

This introduction to reasoning uses Hintikka's interrogative model of inquiry. Logical and informal inferences are construed as steps in the same process of inquiry. All new information enters as answers to questions, as in the Socratic questioning method. Other innovations include (i) we distinguish definitory rules of reasoning (they merely tell what is permissible) from strategic principles (they tell how to reason well); (ii) logical rules are formulated so that they apply directly to ordinary language reasoning; (iii) the interrogative model is used to analyze and to construct arguments; and (iv) novel treatments are given of scientific reasoning and of several fallacies.

(b) (with Gabriel Sandu) *On the Methodology of Linguistics*, Basil Blackwell, Oxford, 1991, 186pp.

(c) (edited) *Wittgenstein in Florida. Proceedings of the Colloquium on the Philosophy of Ludwig Wittgenstein, Florida State University, 7-8 August 1989,* Kluwer Academic Publishers, Dordrecht, 1991, 329pp. (Reprinted from *Synthese* **87**, nos. 1-2, (1991).)

Papers

(a) "Carnap, the Universality of Language and Extremability Axioms", *Erkenntnis* **35,** (1991), 325-336.

Mathematicians' attention was called to extremality (maximality, minimality) assumptions by Hilbert's use of an "Axiom of Completeness", in his 1899 *Foundations of Geometry*. Carnap attempted (in his 1936 paper with F. Bachmann) a logical analysis of extremality axioms. However, they tacitly reinterpreted these axioms and hence failed to solve the real problem. This was due to Carnap's belief that the interpretation of a language always involves one single domain of individuals, which belief in turn followed from his commitment to one element in the complex I have called the idea of language as the universal medium.

(b) "Defining Truth, the Whole Truth, and Nothing But the Truth", in *Reports from the*

Department of Philosophy, University of Helsinki, no. 2, (1991), 74 pp.

(c) "Geach and the Methodology of the Logical Study of Natural Language", in *Peter Geach: Philosophical Encounters,* Harry Lewis, editor, Kluwer Academic Publishers, Dordrecht, 1991, 137-149.

Husserl: The Phenomenological Dimension", in *Phenomenology/Fenomenologia. Proceedings of the Symposium on Phenomenology, Jyväskylä, 5 May 1988*, Matti Kosonen, editor, Department of Philosophy, University of Jyväskylä, 1991, 15-28.

(e) "An Impatient Man and His Papers", *Synthese* **87**, (1991), 183-201.

Because of Wittgenstein's impatience as expositor, the problem background of his philosophical ideas is virtually impossible to gather from his so far published writings. Easy access to Wittgenstein's unpublished writings, especially to his notebooks, is therefore badly needed. The Cornell microfilm edition does not adequately serve this purpose, either, even though its availability means that the legal status of the bulk of Wittgenstein's Nachlass is that of published material. The two successive complete works editing projects (the first by a group led by Heringer and Nedo, the second by Nedo) have been abject failures. A change of the editor is therefore recommended.

(f) "The Languages of Human Thought and the Languages of Artificial Intelligence" in *Language, Knowledge and Intentionality: Perspectives on the Philosophy of Jaakko Hintikka,* Leila Haaparanta, Martin Kusch and Ilkka Niiniluoto, editors, *Acta Philosophica Fennica* **49**, 307-330.

(g) "Overcoming 'Overcoming Metaphysics Through Logical Analysis of Language' Through Logical Analysis of Language", *Dialectica* **45**, (1991), 203-218.

Carnap tried to overcome metaphysics through a distinction between empirical and conceptual truths. The distinction has since been challenged, but not on the basis of a systematic logical analysis of language. It is suggested here that the logical theory of identifiability based on the author's interrogative model will provide the tools for such a systematic analysis. As an example of what the model can do, a criticism is offered of Quine's and Chomsky's implicit assumption that language learning is based on atomistic (quantifier-free) "answers" (input).

(h) "Towards a General Theory of Identifiability", in *Definitions and Definability: Philosophical Perspectives,* James H. Fetzer, et al., editors, Kluwer Academic Publishers, Dordrecht, 1991, 161-183.

(i)"Wittgenstein and the Problem of Phenomenology", *Acta Philosophica Fennica* **49**, 15-46.

(j)(with Charles W. Harvey) "Modalization and Modalities", in *Phenomenology and the Formal Sciences,* Thomas M. Seebohm, editor, Kluwer Academic Publishers, Dordrecht, 1991, 59-77.

In Part I of this essay we articulate Husserl's descriptions for the genesis of the primitive logical connectives, negation and disjunction. In Part II we describe possible worlds models for the use of negation and disjunction. In the final part of the essay, we try to show (a) how an appeal to Husserl's analyses of modalization may buffer possible worlds theories of intentionality from charges that it is engaged in a metaphysically naïve enterprise, and (b) how possible worlds methods of analysis may suggest an advisable tactical maneuver for Husserlian phenomenology in light of recent criticisms of Husserl's philosophy of language.

1992

Papers

(a) "Carnap's Work in the Foundations of Logic and Mathematics in a Historical Perspective", *Synthese* **93**, (1992), 167-189.

Carnap's philosophy is examined from new viewpoints, including three important distinctions: (i) language as calculus vs. languge as universal medium; (ii) different senses of completeness; (iii) standard vs. nonstandard interpretations of (higher-order) logic. (i) Carnap favored in 1930-34 the "formal mode of speech," a corollary to the universality assumption. He later gave it up partially but retained some of its ingredients, e.g., the one-domain assumption. (ii) Carnap's project of creating a universal self-referential language is encouraged by (ii) and by the author's recent work. (iii) Carnap was aware of (iii) and occasionally used the standard interpretation, but was not entirely clear of the nature of the contrast.

(b) "The Concept of Induction in the Light of the Interrogative Approach to Inquiry," in *Inference, Explanation and Other Frustrations: Essays in the Philosophy of Science,* John Earman, editor, University of California Press, 1992, 23-43.

(c) "Different Constructions in Terms of 'Knows'", in *A Companion to Epistemology,* Jonathan Dancy and Ernest Sosa, editors, Basil Blackwell, Oxford, 1992, 99B-104B.

(d) "Eino Kaila's 'Blue Fire'", in *Eino Kaila and Logical Empiricism,* Ilkka Niiniluoto et al., editors, *Acta Philosophica Fennica* **52**, (1992), 152-159.

Eino Kaila is usually pigeon-holed as a logical positivist. However, by philosophical temperament he was not a positivist, but not unlike a romantic *Naturphilosoph* who was primarily interested in the secrets of nature rather than our means of knowing it. What attracted him to logical positivists was their interest in contemporary science. Kaila was interested in epistemology, but as a naturalistic study of our actual processes of knowledge acquisition. The most challenging and ultimately frustrating problem (his "blue fire") was for him to understand the sense of causality involved in the contemporary physics.

(e) "Independence-Friendly Logic as a Medium of Knowledge Representation and Reasoning about Knowledge", in *Information, Modelling and Databases,* S. Ohsuga et al., editors, IOS Press, Amsterdam, Washington, Tokyo, 1992, 258-265.

(f) "The Interrogative Model of Knowledge Acquisition as a Framework for Concept Identification", in *Information, Modelling and Databases,* S. Ohsuga et al., editors, IOS Press, Amsterdam, Washington, Tokyo, 1992, 174-181.

(g) "The Interrogative Model of Inquiry as a General Theory of Argumentation", *Communication and Cognition* **25**, (1992), 221-242.

(h) "Knowledge-Seeking by Questioning", in *A Companion to Epistemology,* Jonathan Dancy and Ernest Sosa, editors, Basil Blackwell, Oxford, 1992, 241A-244A.

(i)"Theory-Ladenness of Observations as a Test Case of Kuhn's Approach to Scientific Inquiry", in *PSA 1992. Proceedings of the 1992 Biennial Meeting of the Philosophy of Science Association* **1**, David Hull et al., editors, Philosophy of Science Association, East Lansing, MI, 1992, 277-286.

Kuhn's basic concepts need closer analysis. For instance, the alleged theory-ladenness of observations has several different interpretations. In one sense, it is trivially built into my interrogative approach to inquiry, in that the consequences of an observational answer by nature depnds crucially on the initial theoretical premises. A more interesting sense is obtained by noting the multi-level character of inquiry. A lower-level (experimental or observational) inquiry will on this view depend on earlier results obtained on the higher (theoretical) level.

