Epistemic Logic

A SURVEY OF THE LOGIC OF KNOWLEDGE

Nicholas Rescher

Epistemic Logic

A Survey of the Logic of Knowledge

Nicholas Rescher

UNIVERSITY OF PITTSBURGH PRESS

Published by the University of Pittsburgh Press, Pittsburgh, Pa., 15260

Manufactured in the United States of America
Printed on acid-free paper
10 9 8 7 6 5 4 3 2 1

Library of Congress Cataloging-in-Publication Data

Rescher, Nicholas.
Epistemic logic : a survey of the logic of knowledge / Nicholas Rescher.
p. cm.
Includes bibliographical references and index.
ISBN 0-8229-4246-1 (alk. paper)
1. Epistemics. 2. Logic. 3. Knowledge, Theory of. I. Title.
BC21.E64R47 2005
160—dc22
2004021645

Preface

Not long ago I was invited to contribute a sketch of epistemic logic offering a general survey of philosophical logic to *A Companion to Philosophical Logic* (Oxford University Press). When preparing this piece, it struck me that there actually is no synoptic account of the field. Despite a great many articles in scholarly journals dealing with special topics in epistemic logic, there does not exist a text that puts the pieces together in a systematic exposition of the field that connects it with the issues of philosophical epistemology. My aim here is to fill this gap. The topic is of sufficiently wide interest—not only to logicians but also to philosophers, cognitive scientists, information scientists, artificial intelligence researchers, and others—that students need and deserve a helping hand in securing convenient access to this domain.

I am grateful for the helpfulness of various individuals: the students in my Epistemology course in the fall of 2001—especially, Dane Roberts—for the opportunity their questions have afforded me to work out this presentation of the relevant ideas; my Pittsburgh colleagues Joseph Camp and Anil Gupta for numerous helpful suggestions; and to my assistant, Estelle Burris, for her patience and competence.

Contents

1

Setting the Stage

The Idea of Epistemic Logic

Epistemic logic is that branch of philosophical logic that seeks to formalize the logic of discourse about knowledge. Its object is to articulate and clarify the general principles of reasoning about claims to and attributions of knowledge—to elucidate their inferential implications and consequences. In pursuing this goal, it deals principally with propositional knowledge (along the lines of "Smith knows *that* coal is black") and secondarily also with interrogative knowledge (along the lines of "Jones knows *where* the treasure is buried and *who* put it there").[1] It is the object of this book to give an overview of the discipline by setting out in a formalized manner the general principles for reasoning about such matters.

The History of Epistemic Logic

Epistemic logic is a product of the second half of the twentieth century. After the preliminary work of Rudolf Carnap's deliberations about belief sentences,[2] epistemic logic was launched in an important 1948 paper by the Polish logician Jerzy Łoś.[3] Łoś developed what he called a logic of "belief" or "acceptance" based on an operator *Lxp* for "the individual *x* believes (or is committed to) the proposition *p*," for which he stipulated axiomatic rules substantially akin to those to be specified for the knowledge operator to be introduced below. A spate of publication during the 1950s by such logicians as Alonzo Church, Arthur Prior, Hilary Putnam, G. H. von Wright, and the present writer further extended the range of relevant deliberations. During the 1960s various authors carried matters forward, with the first book on the topic (by Jaakko Hintikka) appearing in 1962. Since then there has been a small but steady stream of work in the field. (For details, see the bibliography.)

Fundamentals of Notation

The present treatment of epistemic logic undertakes the construction of a deductively formalized system **s** that is adequate for this purpose. As is often the case with axiomatic treatments, the present discussion illustrates how a modest set of basic assumptions provides a great deal of instructive information about the conceptual anatomy of the idea at issue.

