ROLF A. EBERLE

A LOGIC OF BELIEVING, KNOWING, AND INFERRING

It will be our aim to discuss and characterize an interpreted system which generates the necessary truths peculiar to certain concepts of believing, knowing, and inferring: ones which can be realistically ascribed to beings (such as humans or computers) which fall short of deductive infallibility and omniscience. The system will also provide for a distinction between *de re* and *de dicto* interpretations of epistemic contexts. In the pursuit of these goals, some new techniques will be required which might be of general interest in logic.

1. Informal discussion

In Hintikka's system [2], as well as in even the weakest among Lemmon's systems [7] of epistemic logic, we find counterparts of the inference rule

(1) $$\frac{\text{provably: } (\phi \rightarrow \psi)}{\text{provably: } (\mathrm{K}\phi \rightarrow \mathrm{K}\psi).}$$

If we read 'Kϕ' as 'it is known (by some unspecified knower) that ϕ', then the inference (1) expresses, roughly, that all logical consequences of things one knows are again things one knows. Given this reading, the idealized notion of knowledge which is being treated cannot be ascribed to beings whose deductive capacities are fallible or finite. A rule analogous to (1), but regarding beliefs instead of knowledge (as it is discussed in [5]), would be even less intuitive. The prevailing attitudes toward such inference rules seem to fall into two categories:

(A) There is the approach, taken by Hintikka in [2], of endorsing such rules while endowing the formal epistemic notions with new pre-systematic interpretations which seem to fit the rules. Apart from doubts whether this reinterpretation really corresponds to clear intuitions (see [5]), the objection remains that the systems fail to treat of the customary notions of belief and knowledge. And, while idealized epistemic notions may have their uses, epistemologists will also require a precise treatment of 'believ-

Synthese 26 (1974) 356–382. *All Rights Reserved*

ing' and 'knowing' in their customary senses. It is sometimes claimed (e.g. by Purtill in [9]) that this demand is inappropriate: we cannot hope, it is said, to develop a consistent epistemic system which takes account of the incoherent beliefs of the insane or the incomplete knowledge of infants. To the contrary, we shall aim at the construction of a consistent system which provides for total ignoramusses (ones who know nothing), complete idiots (ones who cannot draw even the most elementary inferences), and ultimate fools (ones who believe nothing but contradictions), without ceasing to be of interest to a logician. In this way, we hope to counter criticism of the sort 'it cannot be done!' by doing it.

(B) There is the approach, taken by Hintikka in [5], of restricting inference rules such as (1) to those logical consequences which, to the rational person, would in some sense seem 'obvious'. Whether or not an inference is obvious might be thought to depend on the length of the derivation, on the length of sentences one has to inspect in a derivation, on the logical complexity of these sentences (e.g. Hintikka's 'quantificational depth' in [5]), on the 'naturalness' of the deductive system one employs, on whether the derivation could be discovered by some mechanical procedure, and so forth. However if ones aim is that of characterizing something akin to the customary notion of knowledge, then this entire approach seems hopeless, no matter what factors in proofs one may take into account. For that aim requires a notion of 'obviousness' which is quite subjective and dependent upon the intelligence, training, prejudices, and degree of scepticism of the knower rather than upon a particular system of derivation with which the knower may not even be acquainted.

Instead of trying to explicate some notion of 'obviousness', is appears more fruitful to focus on those inferences (obvious or otherwise) which a person *actually carries out*. Accordingly, it looks promising to introduce, and suitably interpret, another epistemic operator 'I', where sentences of the form '$I(\phi, \psi)$' may be read 'from the assumption that ϕ one rationally infers that ψ'. Then the former inference rule (1) may be replaced by the postulate

$$(2) \qquad I(\phi, \psi) \rightarrow [K\phi \rightarrow K\psi]$$

which says, roughly, that everything which one actually infers from things one knows are again things one knows. Indeed, it seems reasonable to suppose that the body of ones knowledge is closed under those rational inferences which one actually makes.

The analogous postulate regarding beliefs

(3) $I(\phi, \psi) \to [B\phi \to B\psi]$

(where '$B\phi$' is read 'one believes that ϕ') seems equally plausible. For suppose that some person were to believe that ϕ, that it were claimed that he himself had inferred from ϕ that ψ, and that yet he failed to believe that ψ. Then, it seems to me, one would deny that the person had really made the alleged inference. True, he might say that he did, he might even have written down a sequence of lines which constitute a derivation of ψ from ϕ, but the presystematic notion of inferring requires more than such utterance or performance; namely, that a person's own inferences carry his beliefs into beliefs.

In addition, we shall postulate

(4) $I(\phi, \psi) \to K(\phi \to \psi)$.

In words: if a person rationally infers from ϕ that ψ, then he knows that if ϕ then ψ. By the consequent of (4) we do not mean to convey that the knower need understand the meaning of conditionals. Just as a person may know *that* snow is white without understanding the English words 'snow is white', so also a person may know that $(\phi \to \psi)$ without understanding the meaning of the arrow; and one way of knowing that $(\phi \to \psi)$ is by inferring that ψ from the assumption that ϕ. It is also tempting to read into (4) a temporal precedence: upon having first inferred from ϕ that ψ, a person subsequently knows that $(\phi \to \psi)$. In our tenseless system, this reading is illegitimate. Rather, we regard it as analytic of the concepts of knowledge and of rational inference that a person who draws an inference is thereby satisfying the conditions which warrant the corresponding knowledge claim.

The converse of (4) will not turn out to be valid. That is to say, we shall allow for the logical possibility that a person has non-inferential knowledge of the fact that if ϕ then ψ. Generally, non-inferential knowledge of any consistent sort should be logically possible. If the converse of (4) were adopted, then the epistemic primitive 'I' would become superfluous, and we could replace (say) the postulate (2) by the principle

(5) $K(\phi \to \psi) \to [K\phi \to K\psi]$.

However, (5) should not be treated as valid. For (as Professor K. Lehrer has pointed out to me) even though a person who satisfies the antecedents of (5) would be entitled to infer that ψ, he cannot be said to know that ψ until he has actually drawn the inference to which he is entitled.

Especially due to the principle (4), we regard the inferences expressed by 'I' as *rational* (though not necessarily as deductively valid) ones. For, if one knows that if ϕ then ψ, then one must have some rational grounds which justify (not necessarily deductively) the inference of ψ from ϕ.

Among principles which connect sentences of the form '$I(\phi, \psi)$' with one another, we list the following candidates for consideration:

(6) $[I(\phi, \psi) \ \& \ I(\psi, \chi)] \rightarrow I(\phi, \chi)$

Roughly: If one infers from ϕ that ψ and one infers from ψ that χ, then one infers from ϕ that χ.

(7) $I(\phi \ \& \ \psi, \chi) \rightarrow I(\phi, \psi \rightarrow \chi)$

Roughly: If one infers from the conjunction 'ϕ & ψ' that χ, then one infers from ϕ that if ψ then χ.