(j) (with Gabriel Sandu) "The Skeleton in Frege's Cupboard: The Standard vs. Nonstandard Distinction", *Journal of Philosophy* **89**, (1992), 290-315.

Henkin formulated the standard versus nonstandard distinction in 1950, but the idea of the standard interpretation of higher-order variables is virtually equivalent with that of an arbitrary function which was debated by nineteenth-century mathematicians. Frege's disregard of the latter notion and his criticisms of abstraction show that he opted for a nonstandard interpretation. This would make his system an inadequate foundation for mathematics, even if it were consistent. One reason why the nonstandardness of Frege's interpretation has been overlooked is that he did not identify higher-order existence with definability, which is falsely assumed to be the only possible nonstandard interpretation.

1993

Papers

(a) "Gödel's Functional Interpretation in a Wider Perspective", in *Yearbook 1991 of the Kurt Gödel Society,* H.D. Schwabl, editor, Kurt Gödel Society, Vienna, 1993, 1-39.

(b) "A Historical Note on Scott's 'Game-theoretical Interpretation of Logical Formulae'", in *Yearbook 1991 of the Kurt Gödel Society,* H.D. Schwabl, editor, Kurt Gödel Society, Vienna,1993, 45.

(c) "New Foundations for Mathematical Theories", in *Logic Colloquium 90: Lecture Notes in Logic,* no. 2, J. Väänänen and J. Oikkonen, editors, Springer, Berlin, 1993, 122-144.

(d) "The Original *Sinn* of Wittgenstein's Philosophy of Mathematics", in *Wittgenstein's Philosophy of Mathematics* **2**, Klaus Puhl, editor, Hölder-Pichler-Tempsky, Vienna, 1993, 24-52.

(e) "Socratic Questioning, Logic, and Rhetoric", *Revue Internationale de Philosophie* **47**, (1993), 5-30.

The earliest form of Aristotle's methodology was a dialectic modelled on the Socratic method of questioning. Logic originated as a study of such answers received in a dialectical game as were necessitated by earlier answers. Even after the ideas of syllogistic and syllogistically organized science were developed, Aristotle's conception of method contained a major dialectical element. Aristotle's rhetoric is but another variant of the same dialectical methodology. Several of its main features can be understood better in terms of the logic of dialectical (interrogative) inquiry, for instance Aristotle's comments on the role of the speaker's character in rhetorical persuasion.

1994

Books

(a) *Fondements d'une théorie du langage*, traduit de l'américain par Nadine Lavand, Presses Universitaire de France, Paris, 1994, xviii + 436 pp.

(b) *La vérité est-elle ineffable? et autres essais*, traduit de l'anglais par Antonia Soulez et Francois Schmitz, Collection 'tiré à part', dirigée par Jean-Pierre Cometti, Éditions de l'Éclat, Combas, 1994, 126 pp.

Papers

(a) "An Anatomy of Wittgenstein's Picture Theory", in *Artifacts, Representations and Social Practice,* C.C. Gould and Robert S. Cohen, editors, Kluwer Academic Publishers, Dordrecht, 1994, 223-256.

Wittgenstein's so-called picture theory involves several different assumptions: (i) elementary propositions are "pictures", i.e. isomorphic replicas, of the corresponding states of affairs. (ii) The totality of possible combinations of simple objects matches the totality of elementary propositions. (iii) A name shares a logical form with its object. (iv) Elementary propositions are mutually independent. (v) Complex propositions are pictures in the same sense as elementary ones. (vi) The logical forms of propositions can be represented by their syntactical forms. Later, Wittgenstein rejected at least (ii), (iv), and (vi), but not the general idea of propositions as pictures.

(b) "Qu'est-ce que la logique élémentaire? La logique faite pour l' indépendence est le coeur même de la logique" in *Fondements d'une théorie du langage,* traduit de l'américain par Nadine Lavand, Presses Universitaire de France, Paris, 1994, 271-317.

(c) (with Ilpo Halonen) "Quantum Logic as a Logic of Identification", in *Patrick Suppes:*

Scientific Philosopher **3**, Paul Humphreys, editor, Kluwer Academic Publishers, 1994, 125-

145.

(d) (with Gabriel Sandu) "Uses and Misuses of Frege's Ideas", *The Monist* **77**, (1994), 278-293.

Frege's achievement as the creator of contemporary logic should not blind us to the limitations of his approach to logic, language and mathematics. Frege accepted the universality of language and its corollaries, the ineffability of semantics, the one-world view of meaning, and the hypostatization of meanings into meaning entities (Sinne). He assumed compositionality and hence overlooked informationally independent quantifiers. His thesis of the ambiquity of words like "is" is unacceptable in the semantics of natural languages. In higher-order logic, he assumed a nonstandard interpretation, which made it impossible for him to handle the important idea of arbitrary function.

(e) (with Gabriel Sandu) "What is a Quantifier?", *Synthese* **98**, (1994), 113-129.

Quantifiers have been interpreted (i) as higher order predicates (Frege), (ii) substitutionally, and (iii) as deputizing choice (Skolem) functions (Hilbert). Game-theoretical semantics is an implementation of (iii), and its success provides evidence for (iii). It vindicates the notion of informationally independent quantifiers. Such quantifiers cannot be accommodated by interpretations (i) and (ii), which further supports (iii). The theory of so-called generalized quantifiers relies on (i), and hence cannot do justice to independent quantifiers or be a fully general theory of quantifiers.

(f) (with Gabriel Sandu) "Why Parallel Processing?", in *Philosophy and Cognitive Sciences. Proceedings of the Sixteenth International Wittgenstein Symposium,* Robert Casati, Barry Smith and Graham White, editors, Hölder-Pichler-Tempsky, Vienna, 1994, 265-272.

1995

Books

(a) (edited with Klaus Puhl) *The British Tradition in 20th Century Philosophy. Proceedings of the 17th International Wittgenstein Symposium*, Hölder-Pichler-Tempsky, Vienna, 1995, 385 pp.

(b) (edited) *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics*, Kluwer Academic Publishers, Dordrecht, 1995, x + 459pp.

Papers

(a) "Commentary on Allen", *Proceedings of the Boston Colloquium on Ancient Philosophy* **11**,

John J. Cleary and William Wians, University Press of America, Lanham, Maryland, 1995,

206-213.

(b) "Commentary on Smith's 'What Use is Aristotle's *Organon*?'", in *Proceedings of the*

Boston Area Colloquium in Ancient Philosophy **9**, John J. Cleary and William Wians, editors, University Press of America, Lanham, Maryland, 1995, 286-295.

Smith construes Aristotle's *Organon* as a characterization of epistemic virtue, not as a methodological treatise. But for Aristotle the obvious model of the induction of epistemic virtue is the Socratic *elenchus*. Hence Smith's view implies that Aristotle's methodology was an interrogative (dialectical) one. This agrees with the *Topics,* and there also is a hidden dialectical ingredient in the *Analytics*. This interrogative character of Aristotle's methodology provides also a perspective on Aristotle's anti-regress argument Smith discusses: since atomic premises are the best answers to what-questions, a regress would mean that we do not know what we are talking about.

(c) "Constructivism *aufgehoben*", in *Logica '94*, T. Childers and O. Majer, editors, Filosofia, Praha, 1995, 1-15.

(d) "Finnish Philosophy", in *The Oxford Companion to Philosophy*, Ted Honderich, editor, Oxford University Press, 1995, 281-282.

(e) "The Games of Logic and the Games of Inquiry", *Dialectica* **49** (1995), 229-249.

It has been suggested that truth should instead be characterized by reference to the "language-games" of verification and falsification. The author's game-theoretical semantics (GTS) here explained for formal first-order languages, can be thought of as a realization of this idea. More technically speaking, GTS can also be thought of as a systematization of the well-known "epsilon-delta" definitions in the foundations of analysis. In GTS, truth is not defined by reference to winning a play of a game, but a the existence of a winning strategy in the game for the verifier. In a first-order language, the game-theoretical truth-condition of a sentence S can accordingly be expressed by an explicit second-order (sigma-one-one) sentence asserting the existence of the Skolem function of S. (edited)

(f) "The Longest Philosophical Journey: Quest of Reality as a Common Theme in Bloomsbury", in *The British Tradition in Twentieth Century Philosophy. Proceedings of the Seventeenth International Wittgenstein Symposium,* Klaus Puhl and Jaakko Hintikka, editors, Hölder-Pichler-Tempsky, Vienna, 1995, 1-26.