Use will be made here of the familiar resources of propositional and quantificational logic supplemented by the machinery of quantified modal logic. All of the following symbols will accordingly be used in the standard way:

$\sim$, &, $\vee$, $\supset$, and $\equiv$ for the familiar propositional connectives
$\forall$ and $\exists$ for universal and existential quantification
$\Box$ and $\Diamond$ for the modalities of necessity and possibility

Additionally, the following notational conventions will be employed:

Kxp ("*x* knows that *p*")
K^*xp ("*p* is derivable from propositions that *x* knows")
x, *y*, *z*, . . . as variables for knowers: intelligent individual (or possibly groups thereof)
p, *q*, *r*, . . . as variables for propositions (contentions to the effect "that such-and-such is the case")
t, *t'*, *t"*, . . . as variables for specifically true propositions (note that

$(\forall t)F(t)$ amounts to $(\forall p)(p \supset F(p))$ and $(\exists t)F(t)$ amounts to $(\exists p)(p \,\&\, F(p))$, F being an arbitrary propositional function

$u, u', u'', \ldots$ as variables for objects of consideration or discussion

$F, G, H, \ldots$ as variables for properties or features objects or of propositions

$S, S', S'', \ldots$ as variables for sets of objects or propositions

$Q, Q', Q'', \ldots$ as variables for questions

The quantificational logic at work here is type differentiated: *p, q, r,* and so on stand for propositions *x, y, z,* and so on for knowers and so on. One could, in theory, employ a single class of variable α, β, γ and so on, and then render

$(\forall x)Fx$ as $(\forall \alpha)(\alpha \in K \supset F\alpha)$,
where K represents the set of knowers,

$(\exists p)Gp$ as $(\exists \alpha)(\alpha \in P \,\&\, G\alpha)$,
where P represents the set of propositions,

and the like. But this more elaborate style of presentation would make our formulations needlessly complicated and less easily read. It should be noted that the various domains at issue (knowers, propositions, truths, and the like) are all nonempty, so that the inference from "all" to "some" is appropriate in all cases.

Certain special symbols will be employed as follows:

$\vdash p$ for "p is a thesis of our system (**s**)"

$\Vdash p$ for $\vdash (\forall x)Kxp$

$p \vdash q$ for $\vdash p \supset q$

$p \Vdash q$ for $\vdash (\forall x)Kxp \supset Kxq)$

$p @ Q$ for "p answers the question Q"

The symbol $\vdash$ will also be called upon to serve as an index of entailment through the following equivalence:

$p \vdash q$ iff $\vdash (p \supset q)$

Since the antecedent p may disaggregate into the conjunction of a series of propositions, $p_1, p_2, \ldots p_n$, this stipulation renders our system subject to what is standardly called the deduction theorem, on the basis of the following equivalence:

$p_1, p_2, \ldots, p_n \vdash q$ if and only if $p_1, p_2, \ldots, p_{n-1} \vdash p_n \supset q$.

With propositional knowledge of matters of fact, the basic unit of assertion will be a statement of the form "*x* knows that *p*" (*Kxp*). Such propositional knowledge is a matter of a relationship—a *cognitive* relationship—between a person and a true proposition. And just as for an otherwise unidentified individual *x* one can uniformly substitute the name of any individual, so for an otherwise unidentified proposition *p* one can uniformly substitute any other. The use of variables thus affords a gateway to generality by providing for substitution. For example, since it obtains as a general principle that

If *Kxp*, then *p*,

one automatically secures a vast range of such other assertions as

If *Ky*(*p* & *q*), then *p* & *q*,

which results from the preceding via the substitutions *y*/*x* and (*p* & *q*)/ *p*.

Recourse to symbolic representation enables us to achieve greater precision. For instance, in ordinary language "*x* does not know that *p*" is equivocal as between $\sim Kxp$ and p & $\sim Kxp$, which would be more accurately formulated as "*p*, and *x* does not know it."

Theses of the System

As already mentioned, $\vdash$ here serves as an assertion symbol indicating that what follows qualifies as a general principle of the system of epistemic logic (**s**) that is under construction. By convention its employment conveys implicit universality for any free variables. Thus,

$\vdash Kxp \supset p$

asserts that $(\forall x)(\forall p)(Kxp \supset p)$ holds in our system. A proposition that qualifies as a thesis of the system should be seen as being true on logico-conceptual grounds alone. Its validation will rest entirely on the specification of the terms of reference that are employed and thus on the conventions of meaning and usage that are being adopted. These theses accordingly serve to specify the conception of *knowledge* that is to be at issue. And since a "logic" of knowledge must deal in general principles, it is the establishment or refutation of such conceptually grounded generalizations that concern us at present. What is at work here is in fact a somewhat delicate reciprocal feedback process. A certain particular conception of knowledge guides the construction of our epistemic system. And the theses of this system define and precisify the particular conception of knowledge that is at issue.