(8) $[I(\phi, \psi) \ \& \ I(\chi, \theta)] \rightarrow I(\phi \ \& \ \chi, \psi \ \& \ \theta)$

Roughly: If one infers from ϕ that ψ and infers from χ that θ, then one infers from the conjunction 'ϕ & χ' that ψ & θ.

(9) $[I(\phi, \psi) \ \& \ I(\phi, \psi \rightarrow \chi)] \rightarrow I(\phi, \chi)$

Roughly: If one infers from ϕ that ψ and infers from ϕ that if ψ then χ, then one infers from ϕ that χ.

(10) $I(\phi \ \& \ \psi, \phi)$

Roughly: From every conjunction one infers the first conjunct.

In appraising the plausibility of these principles, we ask ourselves whether the very meaning of 'inferring' compels us to accept them; and, where the principles have conditional form, we ask whether a person who has drawn the inference mentioned in the antecedent has *thereby* made the inference mentioned in the consequent. By this test, the proposals (6)–(10) seem to be listed in an order of decreasing plausibility. Accordingly, we shall adopt as a postulate only the principle (6) which asserts the transitivity of inferring. Nevertheless, there are surely contexts in which a logic with somewhat stronger principles would seem both desirable and warranted. This can happen, for example, if we want to apply the theory of inferences to machines which have been programmed to carry out certain deductive steps. In order to allow for this possibility, we shall characterize both the sematics and the deductive apparatus of our system in such a way that it will be immediately obvious how either can be strengthened as much as we please.

It is rather generally agreed that epistemic contexts are non-extensional. However, it is sometimes argued (e.g. in Montague [8]) that such contexts can be adequately treated in the framework of some intensional logic. Yet, the customary notions of knowledge and especially of belief clearly make for contexts which are non-intensional as well, in the sense that they do not admit of the valid interchange even of necessarily equivalents; that is, of terms and formulas which express the same intension. Any person who falls short of logical omniscience will be ignorant of some equivalents of things he knows and will fail to believe some equivalents of his beliefs. Since one's inferences carry beliefs into beliefs, the interchange of equivalents should fail as well in contexts which express one's inferences. Thus, if it should happen that

(11) $\quad \mathrm{I}(\phi, \psi)$,

while ϕ and ψ are respectively equivalent to ϕ' and ψ', one is not entitled to conclude by interchange that either

(12) $\quad \mathrm{I}(\phi', \psi)$

or

(13) $\quad \mathrm{I}(\phi, \psi')$.

So far, we have regarded the notion of inferring as a binary relation between sentences. But, in general, inferences are drawn from several assumptions. Since the body of ones beliefs need not be closed under simplification or adjunction, these assumption should not be represented by their conjunction; and since the order in which premises are being considered may affect a person's beliefs, it seems best to treat the relation of inferring as one obtaining between a sequence of premises and a conclusion. For intuitive reasons, we shall not allow that a person infers conclusions from the empty sequence, i.e. from nothing at all. Accordingly, for every integer n larger then 1, we shall make use of an n-ary operator 'I' which serves to distinguish e.g. the following valid principles:

(14) $I(\phi \;\&\; \psi; \chi) \rightarrow [B(\phi \;\&\; \psi) \rightarrow B\chi]$

Informally: if one infers that χ from the assumption that $(\phi \;\&\; \psi)$ then, if one believes that $(\phi \;\&\; \psi)$ one believes that χ.

(15) $I(\psi \;\&\; \phi; \chi) \rightarrow [B(\psi \;\&\; \phi) \rightarrow B\chi]$

Informally: if one infers that χ from the assumption that $(\psi \;\&\; \phi)$ then, if one believes that $(\psi \;\&\; \phi)$ one believes that χ.

(16) $I(\phi, \psi; \chi) \rightarrow [(B\phi \;\&\; B\psi) \rightarrow B\chi]$

Informally: if one infers that χ from the respective assumptions that ϕ and that ψ, then, if one believes that ϕ and believes that ψ one also believes that χ.

Note that the inferences mentioned by the antecedents of (14), (15), and (16) are all regarded as genuinely different, as are the assumptions regarding ones beliefs.

We shall respect tradition by endorsing the postulates:

(17) $K\phi \rightarrow \phi$

[if one knows that ϕ then ϕ], and

(18) $K\phi \rightarrow B\phi$

[if one knows that ϕ then one believes that ϕ].

Among principles regarding iterated belief and knowledge, the following are frequently proposed:

(19) $B\phi \rightarrow BB\phi$

[if one believes that ϕ then one believes that one believes that ϕ],

(20) $B\phi \rightarrow KB\phi$

[if one believes that ϕ then one knows that one believes that ϕ],

(21) $K\phi \rightarrow BK\phi$

[if one knows that ϕ then one believes that one knows that ϕ], and

(22) $K\phi \rightarrow KK\phi$

[if one knows that ϕ then one knows that one knows that ϕ].

The customary notions of belief and knowledge seem such that (a) if a person acts as if it were the case that ϕ, and the outcome of his actions matters to him, that tends to confirm that he believes that ϕ, (b) his outspoken denial that he believes that ϕ tends to confirm that he fails to believe that he believes that ϕ, and (c) such actions and denials are not treated as both confirming and disconfirming the very same thing. Accordingly, (19), and indirectly also (20), do not seem plausible with respect to customary uses of 'belief' and 'knowledge'. Further, a person may well believe that ϕ and, in fact, possess sufficient evidence to warrant his claim that ϕ, while he may not believe himself to be in possession of such sufficient evidence. In these circumstances, (21), and indirectly also (22), appear to be false. While these considerations are not conclusive, they may serve to indicate that the formulas (19)–(22) are at least controversial and should for that reason not be treated as postulates of a general epistemic logic. However, it will be obvious how one can add any of them to the axioms and to the semantical requirements in contexts where such additions seem plausible.

Up to now, we have confined our discussion to ideas which will find formal expression in the sentential part of the subsequent theory. Of greater interest and difficulty, however, will be the introduction of quantifiers which are allowed to reach into epistemic contexts.

As Quine in [10] and Kaplan in [6] have pointed out, the sentence

(23) Someone is such that Ralph believes that he is a spy

has at least one sense in which its truth might be of interest to the F.B.I., a sense which might be expressed by the formal sentence

(24) $\exists x$[Ralph believes that x is a spy]

provided that quantification is so construed that (24) is not a logical consequence of such almost trivial sentences as

(25) Ralph believes that the shortest spy is a spy.

Similarly, assertions which have been of concern to Hintikka in [3] such as

(26) Tom knows who the President is,

or rather its less ambiguous paraphrase to which alone we shall address ourselves:

(27) Someone is such that Tom knows of him that he is the President

might be translated by

(28) $\exists x$[Tom knows that $x =$ the President],

provided that (28) is so interpreted that it does not follow by Existential Generalization from the trivial assertion that

(29) Tom knows that the President = the President.