(g) "Meinong in a Long Perspective", *Grazer Philosophische Studien* **50**, Rudolf Haller, editor, (listed 1995, appeared in 1996), 29-45.

Meinong's thought is considered in relation to several major conceptual problems, including the Frege-Russell thesis that words like *is* are multiply ambiguous and Aristotle's treatment of existence. This treatment leads to a problem of how to interpret quantifiers. The three main possible interpretations are (i) quantifiers as ranging over actual individuals (or individuals existing in some one world); (ii) quantifiers as ranging over a set of possible individuals; (iii) quantifiers merely as a way of specifying the interdependencies of the concepts (forms) specified by syllogistic terms. The subsequent history of philosophers' and logicians' treatments of existence is characterized by a tension between (i)-(iii). Meinong's position is in the main (iii) whereas Russell in his *On Denoting* defended (i). The contrast between (i)-(iii) has a counterpart in nineteenth-century discussions about foundations of mathematics.

(h) "On Proper (Popper?) and Improper Uses of Information in Epistemology", *Theoria* **59**, 158-165.

The following theses are put forward:

(1) Information is specified by specifying which alternatives concerning the reality it admits.

(2) These alternatives do not normally concern the state or the history of the entire universe but only of some small part of it.

(3) Information and probability are inversely related.

(4) A purely logical definition of information is impossible.

(5) The use of information as a goal is compatible with the use of inductive probabilities.

(6) Nonzero inductive probabilities can be associated with strict generalizations also in infinite universes.

(7) There are several different kinds of information which can serve as utilities in an epistemic decision.

(i) "The Phenomenological Dimension", in *The Cambridge Companion to Husserl*, Barry Smith and David W. Smith, editors, Cambridge University Press, 1995, 78-105.

Husserlian intentionality operates via what is immediately given in experience. Phenomenological reductions are calculated to uncover these given elements in experience and the constitution of our world from them. What is given are not phenomena, but part of the reality. Husserl's term for the medium of immediate givenness is *Anschauung*. Husserl's transcendental reduction can be compared with Russell's reduction to acquaintance. A major difference is that for Russell the unedited given is already categorially structured into objects of different logical types, whereas for Husserl empirical experience yields in the first place only unstructured *hyle* on which we impose forms.

(j) "Standard vs. Nonstandard Distinction: A Watershed in the Foundations of Mathematics", in *From Dedekind to Gödel: Essays on the Development of the Foundations of Mathematics,* Jaakko Hintikka, editor, Kluwer Academic Publishers, Dordrecht, 1995, 21-44.

(k) "What Is Elementary Logic? Independence-friendly Logic as the True Core Area of Logic", in *Physics, Philosophy, and the Scientific Community,* Kostas Gavroglu, John Stachel and Marx W. Wartofsky, editors, Kluwer Academic Publishers, Dordrecht, 1995, 301-326.

(l)(with Ilpo Halonen) "Semantics and Pragmatics for Why-Questions", *Journal of Philosophy* **92**, (1995), 636-657. (Appears also in Jaakko Hintikka, *Inquiry As Inquiry: Toward a Logic of Scientific Discovery, Selected Papers* **V**, Kluwer Academic Publishers, Dordrecht, 1999, 183-204.

Questions like *Why* P(b)? constitute a degenerate case of statement questions in which presuppositions and answers (in the technical sense) collapse. Answering it consists in deriving P(b) interrogatively from the relevant initial premises T. In typical circumstances, there exists a "covering law formula" H[x] such that H[b] is derivable from nature's answers and the universal implication from H[x] to P(x) derivable from T. This H is what in ordinary usuage is the "answer" to a why-question. It is an explanation of P(b) in that it in effect is a summary of the entire argument from T to P(b).

(m) (with Byong-Chul Park) "The Background of Wittgenstein's Phenomenology", *Phenomenological Inquiry* **19** (1995), 134-148.

(n) (with Gabriel Sandu) "The Fallacies of the New Theory of Reference", *Synthese* **104**, (1995), 245-283.

The so-called New Theory of Reference (Marcus, Kripke etc.) is inspired by the insight that in modal and intensional contexts quantifiers presuppose nondescriptive unanalyzable identity criteria which do not reduce to any descriptive conditions. From this valid insight the New Theorists fallaciously move to the idea that free singular terms can exhibit a built-in direct reference and that there is even a special class of singular terms (proper names) necessarily exhibiting direct reference. This fallacious move has been encouraged by a mistaken belief in the substitutional interpretation of quantifiers, by the myth of the *de re* reference, and a mistaken assimilation of "direct reference" to ostensive (perspectival) identification. The *de dicto* vs. *de re* contrast does not involve direct reference, being merely a matter of rule-ordering ("scope"). The New Theorists' thesis of the necessity of identities of directly referred-to individuals is a consequence of an unmotivated and arbitrary restriction they tacitly impose on the identification of individuals.

(o) (with Gabriel Sandu) "What Is the Logic of Parallel Processing?", *International Journal of Foundations of Computer Science* **6** (1995), 27-49.

We can associate with each consistent formula *F* of first-order logic a computing device as its representation. This computing device is one which will calculate the Skolem functions of *F* (for a denumerable domain). When two such devices are operating in parallel, the resulting architecture does not necessarily reprresent any ordinary first-order formula, but it will represent a formula in independence-friendly (IF) logic, which hence can be considered as a true logic of parallel processing. In order to preserve representability by a digital automaton (Turing machine), a nonstandard (constructivistic) interpretation of the logic in question has to be adopted. It is obtained by restricting the Skolem functions available to verify a formula *F* to recursive ones, as in the Gödel's *Dialectica* interpretation.

1996

Books

(a) *The Principles of Mathematics Revisited*, Cambridge University Press, Cambridge, 1996, xii + 288 pp.

A new basic first-order logic is proposed and used to explore the foundations of mathematics. This new logic enables logicians to express on the first-order level such concpets as equicardinality, infinity and truth in the same language. The famous impossibility results by Gödel and Tarski that have dominated the field for the past sixty years turn out to be much less significant than has been thought. All of ordinary mathematics can in principle be done on this first-order level, thus dispensing with all problems concerning the existence of sets and other higher-order entities. (publisher, editor)

(b) *La philosophie des mathematiques chez Kant. La structure de l'argumentation transcendantale,* traduit de l'anglais par Corinne Hoogaert, in the series L'interrogation philosophique, dirigée par Michel Meyer, Presses Universitaires de France, Paris, 1996, viii + 312 pp.

(c) *Ludwig Wittgenstein: Half-Truths and One-and-a-Half Truths*, *Selected Papers* **I**, Kluwer Academic Publishers, Dordrecht, 1996, xiv + 353 pp.

Frequently, a genuine understanding of a thinker's ideas is possible only by following them further than he did himself. Wittgenstein's Viennese contemporary Karl Kraus spoke in a similar context of one-and-a-half truths in contradistinction to half-truths. In this volume of essays, Jaakko Hintikka examines in the spirit of Kraus's "bon mot" the two grand visions concerning the interrelations of language, self and the world that guided Wittgenstein's thought at the different stages of his philosophical development. He shows how one of them, the so-called picture theory of language, was in reality a combination of several independent assumptions, while the other, the idea of language-games as the vehicles of meaning, was the end product of an intriquing development. Alas, the role of these two fundamental visions is in Wittgenstein's published books largely hidden by his legendary impatience as an expositor. To counter this impatience, Hintikka shows that many of Wittgenstein's best-known ideas can, and must, be understood as defenses or rationalizations of his overall visions. In several essays, Wittgenstein's ideas are illuminated through comparisons with other philosophers, including Russell, Husserl and Carnap.

Papers

(a) "Ajatuksia Aristoteleen ajattelua koskevista ajatuksista" ["Thoughts on Aristotle's Thoughts about Thinking"], in *Sielun liikkeitä,* Taina and Toivo Holopainen, editors, Gaudeamus, Helsinki, 1996, 28-42.