In dealing with knowledge and its "logic" we are not, of course, functioning in a realm of total abstraction, as would be the case with "pure" (rather than applied) mathematical or theoretical logic. Instead, we are dealing with the resources of intelligent beings (not necessarily members of *Homo sapiens*) operating substantially within the limits imposed by the realities of this world of ours. Accordingly, the "facts of life" that reflect the cognitive situation of such beings and the conditions that define their situation in this world represent the ultimately factual (rather than purely theoretical) circumstances that a logic of knowledge as such will have to reflect. In particular, knowers have to be construed as finite beings with finite capacities, even though reality, nature, has an effectively infinite cognitive depth in point of detail, in that no matter how elaborate our characterizations of the real, there is always more to be said.[4] The reality of it is that epistemic logic is an applied logic and its theses, being geared to salient feature of the established concept of knowledge, stand correlative to the ways in which we actually do talk and think about the matter.

Propositions as Objects of Knowledge

There is nothing problematic about saying "*p*, but *x* does not know (or believe) it." But in the special case of $x =$ oneself (the assertor), this otherwise viable locution is impracticable. This discrepant state of affairs has become known as "Moore's paradox" after G. E. Moore, who first puzzled over it.[5] Of course, there would be nothing amiss about saying "I surmise (conjecture, suspect) that *p* but do not actually know (or confidently believe) it." But in making a flat-out, unqualified statement we stand subject to the ground rule that this purports knowing the truth of the matter, so that in going on to add "but I do not know (or believe) it" to an assertion of ours, we take the inconsistent line of giving with one hand what we take away with the other. Our categorical (that is, unqualified) assertions stand subject to an implicit claim to truth and knowledge, and we thus authorize the inference from asserting p both to Kip and to p itself. Accordingly, when our system **s** is held to make an explicit assertion, this will be something that we ourselves purport to know, so that we then have it that $\vdash p$ entails $(\exists x)Kxp$.

In general, claims to knowledge regarding individual objects or collections thereof can be reformulated with the machinery of propositional knowledge by means of quantification. Thus, consider

"x knows the identity of Jack the Ripper":

$(\exists p)$ (p identifies who Jack the Ripper was & Kxp)

"x knows the major features of London's topography":

$(\forall p)$ (p states a major feature of London's topography $\supset Kxp$)[6]

Such statements about someone's knowledge of individual objects can be reduced to propositional knowledge by employing either

$$(\exists p)(p \mathbin{@} Q \mathbin{\&} Kxp)$$

or

$$(\forall p)(p \mathbin{@} Q \supset Kxp)$$

when p @ Q abbreviates "p answers the question Q."

By and large, propositional knowledge represents a resource by whose means the other principal versions of the concept of knowledge can be recast and represented. However, some knowledge is not propositionally reducible, specifically, know-how of a certain sort. For we have to distinguish between

performatory know-how: x knows how to do A in the sense that x can do A; and

procedural know-how: x knows how A is done in the sense that x can spell out instructions for doing A.

The second sort of know-how is clearly a matter of propositional knowledge—that *x* knows that *A* can be done by doing such-and-such things; for example, *x* knows that people swim by moving their arms and legs in a certain cycle of rhythmic motions. But, of course, *x* can know how *A* is done without being able to do *A*—that is, without *x* having the performatory skills that enable *x* to do *A*. (For example, *x* may know *that* a certain result is produced when a text is translated from one language to another without actually knowing *how* to make such a translation.) And, therefore, while propositional reduction is practicable with respect to *procedural* know-how, such a reduction will not be practicable with respect to *performatory* know-how, seeing that people are clearly able to do all sorts of things (catch balls, remember faces) without being able to spell out a process or procedure for doing so.[7]

All the same, the different modes of knowledge are inextricably interconnected. To know (propositionally) *that* a cat is on the mat one must know (adverbially) *what* a cat is. And this knowledge rests on knowing how to tell cats from kangaroos.