As Hintikka has made explicit in [4], we cannot endow (24) and (28) with the sense at issue if an unrestricted substitutional interpretation of quantifiers is given. For, if we stipulate that '$\exists x$ [Tom knows that $x = a$]' shall be true just in case for some name 'b', the result of replacing 'x' by 'b' in 'Tom knows that $x = a$' is true, then the given quantified statement (which might be a translation of (28)) will be true by virtue of the truth of 'Tom knows that $a = a$' (which might translate (29)). Hence, if a substitutional interpretation is chosen at all, the class of substitutable names must be restricted. Efforts to do so bring out what I believe to be an ambiguity in the English sentences (23) and (26): it could be that the F.B.I. is interested in Ralph's beliefs if (i):

(30) Ralph believes that Olga Stroganoff is a spy,

assuming that 'Olga Stroganoff' is a name which represents a person *to the F.B.I.* while being sufficiently devoid of descriptive content to make it unlikely that Ralph has his belief by virtue of the connotation of the name alone. Thus, the F.B.I. might show some interest in the truth of (30) even if Ralph has obtained his belief just by having overheard a conversation at the next table near the Russian embassy without having any idea who Olga Stroganoff is. It could also be (ii) that the F.B.I. requires more; namely, that there be some name or description, say 'the pretty blonde who seduced Ralph', which represents a person *to Ralph* (perhaps by reminding him of the occasion at which he met that person) and which Ralph knows to be Olga Stroganoff. But (iii), it seems to me that the F.B.I. would show no interest if it were merely the case that

(31) Ralph believes that the most suspicious-looking female in Ralph's acquaintance is a spy,

where 'the most suspicious-looking female in Ralph's acquaintance' is assumed to be a denoting phrase which represents a particular person *to Ralph only*, but is not representative of a person to the F.B.I. Thus, a name which is private to the epistemic subject does not seem to warrant exportation if we understand 'someone is such that Ralph believes him to be a spy' to mean, roughly, 'some objectively identifiable person is such that Ralph believes him to be a spy'. And this seems so even if the private name in fact uniquely denotes, is representative of a person to Ralph, and is a very 'vivid' name.

Similarly, we might be prepared to admit that Tom knows who the President is (or better, that there is someone whom Tom knows the President to be) if

(32) Tom knows that Richard M. Nixon = the President

in contexts like these: Teacher: 'Can anyone tell me who the President is?', – Tom: 'Richard M. Nixon is the President', – Teacher: 'Good, Tom knows who the President is'. Here 'Richard M. Nixon' is assumed to be a name which is objectively representative of a person and sufficiently free of descriptive content to bar the possibility that Richard M. Nixon's being the President is a trivial bit of knowledge. But there is also a sense in which Tom would not be said to know who the President is unless, in addition to (32), Tom knows Nixon under a description, say 'the man whose hand Tom shook yesterday', which is representative of a particular person for Tom.

We shall formally provide for names which warrant exportation from epistemic contexts; names which are 'special' in the objective sense and ones which are 'special' in the subjective sense, so as to allow for translations of (23) and (26) in both of their senses.

In order to interpret contexts of belief and knowledge, we shall have to employ certain formally definable entities which can go proxy for bodies of belief and knowledge. Since we want to allow for inconsistent bodies of belief, we cannot very well identify a body of beliefs with a possible world (say, one to which access is gained by virtue of ones beliefs) or a set of possible worlds (those which are compatible with the believer's attitudes).

Instead, it seems natural to represent ones body of beliefs simply by a set of (possibly inconsistent) statements; namely those which, intuitively, one believes to be true. Similarly, the body of ones knowledge will be represented by the set of all statements which one knows to be true. Having opted for an interpretation of epistemic contexts by means of sets of statements, it makes for a more homogeneous semantics if 'possible worlds' are generally replaced by arbitrary (and possibly inconsistent) sets of statements. Let us call such sets 'theories' or, better still, 'tales'. Truth, denotation, and other semantical notions will therefore be relativized to tales.

Among semantical theories with which I am acquainted, those which represent possible worlds by certain sets of statements (state descriptions, model sets) share a certain unnaturalness in their interpretation of identities: special stipulative clauses are employed to ensure the reflexivity of identity and the interchange of identicals in atomic formulas. This aesthetic defect can be overcome in a manner which we describe with reference to our 'tales' as follows: We associate with every tale a function which maps certain names and variables (intuitively, those constants which denote and those variables which have values according to the tale) into other names which will be called the *representative* (or, standard) *names* according to the tale. Thus, with respect to a tale regarding arithmetic, the various names of the number four (barring descriptions) might be mapped onto the representative name '4', also, according to the F.B.I.-story, various names of a suspected spy might be represented by the one name 'Olga Stroganoff'; and, for the purposes of a biography of the President, various names of the President might be mapped on the one representative name 'Richard M. Nixon'. If R is such a function and a and b are constants, then the identity statement '$a = b$' should be true (according to a given tale) just in case R is defined on a and b and $R(a) = R(b)$ [i.e. if a and b denote and are represented by the same name]. And, if F is an extensional one-place predicate, we can guarantee its extensionality by the requirement that 'Fa' shall be true just in case '$FR(a)$' is true [i.e. that any name shall be interchangeable with its representative in this context]. This last provision will have the side effect that 'Fa' will be true only if a is a denoting term (one in the domain of R), since otherwise $R(a)$ is the empty set and the expression '$FR(a)$' will degenerate into the ill-formed expression F which can never be true. Generally, identities and basic predicates are treated as expressive of positive qualities and relations (they truly apply only to denoting terms). It is one of my idiosyncrasies that I am pleased with this incidental benefit from the requirement of extensionality, and I have argued for its plausibility elsewhere (in [1]).

In Henkin-type completeness proofs of theories with identity one frequently constructs a universe by picking certain representative names and identifying the individuals with those names. With a view to this possibility, it seems natural to dispense with universes altogether and to work only with the representative names themselves. In this way, we get rid of some metaphysical clutter (the universes), achieve an interpretation of identities which seems just as natural as ones which use assignments of individuals to terms, and at the same time we obtain one category of special names which in epistemic contexts seem to warrant exportation.

Given these preliminaries, it is tempting to propose a substitutional interpretation of quantification of such a sort that the names to be substituted for the variables of quantification are representative names. However, there does not appear to be any easy and natural way of finding axioms corresponding to this interpretation (presumably a predicate would be needed which serves to pick out not just denoting names, as does the predicate 'exists', but, more specifically, representative denoting names). Instead, we shall proceed in roughly the following manner: Suppose that R is an assignment of representative names to both names and variables. Then a statement like '$\exists x\phi(x)$' shall hold if there is a variable 'y' (free in $\phi(y)$) which has values (i.e. which is in the domain of R) such that the formula $\phi(y)$ holds. And $\phi(y)$, in turn, shall hold just in case $\phi(R(y))$ holds. If ϕ is a purely extensional context, then the previously mentioned clause for extensionality will provide that the name $R(y)$ may be replaced in ϕ by any of the names it represents, and '$\exists x\phi(x)$' will have its unrestricted substitutional sense. But if ϕ turns out to be an epistemic context, then the representative name $R(y)$ may not be replaced in ϕ by an arbitrary name and '$\exists x\phi(x)$' will, in effect, have a substitution interpretation which is confined to representative names. As an intermediate step in evaluating the truth of quantified statements, we shall look at the formula $\phi(y)$ where the free variable 'y' serves as a place holder for its representative name; and in axiomatizing the theory, the syntactical difference between free variables and other terms will serve us instead of the purely semantical difference between representative names and other terms, since we cannot express the latter distinction directly in the object-language. Thus, sloppily expressed, Existential Generalization will hold in the form:

(33) $[\phi(x)$ & x exists$] \rightarrow \exists x\phi(x)$,

but only if 'x' is a variable. If ϕ is an extensional context, but not if it is an epistemic context, the generalized form of (33) may be derived by interchange of identicals. To give an epistemic example of (33), the sentence

(34) $\exists x$ [Kepler knew that x numbers the planets]

may be inferred from

(35) Kepler knew that x numbers the planets & x exists,

where the indefinite 'x' is so interpreted that it stands proxy for a suitable representative name; say for the standard name '9'. Under any interpretation which provides that '9' is, in fact, a representative name, (34) will be true provided

(36) Kepler knew that 9 numbers the planets

is true. But the inference of (34) from (36) will not hold under every interpretation, since it might be that '9' is not treated as a representative name. Also, Kepler might not have known 9 by the name '9'. If so, we might still grant (34) if there is some 'it' (called '9' in astronomy) such that Kepler knew 'it' to number the planets; i.e. if (35) is the case. On the other hand, (34) is definitely not warranted by the assumption that

(37) Kepler knew that the number of planets numbers the planets.

Assuming that 'Richard M. Nixon' is a representative name and that a is any term denoting Nixon, it seems that we can express (in one sense) that

(38) Tom knows who a is

by existential generalization from

(39) Tom knows that Richard M. Nixon $= a$,

which seems to work nicely for all names *a*, except for the one representative name *a* = 'Richard M. Nixon' itself. Does the statement

(40) Tom knows that Richard M. Nixon = Richard M. Nixon

really imply that Tom knows who Richard M. Nixon is?

Note, to begin with, that if I ask you whether you know who *a* is, and you reply by saying that he is *b*, I can always continue by asking you next whether you know who *b* is. And, in playing this game, you will either end up by asserting a trivial identity, or go around in circles, or continue forever. Given these alternatives and the assumption that there is a definite answer to my questions, it seems best to stop the game with an assertion of the form '$a = a$', where *a* is a representative name which, in the framework of a given tale, is thought of as a proper name which lacks descriptive content. Observe next that (40) is not entirely trivial; for we interpret identity statements as having existential import, so that '$a = a$' may be read '*a* exists'. Accordingly, what Tom is said to know, by (40), is that Richard M. Nixon exists. But if 'Richard M. Nixon' is thought of as a name which is not given any descriptive content by that tale in whose context we appraise the truth of (40), then one knows who Richard M. Nixon is (in the sense that there is someone whom one knows Richard M. Nixon to be) if one knows (by a knowledge which is not a 'knowledge by descriptions') simply that Richard M. Nixon exists. If still more is required before we are willing to admit that Tom knows who Nixon is, we may have to fall back on a special subclass of representative names: ones which are (subjectively) representative of a person for Tom. To obtain this narrower class of names, we shall employ a special one-place predicate 'A' (which reminds us, somewhat inappropriately, of 'acquaintance'), where formulas like '$\mathrm{A}x$' are read 'there is some name or description by which one knows x'. Thus, that sense of 'one knows what the number of planets is' in which we require that one must be in some special *rapport* with the number 9, might be expressed by the relativization of quantifiers to the predicate 'A' as follows:

(41) $\exists x\,[\mathrm{A}x \;\&\; \mathrm{K}(x \text{ numbers the planets})]$,

which will be true (in a tale regarding numbers and planets) if

(42) A9 & K (9 numbers the planets)

is true; where 'A9', in turn, is true if one knows the number 9 under some description like 'the number of my outstretched fingers' which may be special subjectively by virtue of its vividness or its associations with the acquisition of number-terms; whereas the name '9' is special objectively (i.e. relative to the tale of astronomy) by standardizing reference for the purpose of astronomy.

In order to link up the predicate 'A' with epistemic notions, we postulate:

$$(43) \qquad A\tau \rightarrow K(\tau = \tau)$$

Roughly: if one knows τ under some name or description, then one knows that τ exists; and

$$(44) \qquad (A\tau \ \& \ [K(\zeta = \tau) \vee K(\tau = \zeta)]) \rightarrow A\zeta.$$

Roughly: if one knows τ under some name or description and one also knows either that ζ is τ or that τ is ζ, then (thereby) one also knows ζ under some name or description.

For the sake or simplicity, we shall not undertake to treat of modal contexts. Thus, we shall not be able to express statements like 'ϕ is compossible with everything Tom knows' or (the equally interesting) 'ϕ is compossible with something Tom knows'. Further, it would seem natural to require that '$A\tau$' shall mean 'one has non-inferential knowledge of τ'. This could apparently be done by requiring that one should know that τ exists in all those (epistemically) alternative circumstances which differ from the actual one at most with respect to the assumption that τ's existence has been inferred. For, if one is acquainted with τ, then it must be possible that ones knowledge that τ exists is not the result of an inference one makes. However, without taking account of modal notions, we cannot express this requirement. Further, there is one sense of 'knowing who (or what) *a* is' in which such knowledge does not presuppose the existence or even the possibility of *a*. Thus, we might grant that someone knows who Faust is if the person is thoroughly familiar with the relevant literature, or that he knows what the largest number is if he can define that inconsistent notion. These are senses in which 'one knows what *a* is' does not mean 'there exists some item which one knows to be *a*', their analysis seems to require semantical reference to tales other than the true one and, in the interest of simplicity, we shall not attempt to capture these senses.

Further informal remarks will accompany the axiomatic and semantical developments wherever it seems helpful to elucidate the formalism.

2. Syntactical preliminaries

We shall assume, in the meta-theory, the existence of a proper class of constants and of a proper class of variables. Also, for every natural number n, there shall be a set, of indefinite size, of n-place predicates. We shall use parentheses and brackets, the customary connectives '$\sim$', '$\rightarrow$', '&', '$\vee$', and '$\leftrightarrow$', the quantifers '$\exists$' and '$\forall$', the identity sign '$=$', and the descriptive operator '$\imath$'. In addition, we have the following non-schematic letters: the one-place sentential operator 'B' in contexts like

$B\phi$ [read: one believes that ϕ];

the one-place sentential operator 'K' in contexts like

$K\phi$ [read: one knows that ϕ];

for every integer n (such that $2 \leqslant n$), the n-place sentential operator 'I^n', whose superscript we delete in contexts like

$I(\phi_1, \ldots, \phi_{n-1}; \phi_n)$ [read: one rationally infers that ϕ_n from the respective assumptions that $\phi_1, \ldots, \phi_{n-1}$];

and finally the one-place predicate 'A' in the contexts

$A\tau$ [read: one knows τ under some name or description].