(b) "Cogito, ergo quis est?", *Revue Internationale de Philosophie* **50**, (1996), 5-21.

The "performatory", interpretation of the cogito relies on a logical parallelism between the self-refuting character of asserting "I don't exist" and Descartes's skeptical thought-experiment of trying to think the same. But who is the "I" here? (It is imaginable that someone could perform the cogito and yet not enjoy bodily existence.) In fact only perspectivally identified entities can be proved to exist by the cogito, e.g. Descartes's "I" but not Cartesius. – Along separate lines, it is shown that Descartes's logic of existence in the cogito agrees with Aristotle's treatment of existence in the context of a syllogistic science.

(c) "Contemporary Philosophy and the Problem of Truth", in *Methods of Philosophy and the History of Philosophy,* Simo Knuuttila and Ilkka Niiniluoto, editors, *Acta Philosophica Fennica* **61**, (1996), 23-39.

Contemporary philosophy exhibits a contrast between two grand views: (1) our language (and the thinking it embodies) is inescapable and inexpressible; (2) our language can be discussed, varied and theorized about in language. The view (1) has prompted philosophers like Heidegger to postulate a special hermeneutical approach to philosophy. A test case is the (in)definability of truth. Tarski's 1935 indefinability theorem has been taken to support (1). However, in independence-friendly languages truth is definable for the same language. This shows that irrelevance of Tarski's theorem and deprives the hermeneutical approach much of its motivation.

(d) "Knowledge Acknowledged: Knowledge of Propositions vs. Knowledge of Objects", *Philosophy and Phenomenological Research* **56**, (1996), 251-275.

By allowing a concept to be informationally independent of another one even when it is within its syntactical scope, the first adequate epistemic logic can be formulated, with *knows that* as the only irreducible kind of knowledge. In this logic, two largely independent derivative concepts of knowledge are distinguished: knowledge of propositions and knowledge of objects. (In the former, an existential quantifier is independent of a sentence-initial knowledge operator; in the latter, a disjunction.) Distinctions like *de dicto* vs. *de re* or attributive vs. referential are derivative scope distinctions. There is no unanalyzable knowledge *de re*.

(e) "On the Development of Aristotle's Ideas of Scientific Method and the Structure of Science", in *Aristotle's Philosophical Development*: *Problems and Prospects*, William Wians, editor, Rowman and Littlefield, Savage, Maryland, 1996, 83-104.

Aristotle's early methodology was an interrogative one modelled on the Socratic *elenchus*. His strategic interest led him to study answers that are necessitated by earlier ones. This study became his syllogistic logic, which is thus still a part of the general theory of interrogative inquiry. However, the syllogistic methodology was by itself too narrow to cope with the problems of change or the special role of the widest premises of a science. Aristotle dealt with the latter by his idea that existence assumptions trickle down in a series of syllogisms from wider to narrower terms.

(f) "Ovatko uutiset analyyttisen filosofian kuolemasta liioiteltuja?" (Is the news about the death of analytic philosophy exaggerated?), in I.A. Kieseppä et al., editors, *Tieto, totuus ja todellisuns,* Gaudeamus, Helsinki, 1996, 267-280. (Reprinted in *Filosofian köyhyys ja rikkaus: Nykyfilosofian kartoitusta*, Janne Hiipakka and Risto Vilkko, editors, Art House Oy, Helsinki, 2001, 21-39.)

(g) "The Place of C.S. Peirce in the History of Logical Theory", in *The Rule of Reason: The Philosophy of Charles Sanders Peirce,* Jacqueline Brunning and Paul Forster, editors, University of Toronto Press, Toronto, 1996, 13-33.

In the grand contrast between language as the universal medium and the model-theoretical conception of language, Peirce belongs squarely to the latter camp. This is shown by his own testimony, his work in modal logic, his anticipation of game-theoretical semantics, his acceptance (and cultivation) of metalogic, his insistence on the iconicity of logic, and his theorematic vs. corollarial distinction. Peirce's model-theoretical approach is also a precondition of his idea that human action is constitutive of meaning and indeed a precondition of his entire pragmati(ci)sm. It distinguishes him sharply from such universalists as Frege, Wittgenstein and Quine.

(h) "Possible Worlds – Possible Individuals", in *Philosophy of Language An International Handbook of Contemporary Research* **2**, Marcelo Dascal et al., editors, Walter de Gruyter, Berlin, 1996, 1271-1278.

(i)"Strategic Thinking in Argumentation and Argumentation Theory", *Revue Internationale de Philosophie* **50**, (1996), 307-324.

All rational argumentation, exemplified by the "deductions" of the likes of Sherlock Holmes, can be construed as a question-answer sequence, interspersed by deductive inferences. The implementation of this idea presupposes a new, better logic of questions and answers which uses the notion of informational independence. What is better and worse in argumentation is determined by strategic rules as distinguished from definitory rules which merely specify which steps in the process are admissible. Although the definitory rules for questioning and for deduction are entirely different, the strategic rules governing the two are nearly identical in purely discovery-oriented reasoning.

(j) "Wittgenstein on Being and Time", in Jaakko Hintikka, *Ludwig Wittgenstein: Half-truths and One-and-a-half-truths,* Kluwer Academic Publishers, Dordrecht, 1996, 3-18.

Wittgenstein distinguished between memory time and information time. This distinction is closely related to, but not identical with, a distinction between perspectivally identified and publicly identified time reference and also to one between phenomenological and physical time. The distinction is also related to the problem of integrating different perspectival time frames into a single public (physical) time. In the *Tractatus*, memory time was the basic one (our propositions are verified in the present) but was replaced later by information (physical) time as our basic concept of time.

(k) "World Lines and Their Role in Epistemic Logic", in *Philosophical Logic and Logical Philosophy: Essays in Honour of Vladimir A. Smirnov*, Peter Bystrov and Vadim Sadovsky, editors, Kluwer Academic Publishers, Dordrecht, 1996, 121-137.

(l)(with Marcelo Dascal and Kuno Lorenz) "Games in Language", in *Philosophy of Language: An International Handbook of Contemporary Research* **2**, Marcelo Dascal et al., editors, Walter de Gruyter, Berlin, 1996, 1371-1391.

(m) (with Gabriel Sandu) "Game-Theoretical Semantics", in *Handbook of Logic and Language,* Johan van Benthem and Alice ter Meulen, editors, Elsevier, Amsterdam, 1996, 361-410.

(n) (with Gabriel Sandu) "A Revolution in Logic?", *Nordic Journal of Philosophical Logic* **1**, (1996), 169-183.

Frege formulated his logic so as to rule out some possible and interpretable patterns of dependence and independence between quantifiers. When they are restored, we obtain independence-friendly (IF) first-order logic. This logic is stronger than ordinary first-order logic. It admits a complete disproof method but not semantically complete axiomatization. This incompleteness makes nonsense of Frege's and Hilbert's foundational projects, but opens the possibility of descriptively complete axiomatization where it was previously impossible. This incompleteness does not affect the requirements on formal proofs in mathematics except that all valid inference patterns cannot now be recursively enumerated. New patterns can be discovered, not by intuition, but by model-theoretical considerations. Among other things, the axiom of choice can be vindicated in this way.

1997

Books

(a) *Lingua Universalis vs. Calculus Ratiocinator: An Ultimate Presupposition of Twentieth-Century Philosophy, Selected Papers* **II**, Kluwer Academic Publishers, Dordrecht, 1997, xxii + 268 pp.

The essays collected here explore a fundamental contrast between two overall visions of language and its availability to self-examination. They can be characterized as "language as the universal medium" nd "language as calculus" (or the model-theoretical view). The former normally includes the ineffability of semantics and a one-world ontology. This contrast has dominated twentieth-century philosophy but has scarcely been acknoledged before. Philosophers examined here from the vantage point of the contrast include Peirce, Frege, Wittgenstein, Carnap, Quine, Husserl and Heidegger. Tarski's famous result concerning the indefinability of truth seems to decide the issue in favor of the universalists. Hintikka nevertheless shows that Tarski's result is inconclusive and that truth can in fact be defined in languages which are in certain respects comparable to ordinary language. This unique volume is a must for every contemporary philosopher and for everyone interested in the semantics of our language. (publisher)

Papers

(a) "Commentary on Allen", in *Proceedings of the Boston Area Colloquium in Ancient Philosophy* **11**, John J. Cleary and William Wians, editors, University Press of America, Lanham, Maryland, 1997 (for 1995), 206-213.