2

Basic Principles

Acceptance and Assertion

The distinction between four modes of propositional acceptance/assertion will be serviceable in characterizing the present system of epistemic logic (**s**):

Type 1: $\Vdash p$, that is, iff $\vdash (\forall x)Kxp$. This represents "obvious knowledge" coordinate with acceptance by our epistemic system (**s**) of *p* as something *universally recognized* among the knowers at issue.

Type 2: $\vdash p$, that is, iff *Ksp*. On this basis we propose to accept *p* as part of our epistemic system (**s**), itself now regarded as a "knower" of sorts.[1] Such acceptance represents "patent knowledge" coordinate with acceptance as certain on epistemico-logical grounds. (It, accordingly, includes the theses of both epistemic and standard logic.)

Type 3: *p*, that is, *Kip*, where i = oneself. Note that, as already discussed in chapter 1, when one makes a flat-out assertion of *p*, one claims to know it, so that asserting *p* comes to claiming *Kip*.

Type 4: $\ast p$, that is, a qualified assertion of *p*. This represents "merely accepted belief" coordinate with a tentative or provisional acceptance as true, conceivably on the basis of conjecture rather than actual knowledge.

In the fourth case we do indeed regard the proposition in question as acceptable but without claiming assured *knowledge* of the matter. A type-4 assertion has to be seen as a truth that is not actually known but only (and perhaps hesitatingly) endorsed. Such theses may be surmised or presumed, regarded as plausibly putative truths—as was the case with the thesis "There are mountains on the far side of the moon" in the cognitive state of the art of the nineteenth century. What is at issue is a tentatively adopted mere belief that, as such, contrasts decidedly with actual knowledge. In claiming p in the mode of $\ast$-assertability we do no more than to characterize it as a putative truth, rather than as one that is deemed to be certifiable as such. (Such theses will not form part of one's purported knowledge, let alone of the assertions of **s** itself.) The existence of the fourth level of assertion is a reminder that epistemology is broader than the theory of *knowledge.* For matters of presumption, conjecture, reasonable belief, and plausible assertability also clearly fall within its purview.

On this basis, then, all four of these modes of acceptance do indeed convey a commitment—an *assertion.* A claim that p is the case obtains in every instance, but with different epistemic modalities—conjecture, plausible supposition, or the like. Such tentative endorsements say that something is true, all right, but in a substantially less firm and confident tone of voice.

It deserves noting that one can maintain that

$$\ast(p \;\&\; {\sim}(\exists x)Kxp)$$

without self-contradiction. The claim "p is presumably true, although no one actually knows it (for sure)" is viable. But, of course,

$$Ki(p \;\&\; {\sim}(\exists x)Kxp)$$

is self-contradicting in claiming both that one knows p and that nobody does. Accordingly,

$$p \;\&\; {\sim}(\exists x)Kxp$$

can be maintained as a type-3 assertion but not as a type-2 one (let alone a type 1 one).

Some Fundamental Principles

Four principles will play a leading role in the system **s** that we are engaged in elaborating:

1. *Knower capacity:* We are dealing with actual knowers, individuals who know at least *something:*
   $\vdash (\forall x)(\exists p)Kxp$
2. *Knower limitation:* We are dealing with knowers of limited capacity, none of whom are omniscient:
   $\vdash (\forall x)(\exists p)(p \,\&\, {\sim}Kxp)$

Thus, a godlike being who knows everything is left out of the range of present concern. This does not, of course, mean that omniscience itself is a logical impossibility. It is merely excluded as a viable prospect for the particular conception of knowledge and knowers that is at issue in the present context of deliberation.

3. *Veracity:* Whatever is actually known by someone will have to be the case. Where we credit knowledge we must credit truth as well: there is no such thing as a *knowledge* of falsehoods; that sort of thing would have to be classed as merely *putative* knowledge.
   $\vdash Kxp \supset p$

Actual knowledge, in short, must be veridical.