All categories of symbols are disjoint, and no symbol shall be the empty set 0.

The syntactical predicates 'is a *formula*', 'is a *term*', and 'is *free* in' are assumed to be defined in the expected manner. *Sentences* are formulas which contain no free variables, and *names* are terms without free variables. Given any set T of formulas, the *formulas of* T (*terms of* T, *sentences of* T, *names of* T) shall be the formulas (terms, sentences, names) which are such that every variable, individual constant and predicate (excepting 'A') which occurs in them also occurs in some member of T.

If θ is a well-formed expression, and R is a function whose domain is a set of constants and variables and whose range is a set of terms then $\theta\{R\}$ [read: the proper simultaneous *substitution*, in θ, of terms for free variables in accord with R] is recursively defined as follows:

(1) if θ is a variable, then $\theta\{R\} = R(\theta)$, if θ is in the domain of R; and $\theta\{R\} = \theta$ otherwise,

(2) if θ is an individual constant, then $\theta\{R\} = \theta$,

(3) if for some n-place predicate π and n-term sequence $\tau_0, \ldots, \tau_{n-1}$ of terms $\theta = \pi\tau_0 \ldots \tau_{n-1}$, then $\theta\{R\} = \pi\tau_0\{R\} \ldots \tau_{n-1}\{R\}$,

(4) if for some terms τ_0, τ_1, $\theta = [\tau_0 = \tau_1]$, then $\theta\{R\} = [\tau_0\{R\} = \tau_1\{R\}]$,

(5) if for some term τ, $\theta = A\tau$, then $\theta\{R\} = A\tau\{R\}$,

(6) if for some formula ϕ, $\theta = B\phi$, then $\theta\{R\} = B\phi\{R\}$,

(7) if for some formula ϕ, $\theta = K\phi$, then $\theta\{R\} = K\phi\{R\}$,

(8) if for some n-term $[2 \leqslant n]$ sequence of formulas $\phi_1, \ldots, \phi_n$ $\theta = I(\phi_1, \ldots, \phi_{n-1}; \phi_n)$, then $\theta\{R\} = I(\phi_1\{R\}, \ldots, \phi_{n-1}\{R\}; \phi_n\{R\})$,

(9) if for some formula ϕ, $\theta = \sim\phi$, then $\theta\{R\} = \sim\phi\{R\}$,

(10) if for some formulas ϕ and ψ, $\theta = (\phi \,\&\, \psi)$, then $\theta\{R\} = (\phi\{R\} \,\&\, \psi\{R\})$, and so forth for other sentential compounds,

(11) if for some formula ϕ and variable α, $\theta = \exists\alpha\phi$, then $\theta\{R\} = \exists\alpha\phi\{R'\}$, where $R' = R$ minus the set of all pairs $\langle x, y\rangle$ where either $x = \alpha$ or α is free in y, and similarly for universal statements and descriptive terms.

We say that $\theta\{R\}$ is *proper* if θ contains no well-formed expression of the form $\exists\alpha\phi$, $\forall\alpha\phi$, or $\iota\alpha\phi$ such that for some variable β, β is free in ϕ and α is free in $R(\beta)$.

$R^{x_0 \ldots x_{n-1}}_{y_0 \ldots y_{n-1}}$ (and, as an intended special case R^x_y) is that assignment which differs from R, if at all, only by assigning to each x_i the item y_i; note that all the x's are in the domain of this variant of R.

3. The axioms and inference rules

Suppose that n is a natural number, m and k are positive integers, $2 \leqslant m$, α and β are distict variables, ϕ and ψ are formulas, β is not free in ϕ, π is an n-place predicate, χ is an m-term and ξ a k-term sequence of formulas, ζ, η, and θ are terms, τ is an n-term sequence of terms, and d is the identity relation confined to the variables in ϕ. Then, an *axiom* is a formula of one of the following forms:

(Ax1) ϕ, if ϕ is a tautology

(Ax2) $\forall\alpha(\phi \to \psi) \to (\forall\alpha\phi \to \forall\alpha\psi)$

(Ax3) $\exists\alpha\phi \leftrightarrow \sim\forall\alpha\sim\phi$

(Ax4) $(\forall\alpha\phi \,\&\, \zeta=\zeta) \rightarrow \phi\{d^{\alpha}_{\zeta}\}$, provided that ζ is a variable and $\phi\{d^{\alpha}_{\zeta}\}$ is proper

In words: If everything satisfies the condition ϕ and if 'it' is self-identical (or, 'it' exists), then 'it' satisfies the condition ϕ. (Ax4) is the principle of Universal Instantiation confined to variables which have values.

(Ax5) $\forall\alpha\, \alpha=\alpha$

In words: everything (i.e. every actual thing) is self-identical (or, exists).

(Ax6) $\pi\tau_0 \ldots \tau_{n-1} \rightarrow (\tau_0=\tau_0 \,\&\ldots\&\, \tau_{n-1}=\tau_{n-1})$

In words: If the respective items $\tau_0, \ldots, \tau_{n-1}$ enter into the (positive) relation π, then each of these items is itself-identical (or, exists). Informally: atomic formulas of the sort displayed in the antecedent of (Ax6) express the possession of positive qualities or the entering into positive relations.

(Ax7) $\zeta=\eta \rightarrow \eta=\zeta$

In words: If ζ is identical with η then η is identical with ζ: identity is symmetric.

(Ax8) $(\zeta=\eta \,\&\, \eta=\theta) \rightarrow \zeta=\theta$

In words: If ζ is the same as η and η is the same as θ then ζ is the same as θ: identity is transitive.

(Ax9) $(\pi\tau_0 \ldots \tau_{n-1} \,\&\, \tau_i=\zeta) \rightarrow \pi\tau_0 \ldots \tau_{i-1}\,\zeta\tau_{i+1} \ldots \tau_{n-1}$, for all $i<n$.

In words: If the respective items $\tau_0, \ldots \tau_{n-1}$ enter into the relation π and the i-th item τ_i is identical with ζ then the sequence of items in which τ_i is replaced by ζ will also enter that relation. (Ax7)–(Ax9) together provide that all atomic formulas which are free of epistemic operators form extensional contexts.

(Ax10) $\forall\beta\,[\beta=\imath\alpha\phi \leftrightarrow \forall\alpha(\alpha=\beta \leftrightarrow \phi)]$

In words: Exactly those items are identical with the thing which is ϕ which are such that they, and they alone, are ϕ.

(Ax11) $\zeta=\zeta \rightarrow \exists\beta\,\zeta=\beta$, if ζ is a constant or $\zeta=\imath\alpha\phi$,

In words: If ζ is self-identical then there is something which it is.

(Ax12) $(\alpha = \beta \;\&\; \phi) \rightarrow \phi'$, where ϕ' results from ϕ by replacing one or more free occurrences of α by free occurrences of β. Note that this principle of interchange holds only with respect to *variables* α and β.

(Ax13) $K\phi \rightarrow \phi$

In words: If one knows that ϕ, then ϕ. Knowledge implies truth.