Allen correctly describes the *Topics* as a handbook of dialectical argumentation, and asks how Aristotle's logic developed from such argumentation. But passages like *An. Post*. I, 75a 18-27 show that Aristotle continued to think of all inferences as steps in interrogative inquiry. Logical inferences are merely those question-answer steps where the answer is necessitated by earlier answers. Yet it can still be necessary *ad hominem*. The problem of Aristotle's development concerns the separation of such necessitated answers from others. The difficulty in interpreting Aristotle is largely due to his failure to distinguish definitory and strategic rules from each other.

(b) "A Game Theory of Logic – A Logic of Game Theory", *Vienna Circle Institute Yearbook* **5**, (1997), 315-323.

Game theory is sometimes claimed to be a general theory of rationality. If so, it should have applications outside its original scope of competitive and ecomonic activities. The author's game-theoretical semantics is a case in point. It has led to significant new developments, including independence-friendly logic, where independence means informational independence as in game theory. All games must nevertheless be actually playable. One corollary is that human players must be able to form their strategies in response to other players' moves. The question is also raised whether playability requires that one's strategies be computable.

(c) "Hilbert Vindicated?", *Synthese* **110**, no. 1, (1997), 15-36.

The professed reasons for classifying Hilbert a formalist are largely mistaken. The only half-way valid reason is Hilbert's preference for concrete symbols as the objects dealt with in the foundations of mathematics. This preference was nevertheless only a small part of Hilbert's campaign against general concepts and for concrete individuals in logic and in the foundations. For Hilbert, mathematics is not a set-theoretical but a combinatorial enterprise. The same orientation is illustrated by Hilberts' use of the epsilon-symbol. Hilberts' ideas are partly vindicated by the author's independence-friendly first-order (and hence "nominalistic") logic to which all usual mathematical truths can in a sense be reduced.

(d) "The Idea of Phenomenology in Wittgenstein and Husserl", in *Phänomenologie und Logischer Empirismus: Zentnarium Felix Kaufman*, Friedrich Stadler, editor, Springer-Verlag, Vienna, 1997, 127-151.

(e) "No Scope for Scope?", *Linguistics and Philosophy* **20**, (1997), 515-544

The notion of scope is ambiguous, indicating both logical priority and the limits of binding. A separation of these two in first-order logic yields a stronger logic which automatically solves the donkey sentence problem. In natural language, priority scope is determined by ordering principles which do not reduce to bracketing. Binding scope disappears, for natural-language quantifier phrases do not operate like variable-binders but as scopeless description-like terms whose quantifiers range over certain choice sets that depend on the stage a semantical game has reached and hence cannot be indicated by bracketing.

(f) "On Creativity in Reasoning", in *The Complexity of Creativity*, Å.E. Andersson and N.-E. Sahlin, editors, Kluwer Academic Publishers, Dordrecht, 1997, 67-78.

(g) "Replies", in *Knowledge and Inquiry: Essays on Jaakko Hintikka's Epistemology and Philosophy of Science,* Matti Sintonen, editor, Poznan Studies in the Philosophy of Science, Rodopi, Amsterdam, 1997, 311-340.

This is a series of answers to, and comments on, all the other contributions to the same volume. The topics touched on include Aristotle and plenitude; Aristotle's dialectic; theories of questions in German philosophy around 1900; relationship between Wittgenstein and Ramsey; inductive logic; the atomistic postulate; caution and nonmonotonic inference; identifiability; questions and natural kinds; structure of inquiry; semantics of questions; science and games; explanation; interrogative approach to inquiry; and the logical structure of learning models.

(h) "A Revolution in the Foundations of Mathematics?", *Synthese* **111**, no. 2, (1997), 155-170.

The received picture of the foundations of mathematics consists of (semantically complete) first-order logic supplemented by higher-order logic or, usually, set theory. But ordinary first-order logic involves needless restrictions on quantifier interplay whose removal results in a new *independence-friendly first-*

order logic. It captures many mathematical notions (equicardinality, infinity, etc.) ordinary first-order logic does not capture. It is semantically incomplete, but it facilitates descriptively complete axiomatization of many mathematical theories. In a sense, any mathematical theorem is interpretable as a truth of this logic. In brief, it makes in principle set theory and higher-order largely dispensable.

(i)"Three Dogmas of Quine's Empiricism", *Revue Internationale de Philosophie* **51**, (1997),

457-477.

The thought of W.V. Quine is shaped by three important but largely unarticulated presuppositions which he shares with many other philosophers.

(1) Quine assumes that our discourse is meaningful only insofar as it pertains to the actual world. (The one-world assumption.) Quine sees no use for concepts which for instance serve to locate the actually realized scenario on the map of numerous possible ones. This excludes him from realistic modal logic, philosophically and linguistically significant model theory, realistic conceptions of probability etc.

(2) A related assumption in Quine is that we cannot meaningfully and nontrivially speak in our language of its own semantics. (The ineffability of semantics.) This reduces semantics for Quine to the study of the linguistic behavior of language users. It is likewise this ineffability thesis that lies at the bottom of Quine's rejection of the notion of analyticity. If genuine meaning attributions could be expressed and discussed in language, they would be on the same level as any other empirical hypotheses, and so would be e.g. meaning postulates.

(3) Like many other philosophers Quine in effect thinks that the most basic input into our cognitive processes consists of information codifiable in particular propositions. (The atomistic postulate.) Because of the atomistic postulate and of the one-world assumption Quine's projected behavioristic language theory labors under several serious handicaps.

None of the three presuppositions is acceptable. The one-world assumption misconstrues the actual semantics of our language. The ineffability thesis is disproved by recent developments in logical semantics, and the atomistic postulate is contradicted as an epistemological thesis by the method of controlled experiment and contradicted as a psychological thesis by the phenomenology of human cognition, including the phenomenology of perception.

(j) "What was Aristotle Doing in his Early Logic, Anyway? A Reply to Woods and Hansen", *Synthese* **113**, no. 2, (1997), 241-249.

Allegedly against my earlier interpretations, Woods and Hansen argue that several typical Aristotelian fallacies are mistakes in reasoning rather than in questioning. But this contrast did not exist for Aristotle. For him our "logical inferences" were merely special kinds of steps in interrogative inquiry, viz. answers that are (as we would say) logically implied by early answers. In Aristotle's terminology, such interrogative steps are necessary *ad argumentum*, and not merely *ad hominem*. Either kind of step instantiates both definitory and strategic rules. This analysis illuminates such Aristotelian fallacies as *petitio principii* and the fallacy of many questions.

(k) "Who Is About to Kill Analytic Philosophy?", in *The Story of Analytic Philosophy*, Anat Biletzki and Anat Matar, editors, Routledge, London, 1997, 253-269.

1998

Books

(a) *Language, Truth and Logic in Mathematics, Selected Papers* **III**, Kluwer Academic Publishers, Dordrecht, 1998, x + 247 pp.

(b) *Paradigms for Language Theory and Other Essays, Selected Papers* **IV,** Kluwer Academic Publishers, Dordrecht, 1998, x + 310 pp.

(c) *Questions de logique et de phénoménologie*, Élisabeth Rigal, editor, Élisabeth Rigal, et al., translators, in the series Problèmes et Controverses, Jean-François Courtine, directeur, Librairie Philosophique, J. Vrin, Paris, 1998, 338pp.

(d) *El viaje filosófico más largo: De Aristóteles a Virginia Woolf*, Marcelo M.M. Hurtado, translator, Gedisa Editorial, Barcelona, 1998, 287pp.

Papers

(a) "*Argumentum ad hominem:* Will the Real Fallacy Please Stand Up?", *Armenian Mind* II, no. 1, (1998), 45-60.

(b) "Der Formelkram ist nur eine Sprache", in *Einladung zum Denken*: *Ein kleiner Streifzug durch die Analytische Philosophie,* Dagmar Borchers, Olaf Brill and Uwe Czaniera, editors, Verlag Hölder-Pichler-Tempsky*,* Wien, 1998, 133-42.