4. *Conjunctivity:* Knowers have the (rather minimal) competence of being able to "put two and two together." When a knower knows two facts separately, the knower knows them conjunctively. If Kxp and Kxq, then $Kx(p \,\&\, q)$:
   $\vdash (Kxp \,\&\, Kxq) \supset Kx(p \,\&\, q)$

Thesis (4) is controversial. What renders it problematic is that its analogue clearly fails for merely probable (that is, less than certain) propositions: a conjunction of probable truths need not itself be probable.[2] But, of course, knowledge should be certain. And in any case the present discussion construes "knowing" in a generous sense of potential accessibility that goes somewhat beyond the realities of ordinary performance. The thesis at issue is thus less a consequence of theoretically geared principles than an artifact of the particular sense of *knowledge* at work in the present systematization of the concept.

Moreover, the converse of thesis 4 also obtains:

5. *Converse conjunctivity: Knowledge is disaggregative:* what is known conjunctively is also known separately:
   $\vdash Kx(p \,\&\, q) \supset (Kxp \,\&\, Kxq)$

Two important consequences follow from the preceding principles:

Knowledge is consistent: If Kxp and Kyq, then p is bound to be compatible with q.

This follows from *veracity* and *conjunctivity* in view of the overall consistency of the truth. For Kxp and Kyq yields $p \,\&\, q$. Accordingly,

If $Kxp \,\&\, Kxq$, then never $p \vdash {\sim}q$.

Falsehoods are unknowable. Knowledge requires conformity to fact, and falsehoods are not factual.

Accordingly, falsehoods cannot be known: If p is false, then no one can know that p: If ${\sim}p$, then ${\sim}(\exists x)Kxp$, or equivalently: If $(\exists x)Kxp$, then p. This can be demonstrated via thesis 3, *veracity*. An unknown truth is a real loss, but the unknowability of falsehoods is simply a matter of excluding error from the cognitive realm.

We must, of course, credit our knowers with the modest intelligence requisite for grasping certain elemental truths. Thus, the p-universalized thesis

$Kx(p \vee {\sim}p)$

is quite appropriate, unlike the p-universalized thesis

$Kxp \vee Kx{\sim}p$.

It is one thing to know *that* one of several propositions is true according to $K(p \vee q)$, but it is something quite different to know *which* of several propositions is true according to $Kxp \vee Kxq$.

One helpful way of looking at the matter is that these various principles are definitively constitutive of the sense of knowledge at issue here through serving to spell out its conceptual ramifications.

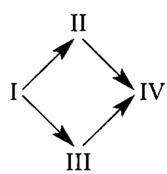
Some Theses about Knowers and Knowledge

Consider the following four theses about knowers:

(A)
I. $(\forall x)(\forall t)Kxt$
II. $(\forall x)(\exists t)Kxt$
III. $(\exists x)(\forall t)Kxt$
IV. $(\exists x)(\exists t)Kxt$

Here thesis IV, that somebody knows something, is, of course, at odds with radical skepticism.

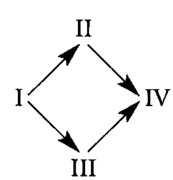
By the rules of quantificational logic we have



where the arrows represent entailment relationships. Since here we are concerned with imperfect knowers of limited capacity and prescind from an omniscient knower, we reject thesis III and hence thesis I in its wake. And since we are concerned with actual knowers, we accept thesis II and hence thesis IV in its wake. Accordingly, our stance toward the four theses at issue is $- + - +$.

It is informative also to consider the situation in regard to ignorance:

(B)
I. $(\forall x)(\forall t){\sim}Kxt$ ($\equiv$ not AIV)
II. $(\forall x)(\exists t){\sim}Kxt$ ($\equiv$ not AIII)
III. $(\exists x)(\forall t){\sim}Kxt$ ($\equiv$ not AII)
IV. $(\exists x)(\exists t){\sim}Kxt$ ($\equiv$ not AI)



Given our rejection of skepticism, we reject thesis I, and since our knowers are specifically *limited* knowers, we accept thesis II and thus