(Ax14) $K\phi \rightarrow B\phi$

In words: If one knows that ϕ then one believes that ϕ.

(Ax15) $B\phi \rightarrow \alpha = \alpha$, provided that α is free in ϕ.

In words: If one believes that 'it' is ϕ, then the 'it' in question exists. Very loosely: A free variable α in the context 'one believes that α is ϕ' serves to express that one believes *of* the actual object α that it is ϕ. (Ax15) is used in proving, e.g. that if one believes that α is ϕ then there exists something which one believes to be ϕ.

(Ax16) $I(\chi_0, \ldots, \chi_{m-2}; \chi_{m-1}) \rightarrow \alpha = \alpha$, if for some $i < m$, α is free in χ_i.

Loosely: If one carries out a certain rational inference and the variable α is free in one of the formulas representing a step in that inference, then one is inferring something 'about' the actual object α.

(Ax17) $I(\chi_0, \ldots, \chi_{m-2}; \chi_{m-1}) \rightarrow ([B\chi_0 \;\&\ldots\&\; B\chi_{m-2}] \rightarrow B\chi_{m-1})$

In words: If one infers that χ_{m-1} from the respective assumptions that $\chi_0, \ldots, \chi_{m-2}$ and if one believes each of these assumptions, then one also believes that χ_{m-1}. Or: the body of ones beliefs is closed under those inferences which one actually carries out.

(Ax18) $I(\chi_0, \ldots, \chi_{m-2}; \chi_{m-1}) \rightarrow ([K\chi_0 \;\&\ldots\&\; K\chi_{m-2}] \rightarrow K\chi_{m-1})$

In words: If one rationally infers that χ_{m-1} from the respective assumptions that $\chi_0, \ldots, \chi_{m-2}$ and if one knows that each of these assumptions is true, then one also knows that the conclusion is true. Or: the body of ones knowledge is closed under those rational inferences which one actually makes.

(Ax19) $I(\chi_0, \ldots, \chi_{m-2}; \chi_{m-1}) \rightarrow K[(\chi_0 \;\&\ldots\&\; \chi_{m-2}) \rightarrow \chi_{m-1}]$

In words: If one rationally infers that χ_{m-1} from the respective assumptions that $\chi_0, \ldots, \chi_{m-2}$, then one knows that the conjunction of those assumptions is true only if the conclusion χ_{m-1} is true.

(Ax20) $[\mathrm{I}(\xi_0, \ldots, \xi_{k-1}; \chi_0) \mathbin{\&} \ldots \mathbin{\&} \mathrm{I}(\xi_0, \ldots, \xi_{k-1}; \chi_{m-1})] \rightarrow$
$\rightarrow [(\chi_0, \ldots, \chi_{m-1}; \phi) \rightarrow \mathrm{I}(\xi_0, \ldots, \xi_{k-1}; \phi)]$

Informally- the relation of inferring is transitive. Or: the set of all inferences which one draws from given assumptions is closed under further inferences which one actually carries out.

(Ax21) $\mathrm{A}\zeta \rightarrow \mathrm{K}(\zeta = \zeta)$

In words: If one knows ζ by some name or description, then one knows that ζ is ζ (or, that ζ exists).

(Ax22) $(\mathrm{A}\zeta \mathbin{\&} [\mathrm{K}(\zeta = \eta) \vee \mathrm{K}(\eta = \zeta)]) \rightarrow \mathrm{A}\eta$

In words: If one knows ζ by some name or description and if one either knows that ζ is η or that η is ζ, then one also knows η under some name or description.

This completes the list of axioms. No special effort has been made to streamline the axioms or to prove their independence.

We say that ϕ *follows by Modus Ponens* from ψ and χ just in case either $\psi = (\chi \rightarrow \phi)$ or $\chi = (\psi \rightarrow \phi)$. And ϕ *follows by Universal Generalization* from ψ if and only if there exist χ, θ, and α such that $\psi = (\chi \rightarrow \theta)$, $\phi = (\chi \rightarrow \forall\alpha\theta)$, and α is not free in χ. A finite non-empty sequence $\langle s_0, \ldots, s_{n-1}\rangle$ is a *proof* of ϕ just in case $\phi = s_{n-1}$ and, for every $i < n$, s_i is a formula and either (a) s_i is an axiom, or (b) there are $j, k < i$ such that s_i follows by Modus Ponens from s_j and s_k, or (c) for some $j < i$, s_i follows by Universal Generalization from s_j; and such a sequence is a *derivation of ϕ from* a set K of formulas just in case $\phi = s_{n-1}$ and, for every $i < n$, s_i is a formula and either (a) s_i is in K, or (b) there is a subsequence t of s such that t is a proof of s_i, or (c) there are $j, k < i$ such that s_i follows by Modus Ponens from s_j and s_k. We say that ϕ is a *theorem* (that ϕ is *derivable* from K) if there is a proof of ϕ (a derivation of ϕ from K). A set K of formulas is *consistent* if there is no formula such that both it and its negation are derivable from K; and K is *maximally consistent with respect to* the formulas in K if K is not properly contained in any set K' which is consistent and comprises only formulas of K (i.e. formulas, roughly, whose language is that of K), while K is consistent.

Due, e.g., to the fact that we have a proper class of variables, we also have a proper class of theorems and no set is maximal in the sense of comprising all theorems of our unrestricted language.

4. THE SEMANTICS

We begin by defining and subsequently explaining the notion of a model.

DEF. M is a *model* if and only if there exist T, B, K, I, A, H, and R such that $M = \langle T, B, K, I, A, H, R \rangle$ and

(1) T is a set of formulas,

(2) B is a set of sentences of (in the language of) T,

(3) K is included in B,

(4) I is a set of at least 2-term finite sequences of sentences of T, and
   (a) whenever $\langle \phi_0, ..., \phi_{n-1}, \phi_n \rangle$ is in I and for every $i < n$, ϕ_i is in B, then ϕ_n is in B,
   (b) whenever $\langle \phi_0, ..., \phi_{n-1}, \phi_n \rangle$ is in I and for every $i < n$, ϕ_i is in K, then ϕ_n is in K,
   (c) whenever $\langle \phi_0, ..., \phi_{n-1}, \phi_n \rangle$ is in I and for every $i < n$, $\langle \psi_0, ..., \psi_{m-1}, \phi_i \rangle$ is in I, then $\langle \psi_0, ..., \psi_{m-1}, \phi_n \rangle$ is in I,
   (d) whenever $\langle \phi_0, ..., \phi_{n-1}, \phi_n \rangle$ is in I, then $[(\phi_0 \mathbin{\&} ... \mathbin{\&} \phi_{n-1}) \rightarrow \phi_n]$ is in K;

(5) A is a set of names of (in the language of) T, and
   (a) whenever τ is in A, then $[\tau = \tau]$ is in K,
   (b) whenever τ is in A and either $[\tau = \zeta]$ or $[\zeta = \tau]$ is in K, then ζ is in A;

(6) H is a set of sentences of (in the language of) T of the form $\pi\tau_0 ... \tau_{n-1}$, where π is an n-place predicate and τ is an n-term sequence of constants; and

(7) R is a function whose domain is a set of variables and of constants of (in the language of) T, whose range is a set of constants of T, and such that, for every term τ in its domain, $R(R(\tau)) = R(\tau)$, and for some variable α, $R(\tau) = R(\alpha)$.