(c) "On Gödel's Philosophical Assumptions", *Synthese* **114**, (1998), 13-23.

Gödel was a one-world theorist who did not use the idea of other possible worlds or scenarios. Logical truths were for him not truths in all possible worlds, but truths about certain abstract entities in *this* world. As a consequence, Gödel failed to distinguish between different kinds of (in)completeness. He proved the *deductive* incompleteness of elementary arithmetic, but this implies *descriptive* incompleteness only if the underlying theory is *semantically* complete. Because of the same one-world stance, Gödel had to postulate a special supersensory access to his abstract entities, viz. mathematical intuition.

(d) "Perspectival Identification, Demonstratives and 'Small Worlds'", in Jaakko Hintikka, *Paradigms for Language Theory and Other Essays*, *Selected Papers* **IV**, Kluwer Academic Publishers, Dordrecht, 1998, 219-249.

Demonstratives are characterized by their reliance on perspectival rather than public identification method. These two differ in different alternatives to some given situation of language use, not in different situations of use. Many noun phrases can be used demonstratively, i.e. with a *de re* construction with respect to perspectival identification (type three demonstratives). Words like "this" and "that" have their reference fixed ostensively. They are called "type two demonstratives." The reference of "I", "here", and "now" seems already fixed by the situation (type one demonstratives). Yet they too can be thought to rely on tacit ostension.

(e) "The Pragmatic Fallacies of the New Theory of Reference," *Pragmatics and Cognition* **6**, nos. 1-2, (1998), 9-20.

As is well known, according to the "new" theory of reference, the reference relation can be carried out by means of rigid designators whose relationship with the object they designate cannot be analyzed away. Moreover, the new theorists claim, the category of proper names in a natural language marks almost invariabley rigid designators. In this paper, both claims are rejected. Using distinctions between the referential system (which determines which entities the primitive symbols of language refer to in each possible world) and the identification system (which determines which member of one world is identical with which member of another), and between two types of object identification (public and perspectival), it is argued that the use of a noun phrase as a rigid designator is predicated on the assumption that a language user knows who (or what) the noun phrase refers to in the actual world. The conclusion is that rigid designation is not a conceptually irreducible reference relation, nor are proper names always used as rigid designators.

(f) "Ramsey Sentences and the Meaning of Quantifiers", *Philosophy of Science* **65**, (1998), 289-305.

From a (first-order) theory T, its Ramsey reduct r(T) is obtained by replacing all theoretical terms to variables bound to initial second-order existential quantifiers. T and r(T) are not equivalent, which seems to imply that Ramsey reduction is not a genuine elimination of theoretical terms. However, if the basic logic is independence-friendly first-order logic, the reduct is equivalent to a first-order sentence. Does that show that theoretical concepts are eliminable? No, because even first-order quantifiers introduce in effect theoretical concepts, viz. the Skolem functions of the sentence in which they occur.

(g) "Réponses et commentaires", Élisabeth Rigal, translator, in Jaakko Hintikka, *Questions de logique et de phénoménologie*, Élisabeth Rigal, editor, in the series Problèmes et Controverses, Jean-François Courtine, directuer, Librairie Philosophique, J. Vrin, Paris, 1998, 309-329.

(h) "Truth Definitions, Skolem Functions and Axiomatic Set Theory," *Bulletin of Symbolic Logic* **4**, (1998), 303-337.

In defining truth for a language, quantifiers ranging over numbers as numbers must be independent of quantifiers over numbers as Gödel numbers. Such independence is not expressible in ordinary logic, only in independence-friendly logic, where the truth of a sentence equals the existence of its Skolem functions. A predicate asserting such existence can be formulated in an axiomatic set theory AX. Since it cannot be a truth predicate, there is in any model of AX some true sentence S whose Skolem functions do not all exist. This is paradoxical, for Skolem functions produce the "witness individuals" that guarantee the truth of S.

(i) "What is Abduction? The Fundamental Problem of Contemporary Epistemology", *Transactions of the Charles Peirce Society* **34**, (1998), 503-533.

Peirce distinguished between deduction, abduction and induction. Abduction is an inferential process. Only in abduction are new hypotheses introduced into inquiry. Although abductions are inferences, they are virtually identified by Peirce with conjectures. It is argued that abductions cannot be identified with "inferences to the best explanation." Furthermore, the requirement of rationality implies that abductions can always be construed as answers to the inquirer's (explicit or tacit) questions. This vindicates Peirce's claims; for instance, it is natural to call abductions inferences, for the strategic principles of abduction are virtually identical with the strategic principles of deduction.

(j)(with Ilpo Halonen) "Epistemic Logic", in *Routledge Encyclopedia of Philosophy* **3**, Peter Klein and R. Foley, editors, Routledge, London, 1998, 354-359.

(k) (with Arto Mutanen) "An Alternative Concept of Computability", in Jaakko Hintikka, *Language, Truth and Logic in Mathematics, Selected Papers* **III***,* Kluwer Academic Publishers, Dordrecht, 1998, 174-188.

(l)(with Gabriel Sandu) "Quantifiers", in *Routledge Encyclopedia of Philosophy* **7**, Peter Klein and R. Foley, editors, Routledge, London, 1998, 870-873.

(m) (with Gabriel Sandu) "Tarski's Guilty Secret: Compositionality" in *Alfred Tarski and the Vienna Circle*, *Vienna Circle Institute Yearbook* **6**, Jan Wolenski and Eckehart Köhler, editors, Kluwer Academic Publishers, Dordrecht, 1999, 217-230.

Tarski claimed that the concept of truth cannot be used coherently in colloquial language. This claim he initially presented as a consequence of Lesniewski's ideas, not of his own impossibility theorem. Lesniewski's ideas turn on compositionality, whose failure in natural language is therefore Tarski's real reason for his incoherence claim. Accordingly, when compositionality is given up, as in the authors' independence-friendly logic, a truth-predicate becomes definable in the same language. Compositionality therefore should not be a desideratum in a semantical theory.

1999

Books

(a) *Inquiry As Inquiry: Toward a Logic of Scientific Discovery, Selected Papers* **V**, Kluwer Academic Publishers, Dordrecht, 1999, xiii+289 pp.

Papers

(a) "The Emperor's New Intuitions", *The Journal of Philosophy* **96**, (1999), 127-147.

The current practice of appealing to intuitions in philosophical argumentation originated as an imitation of (what was taken to be) Chomsky's methodology in linguistics. Unlike Chomsky, who is a Cartesian, contemporary philosophers do not have a satisfactory theoretical rationale for their appeals to intuitions. They also fail to realize the limitations of their own intuitions and the possibility of interpreting them in different ways. As an example, Kripke's intuitions on which his "new theory of reference" is based are analyzed and criticized. A constructive perspective on intuitions is to think of them as answers one obtains by the usual methods of ascertaining a person's conceptual assumptions, applied to one's own case. In order to be useful, intuitions must also be admitted to have at least implicit generality.

(b) "Is the Axiom of Choice a Logical or Set-Theoretical Principle?", *Dialectica* **53**, (1999), 283-29.

A generalization of the axioms of choice says that all the Skolem functions of a true first-order sentence exist. This generalization can be implemented on the first-order level by generalizing the rule of existential instantiation into a rule of functional instantiation. If this generalization is carried out in first-order axiomatic set theory (FAST), it is seen that in any model of FAST, there are sentences S which are true but whose Skolem functions do not exist. Since this existence is what the truth of S means in a combinational (model-theoretical) sense, in any model of FAST there are sentences which are set-theoretical "true" but false in the normal sense of the word. This shows that the assumptions on which the axiom of choice rests cannot be fully implemented in FAST. The axiom of choice is not a set-theoretical principle.

(c) "Is Logic the Key to all Good Reasoning?", in Jaakko Hintikka, *Inquiry as Inquiry: A*

Logic of Scientific Discovery: Selected Papers **V**, Kluwer Academic Publishers,

Dordrecht, 1999, 1-24.

(d) "On Aristotle's Notion of Existence", *The Review of Metaphysics* **52**, (1999), 779-805.

(e) "Quine's Ultimate Presuppositions", *Theoria* **65**, (1999), 3-24.

(f) (with Ilpo Halonen) "Interpolation as Explanation", *Philosophy of Science* **66**, (1999), 414-423.

A (normalized) interpolant I in Craig's theorem is a kind of explanation why the consequence relation (from F to G) holds. This is because I is a summary of the interaction of the configurations specified by F and G, respectively, that shows how G follows from F. – If explaining E means deriving it from a background theory T plus situational information A and if among the concepts of E we can separate those occurring only in T or only in A, then the interpolation theorem applies in two different ways yielding two different explanations and two different covering laws.

(g) (with Ilpo Halonen) "Unification — It's Magnificent But is it Explanation?", in *Proceedings of the Lund Conference on Explanation,* J. Persson, editor, *Synthese* **120**,

no. 1, (1999), 27-47.

(h) (with Ilpo Halonen and Arto Mutanen) "Interrogative Logic as a General Theory of Reasoning", in Jaakko Hintikka, *Inquiry as Inquiry: A Logic of Scientific*

Discovery, Selected Papers V, Kluwer Academic Publishers, Dordrecht, 1999, 47-90.

2000

Books

(a) *On Gödel, Wadsworth Philosophers Series,* Wadsworth/Thomson Learning, Belmont,

CA, 2000, i-ii + 74 pp.

(b) *On Wittgenstein, Wadsworth Philosophers Series,* Wadsworth/Thomson Learning,

Belmont, CA, 2000, i-ii + 65 pp.

Papers

(a) "Epistemology: Introduction" in *The Examined Life: Readings from Western Philosophy from Plato to Kant*, Stanley Rosen, editor, Random House, New York, 2000, 401-414.

(b) "Gadamer: Squaring the Hermeneutical Circle" *Revue de Internationale de Philosophie* **54,** (2000), 487-497.

(c) "Game-Theoretical Semantics as a Challenge to Proof Theory", *Nordic Journal of Philosophical Logic* **4**, (2000), 127-141.

(d) "History of Logic Before and After Bochenski", in *Joseph (I.M.) Bochenski: Life and Work*, J. Kozak and G. Küng, editors, Verlag A. Stanic Scientific Publishers, 2000.

(e) "Intuitions as Model-theoretical Insights", in *Intuitive Formation of Meaning: Symposium Held in Stockholm, April 20-21,1998,* Sven Sandström, editor, *Konferenser* **48**, 2000, 75-90.

(f) "Knowledge Functions in the Growth of Mathematical Knowledge", in *The Growth of Mathematical Knowledge,* E. Grosholz and H. Berger, editors, Kluwer Academic Publishers, Dordrecht, 2000, 1-15.

(g) "Language as a "Mirror of Nature", *Sign Systems Studies* **28**, (2000), 62-72.

(h) "On the Educational Missions of Philosophy", *Diogenes* **48/4**, no. 192, (2000), 63-70.

(i)"Questioning as a Philosophical Method" in *The Examined Life: Readings from Western Philosophy from Plato to Kant*, Stanley Rosen, editor, Random House, New York, 2000, 453-470.

(j)"The Theory-Ladenness of Intuitions" in *Logique en perspective: Mélanges offerts àPaul Gochet*, François Beets and Éric Gillet, editors, Ouisia, Bruxelles, 2000, 259-287.

(k) "What is IF Logic and Why Do We Need It?", in Chinese translation by Chen Bo, *Journal of Dialectics of Nature* **22**, no. 3, (2000), 20-28.

(l)"What is True and False about So-Called Theories of Truth?", in *Analytic Philosophy and Logic*, Akihiro Kanamori, editor, *Proceedings of the Twentieth World Congress of Philosophy* **6**, 2000, 155-160.

(m) "Review: *Routledge Encyclopedia of Philosophy,* Edward Craig, General Editor, Routledge, London/New York, 1998, vols. 1-10", in *Synthese* **124,** no. 3, (2000), 433-445.

(n) (with Ilpo Halonen) "Aristotelian Explanations", *Studies in History and Philosophy of Science* **31**, (2000), 125-136.

2001

Books

(a) *Filosofian köyhyys ja rikkaus: Nykyfilosofian kartoitusta*, Janne Hiipakka and Risto
Vilkko, editors, Art House Oy, Helsinki, 2001, 400pp. (*The Poverty and Richness of
Philosophy: Perspectives on Contemporary Philosophy*)

Papers

(a) "Ernst Mach at the Crossroads of Twentieth-Century Philosophy" in *Future Pasts:*
Perspectives on the Analytic Tradition*, Juliet Floyd and Sanford Shieh, editors, Oxford
University Press, Oxford, March, 2001, 81-100.*

(b) "Introduction and Postscript: Defining Truth and its Difficulties", *Synthese* **126**, nos. 1-2,
(2001),1-16.

(c) "Intuitionistic Logic as Epistemic Logic", *Synthese* **127**, no. 1, (2001), 7-19.

(d) "Is Logic the Key to All Good Reasoning?", *Argumentation* **15**, (2001), 35-57.

(e) "Post-Tarskian Truth", *Synthese* **126**, no.1 (2001), 17-36.

Using Gödel numbering means speaking of numbers in two different roles, as numbers and as codifications of formulas of the same arithmetical language. If this is done, quantifiers ranging over numbers in the two roles must be informationally independent. This cannot be done in ordinary first-order logic, which explains why Tarski's impossibility theorem holds. It can be done in a suitable independence-friendly first-order language, which means that a self-applied truth-predicate can be defined in it.

This puts the entire theory of truth to a new light. It shows that the previous difficulties in trying to define truth explicitly are not due to the excessive strength of the languages in question, but to their poverty. It deprives such "theories of truth" as the so-called coherence theory much of their motivation. It shows that minimalist approaches to truth have a point, but need IF logic in order to be carried out explicitly.

(f) "What Is Truth? Stay for an Answer"in *What Is Truth*, Richard Schantz, editor, Walter de Gruyter, Berlin, 2002, 238-245.

(g) (with Ilpo Halonen) "Toward a Theory of the Process of Explanation" in Ilpo Halonen, *Interrogative Model of Explanation and Covering Laws* (dissertation), Department of Philosophy, University of Helsinki, Vantaa, 2001, 141-212. (Forthcoming in *Synthese*.)

(h) "The Proper Treatment of Quantifiers in Ordinary Logic", in *Collected Papers of Stig Kanger: With Essays on His Life and Work,* Vol. II, Ghita Holmström-Hintikka, Sten Lindström and Rysiek Sliwinski, editors, Kluwer Academic, Dordrecht, 2001, 87-95.

2002

Papers

(a) "Causes, Causes, Causes: Three Aspects of the Idea of Cause", in *Infinity, Causality and Determinism: Cosmological Enterprises and their Preconditions*, Eeva Martikainen, editor, Peter Lang, Frankfurt, 2002, 111-118.

(b) "Comment on Eklund and Kolak", *Synthese* **131**, no. 3, (2002), 389-393.

(c) "Die Dialektik in Gödels *Dialectica* Interpretation", in Bernd Buldt et al., editors, *Kurt Gödel: Wahrheit und Beweisbarkeit* **2**, öbv & hpt, Vienna, 2002, 67-90. (A corrected and expanded version in German translation of "Gödel's Functional Interpretation in a Wider Perspective", in *Yearbook 1991 of the Kurt Gödel Society Yearbook*, H.D. Schwabl, editor, Kurt Gödel Society, Vienna, 1993, 1-39.)

(d) "Hyperclassical Logic (a.k.a. IF Logic) and Its Implications for Logical Theory", *Bulletin of Symbolic Logic* **8**, (2002), 404-423.

(e) "Looginen empirismi kuusi vuosikymmentä myöhemmin" in *Wienin piiri,* Ilkka Niiniluoto and Heikki J. Koskinen, editors, Gaudeamus, Helsinki, 2002, 250-260.

(f) "Negation in Logic and in Natural Language", *Linguistics and Philosophy* **25**, (2002), 585-600.

(g) "Quantum Logic as a Fragment of Independence-Friendly Logic", *Journal of Philosophical Logic* **31**, (2002), 197-209.

(h) (with Anna-Maija Hintikka) "Wittgenstein the Bewitched Writer" in Rudolf Haller and Klaus Puhl, editors, *Wittgenstein and the Future of Philosophy: A Reassessment after 50 Years. Proceedings of the 24th International Wittgenstein Symposium,* öbv & hpt, Wien, 2002, 131-150.