Briefly, the constituent T in a model is regarded as the tale to which truth is relativized. T is allowed to be any set of formulas, whether they be consistent or inconsistent, deductively closed or not, and of any cardinality.

The constituent B in a model represents the body of all beliefs which one has according to T. B may be any set of sentences of T whatever (including the empty body of beliefs). But we do require that B shall be a set of *sentences*, since it is not intuitively clear what items in the body of a person's beliefs shall be presented by open formulas.

The set K in a model represents the body of one's knowledge. Since one knows only what one believes, K is included in B. By implication, K is also a set of sentences in the language of T.

The constituent I of a model represents the inferences which (according to the tale T) one actually carries out. Accordingly, each member of I is a sequence whose last term represents the conclusion one draws, and whose other terms represent the respective premises from which one draws the conclusion. Since the intuitive relation of inferring is at least 2-place, we exclude 0-term and 1-term sequences from I. The condition (4) (a) provides that the body B of one's beliefs is closed under one's inferences. According to (4) (b), the same is true of the body K of one's knowledge. (4) (c) requires that the set of inferences drawn from given premises $\psi_0, \ldots, \psi_{m-1}$ is closed under further inferences which one cares to make. The condition (4) (d) has this effect: whenever one draws a conclusion from given premises, then one knows that conditional to be true whose antecedent is the conjunction of the respective premises and whose consequent is the draw conclusion. Without this connection between inferring and knowing, we would not feel that I represents *rational* inferences.

The set A in a model is intended to contain those names of T (whether they be constants or descriptions) which represent things to the epistemic subject: they are to be names by which things are known. According to (5) (a), if τ is such a name, then the information that $\tau = \tau$ (which turns out to be equivalent, in our system, to be information that τ exists) must be in the body K of one's knowledge; and (5) (b) has the effect that whenever τ is a subjectively representative name in A and ζ is a name whose designatum is known to be identical with that of τ, then ζ is again one of the names in A.

The constituent H in a model consists of certain relational atomic sentences in the language of T; namely those which, intuitively, *hold* in the tale T. H will take care of the base-step in the recursive definition of truth and hence should comprise those sentences in whose further analysis we are not interested.

The function R in a model is the assignment of *representative names* to constants and variables. Constants which are not in the domain of R are thought of as *non-denoting*, and variables not in its domain are ones *without values*. The condition that for every denoting term τ, $R(R(\tau)) = R(\tau)$ says that every representative name represents itself, and implies also that every representative name $R(\tau)$ denotes [for, if τ is in the domain of R while $R(\tau)$ is not, then $R(R(\tau)) = 0$; and 0 is not a constant, as is $R(\tau)$]. The condition that for every τ in the domain of R there is a variable α such that $R(\tau) = R(\alpha)$ provides that for every representative name there will be a variable which can go proxy for that name.

If $M = \langle T, B, K, I, A, H, R\rangle$, then by M^x_y we mean $\langle T, B, K, I, A, H, R^x_y\rangle$.

DEF. Suppose that $M = \langle T, B, K, I, A, H, R\rangle$ is a model, m and n are natural numbers, $2 \leqslant m$, α is a variable, ζ and η are terms, τ is an n-term sequence of terms, ϕ, ψ, are formulas, χ is an m-term sequence of formulas, π is an n-term predicate, and I is an m-term operator. Then the notions *Val* (M, ζ) [in words: the value, according to M, of ζ] and *M sat* ϕ [in words: M satisfies ϕ] are recursively characterized as follows:

(1) $\mathrm{Val}(M, \zeta) = R(\zeta)$, if ζ is either a constant or variable of T,

(2) M sat $\pi\tau_0 \dots \tau_{n-1}$ just in case $\pi\mathrm{Val}(M, \tau_0) \dots \mathrm{Val}(M, \tau_{n-1})$ is in H,

(3) M sat $[\zeta = \eta]$ just in case $\mathrm{Val}(M, \zeta) = \mathrm{Val}(M, \eta) \neq 0$,

(4) M sat $A\zeta$ just in case for some term η, $\eta\{R\}$ is in A and M sat K $[\zeta = \eta]$,

(5) M sat Bϕ just in case $\phi\{R\}$ is in B,

(6) M sat Kϕ just in case M sat ϕ and $\phi\{R\}$ is in K,

(7) M sat $\mathrm{I}(\chi_0, \dots, \chi_{m-2}; \chi_{m-1})$ just in case $\langle\chi_0\{R\}, \dots, \chi_{m-1}\{R\}\rangle$ is in I and for every $i < m-1$ M sat χ_i, only if M sat χ_{m-1},

(8) M sat $\sim\phi$ just in case ϕ is a formula of T and it is not the case that M sat ϕ,

(9) M sat $(\phi \mathbin{\&} \psi)$ just in case M sat ϕ and M sat ψ, and similarly for other sentential compounds,

(10) M sat $\exists\alpha\phi$ just in case there is a variable β in the domain of R such that $M^\alpha_{R(\beta)}$ sat ϕ,

(11) M sat $\forall\alpha\phi$ just in case for every variable β in the domain of R, $M^\alpha_{R(\beta)}$ sat ϕ,

(12) $\mathrm{Val}(M, \imath\alpha\phi) = R(\beta)$ provided that $\imath\alpha\phi$ is a term of T, β is a variable in that domain of R and for every variable γ in the domain of R: $M^\alpha_{R(\gamma)}$ sat ϕ just in case $R(\gamma) = R(\beta)$; and $\mathrm{Val}(M, \imath\alpha\phi) = 0$ otherwise.

Recalling the $R(\tau) = 0$ whenever τ is not in the domain of R, let us say that a term τ *denotes* (according to M) if the value of τ (according to M) differs from the empty set 0. According to clauses (1) and (12), the values of all denoting terms are constants of T (representative names), and the common value of all non-denoting terms is the empty set 0. Roughly, the value of a denoting description $\imath\alpha\phi$ is the representative name (among those which represent variables) which may be substituted for α in ϕ (making the result true).

According to clause (2), an atomic formula of the form $\pi\tau_0 \ldots \tau_{n-1}$ is satisfied just in case the result of replacing in it all terms τ_i by their representative names holds; and the result of that replacement can hold only if the original formula is in the language of T and if all terms τ_i are denoting terms. For example, if π is a one-place predicate, M sat $\pi\tau$, and we assumed that $\mathrm{Val}(M, \tau) = 0$, then the ill-formed formula π would have to be in H, which is ruled out by the condition (6) on models.

Clause (3) provides that identities are true if both terms flanking the identity sign denote and are represented by the same name. Using the conditions (2) and (3), it follows easily that all predicates form extensional contexts.