2003

Books

(a) (edited with T. Czarnecki, K. Kijania-Placek and A. Rogszczak) *Philosophy and Logic: In Search of the Polish Tradition:Essays in Honour of Jan*

Woleński on the Occasion of his 60th Birthday, Synthese Library **323**, Kluwer Academic Publishers, Dordrecht, 2003, i-xiii+290pp.

Papers

(a) "A Distinction Too Few or Too Many? A Vindication of the Analytic vs. Synthetic Distinction", in *Constructivism and Practice: Toward a Historical Epistemology,* Carol C. Gould, editor, Roman & Littlefield, Lanham, Maryland, 2003, 47-74.

(b) "The Notion of Intuition in Husserl", *Review Internationale de Philosophie*, no.224, (2003), 169-191.

(c) "On the Epistemology of Game-Theoretical Semantics", in *Philosophy and Logic: In Search of the Polish Tradition:Essays in Honour of Jan Woleński on the Occasion of his 60th Birthday,* Synthese Library **323**, J. Hintikka, T. Czarnecki, K. Kijania-Placek and A. Rogszczak, editors, Kluwer Academic Publishers, Dordrecht, 2003, 57-66.

(d) "On Tarski's Assumptions", *Synthese*, (2003), 1-17.

(e) "A Second Generation Epistemic Logic and its General Significance", in *Knowledge Contributers,* Synthese Library, **322**, Vincent F. Hendricks, Klaus Frovin Jørgensen and Stig Andur Pedersen, editors, Kluwer Academic Publishers, Dordrecht, 2003, 33-56.

(f) "Squaring the Vienna Circle with Up-to-date Logic and Epistemology", in *Language, Truth and Knowledge: Contributions to the Philosophy of Rudolf Carnap*, Thomas Bonk, editor, Kluwer Academic Publishers, Dordrecht, 2003, 149-166.

(g) "What Does the Wittgensteinian Inexpressible Express?", *The Harvard Review of Philosophy* **XI**, (2003), 9-17.

(h) (with John Symons) "Systems of Visual Identification and Neuroscience: Lessons from Epistemic Logic", *Philosophy of Science* **70**, (2003), 89-104.

Shows how developments in epistemic logic can play a nontrivial role in cognitive neuroscience. Argues that the striking correspondence between two modes of identification, as distinguished in the epistemic context, and two cognitive systems distinguished by neuroscientific investigation of the visual system (the "where" and "what" systems) is not coincidental, and that it can play a clarificatory role at the most fundamental levels of neuroscientific theory.

2004

Books

(a) *Analyses of Aristotle, Selected Papers* **VI**, Kluwer Academic Publishers, Dordrecht, 2004, I-xii+238pp.

Papers

(a) "Aristotle's Theory of Thinking and Its Consequences for his Methodology", in Jaakko Hintikka, *Analyses of Aristotle: Selected Papers* **VI**, Kluwer Academic Publishers, Dordrecht, 2004, 45-85.

(b) "Did Wittgenstein Follow the Rules? (Or Was He Guided by Them?) in *Experience and Analysis: Contributions of the Austrian Ludwig Wittgenstein Society*, 27th International Wittgenstein Symposium, Elisabeth Leinfellner, Rudolf Haller, Werner Leinfellner, Klaus Puhl and Paul Weingartner, eds., Austrian Ludwig Wittgenstein Society, Kirchberg am Wechsel, 2004, 140-141.

(c) "A Fallcious Fallacy?", *Synthese* **140**, (2004), 25-35.

(d) "Independence-friendly Logic and Axiomatic Set Theory", *Annals of Pure and Applied Logic* **126**, (2004), 313-333.

(e) "Hintikka, Merrill Bristow", in *Dictionary of Modern American Philosophers,* Thoemmes Press, Bristol, 2004.

(f) "On the Different Identities of Identity: A Historical and Critical Essay" in *Language, Meaning, Interpretation,* Guttorm Fløistad, editor, Kluwer Academic Publishers, Dordrecht, 2004, 117-139.

(g) "What Does the Wittgensteinian Inexpressible Express?" *The Harvard Review of Philosophy* **11**, (2003), 9-17.

(h) "What Is the True Algebra of Logic?", in *First-Order Logic Revisited,* Vincent Hendricks et al., editors, Logos Verlag, Berlin, 2004, 117-128.

(i)"Wittgenstein's Demon and His Theory of Mathematics", in *Essays on Wittgenstein and Austrian Philosophy:In Honour of J.C. Nyiri,* Tamás Demeter, editor, Rodopi b.v., Amsterdam - New York, 2004, 89-107.

Forthcoming

Papers

(a) "The Crash of the Philosophy of the *Tractatus:* Wittgenstein's Change of Mind in 1929".

(b) "Epistemology without Knowledge and without Belief".

The nature of epistemology is revealed by two questions: What is it that we are doing in epistemological inquiry? and, What can the product of such an inquiry do for us? The concrete function of the notion of knowledge is to indicate what information we are entitled to act on. What we are doing in epistemological inquiry is shown by the interrogative model of inquiry that has recently been developed. A survey of this model reveals that neither the concept of knowledge nor that of belief are needed in it. Instead of knowledge we are dealing with information, and instead of belief we are dealing with acceptance. The notion knowledge enters only through the question whether the output of inquiry entitles us to act on it. The answer to this question — this is the applicability of the notion of knowledge — depends on the subject matter and not only on the structure of the inquiry, and hence does not belong to general epistemology. Likewise, belief should be construed as a product of inquiry, not as free choice of propositions to accept. As such, it does not belong to general epistemology any more than the concept of knowledge.

(c) "The Indispensability of Mathematics and the A Priori Element in Experimental Science", forthcoming.

Mathematics is indispensable in science because mathematical knowledge is needed for the purpose of answering adequately scientific questions, especially experimental ones. In such a question, the inquirer tries to find out how a variable (say, y) depends on another one (say, x). The desideratum of such a question is therefore of the form

(1) $K_I(\forall x)(\exists y/K_I)\, S[x,y]$

in words, "I know which value of y is related to each value of x as in S[x,y]." An experiment provides ideally a function-in-extension, in other words, enables the inquirer to assert

(2) $K_I(\forall x)\, S[x,g(x)]$

for some function g defined by a class of observations, that is, by a class of pairs of correlated argument-values and function-values. But (2) does not yet satisfy the questioner, i.e. does not entail (1). It does so only in conjunction with the further premise

(3) $K_I(\exists f/K_I)(\forall x)(g(x)=f(x))$

which can be written as

(4) $K_I(\exists f/K_I)(g=f)$

and which says that the inquirer ("myself") knows which function g is. This is an eminently natural requirement, which also follows from epistemic logic. What it means is that according to the logic of questions and answers the knowledge expressed by (3) is indispensable for an adequate answer to the original experimental question. But the knowledge expressed by (3) is mathematical, not factual. Hence certain mathematical knowledge is needed for the purpose of answering experimental questions; mathematics is indispensable in science.

This argument is a straightforward application of the basic logic of questions and answers. It is independent of the role (if any) of mathematics as a means of interconnecting and systematizing scientific propositions.

As a corollary, a new discussion of the problem of induction comes to light. Even if an experimentalist can establish an entire function-in-extension (graph of the function) g, as in (2), the inquirer does not fully know how y depends on x unless and until he or she has figured out what this function is, mathematically speaking.

(d) "Omitting data — ethical or strategic problem?", *Synthese*.

(e) "On Argumentation in a Multicultural Setting" in *Proceedings of the New Delhi Meeting of the IIP*, KluwerAcademic Publishers.

(f) “Presuppositions of Questions — Presuppositions of Inquiry”, *Proceedings of the 2001 IIP Annual Meeting*, Kluwer Academic Publishers, Dordrecht.

(g) “Truth, Negation and Other Basic Notions of Logic”

(h) “Who Has Kidnapped the Notion of Information?”

(i)“Will the Real Ludwig Wittgenstein Please Stand Up”.

(j)“Wittgenstein on Knowledge and Skepticism”.

(k) (with Risto Vilkko) “The Concept of Existence in Aristotle and Frege”

Index