The condition (4) states that Aζ is true if ζ is known to be η, where the result of replacing the free variables in η by their representative names is in the set A of those names which are representative of things for the epistemic subject.

Clause (5) tells us that one believes that ϕ just in case the result $\phi\{R\}$ of replacing the free variables in ϕ by their representative names (according to R) is in the body B of ones beliefs. Here, as indeed in all epistemic contexts, free variables go proxy for their representative names.

According to (6), one knows that ϕ if ϕ is true and the sentence $\phi\{R\}$, obtained from ϕ by replacing free variables by their representative names, is in the body K of ones knowledge. Similarly, by (7), formulas which express that certain inferences are drawn by the epistemic subject shall be satisfied if the corresponding sequence of sentences is in I and the conditional, formed by conjoining the assumptions of the inference and letting that conjunction imply the conclusion of the inference, is true. Due to this latter provision, together with (4) (d) in the definition of 'model', the conditionals in question will be known to be true whenever the corresponding inference is made.

According to the conditions (10) and (11), quantifiers 'range over' items which are thought of as named by representative names (which are assigned to variables). A straight-forward substitutional interpretation would have been possible at the expense of much greater complexity.

Some further definitions: Sentences shall be *true* in a model just in case they are satisfied by it. A formula is *valid* exactly in case that it is satisfied by every model. A set K of formulas *yields* a formula ϕ if and only if ϕ is satisfied by every model which satisfies all members of K.

Since we have a proper class of names and of variables, a set K could never comprise all instances of a quantified statement; and every set of such instances which has a model will also have a model whose langugage is richer than that of the set. For these reasons, the semantical counterpart of omega-completeness fails, even though we give to quantifiers something akin to a substitutional interpretation.

5. Semantical Adequacy

Our system is semantically adequate in the strong sense: given any set K of formulas and any formula ϕ, ϕ is derivable from K if and only if K yields ϕ. In order to minimize technical details, we shall rest content in merely listing the main lemmas which need to be proved and in giving the barest outline of the completeness argument.

The proofs of the following three lemmas proceed by induction on the logical complexity of θ:

Lemma 1. Suppose that (a) θ is a well-formed expression, (b) α is a variable which is not free in θ, (c) $M = \langle T, B, K, I, A, H, R\rangle$ is a model, and (d) β is a variable in the domain of R. Then,

(1) if θ is a term of T, then $\mathrm{Val}(M, \theta) = \mathrm{Val}(M^{\alpha}_{R(\beta)}, \theta)$.

(2) if θ is a formula of T, then M sat θ just in case $M^{\alpha}_{R(\beta)}$ sat θ, and

(3) $\theta\{R\} = \theta\{R^{\alpha}_{R(\beta)}\}$.

Lemma 2. Suppose that (a) θ is a well-formed expression, (b) $M = \langle T, B, K, I, A, H, R\rangle$ is a model, (c) α is an n-term sequence of distinct variables, (d) β is an n-term sequence of variables in the domain of R, (e) $S = R^{\alpha_0}_{R(\beta_0)} \dots R^{\alpha_{n-1}}_{(\beta_{n-1})}$i, (f) $N = \langle T, B, K, I, A, H, S\rangle$, (g) d is the identity relation confined to the variables in θ, and (h) $\theta' = \theta\{d^{\alpha_0 \dots \alpha_{n-1}}_{\beta_0 \dots \beta_{n-1}}\}$ which is proper. Then,

(1) if θ is a term of T, then $\mathrm{Val}(M, \theta') = \mathrm{Val}(N, \theta)$,

(2) if θ is a formula of T, then M sat θ' just in case N sat θ, and
(3) $\theta\{S\} = \theta'\{R\}$.

Lemma 3. Suppose that (a) θ is a well-formed expression, (b) $M = \langle T, B, K, I, A, H, R\rangle$ is a model, (c) α and β are variables, (d) θ' results from θ by replacing one or more free occurrences of α by free occurrences of β, and (e) $R(\alpha) = R(\beta)$. Then,

(1) if θ is a term of T, then $\mathrm{Val}(M, \theta) = \mathrm{Val}(M, \theta')$,
(2) if θ is a formula of T, then M sat θ just in case M sat θ', and
(3) $\theta\{R\} = \theta'\{R\}$.

Using these three lemmas, it is not difficult to prove that each of the axioms is valid and that validity is preserved under the rules of inference, thereby establishing the soundness of the system. In order to prove its completeness, a number of further lemmas are needed. Most of those are familiar from completeness proofs of ordinary predicate calculus, and their derivation in our system requires at most slight variations. For this reason, we shall only list those lemmas which are either a bit different from customary ones or are not usually needed in demonstrating completeness:

Lemma 4. If ζ and η are terms, then $[\zeta = \eta \rightarrow (\zeta = \zeta \,\&\, \eta = \eta)]$ is a theorem.

Lemma 5. If ζ is a term and α is a variable which is not free in ζ, then $[\zeta = \zeta \leftrightarrow \exists \alpha \zeta = \alpha]$ is a theorem.

Lemma 6. If K is a consistent set of formulas then there exists a consistent set K' of formulas such that K is included in K' and for all formulas of K' [i.e. in the language of K'] of the form $\forall \alpha \phi$ there is a variable β such that β does not occur in $\forall \alpha \phi$ and $[(\phi\{d^{\alpha}_{\beta}\} \vee \sim \beta = \beta) \rightarrow \forall \alpha \phi]$ is in K' [where d is identity confined to the variables in ϕ].

Lemma 7. Suppose that ζ is a constant, α is a variable, and ϕ^{ζ}_{α} is the result of replacing in ϕ all occurrences of ζ by free occurrences of α (if possible, and ϕ otherwise). Then, if ϕ is a theorem, so is ϕ^{ζ}_{α}. (The argument proceeds by induction on the length of the proof of ϕ).

Lemma 8. Suppose that K is a consistent set of formulas, f is a one-one mapping of the free variables of K onto a set of constants not occurring in any member of K, and L is the set of all biconditionals $[\phi \leftrightarrow \phi\{d^{\alpha_0 \dots \alpha_{n-1}}_{f(\alpha_0) \dots f(\alpha_{n-1})}\}]$, where ϕ is a formula of K, d is identity restricted to the variables in ϕ, and $\alpha_0, \dots, \alpha_{n-1}$ are distinct variables comprising some or all of the variables which are free in ϕ. Then the union of K and L is consistent. (Suppose not, consider, finite subsets of K and L, and use Lemma 7).

Lemma 9. Every consistent set of formulas is included in a set of formulas which is maximally consistent with respect to the formulas of K.

Theorem [Strong Completeness]. For every consistent set Γ of formulas there exists a model statisfying all members of Γ.

In its barest essentials, the Henkin-type completeness proof is as follows:

(1) By Lemma 7, there is a consistent extension Γ' of Γ such that for all formulas of Γ' of the form $\forall\alpha\phi$ there is a variable β not occurring in $\forall\alpha\phi$ such that $[(\phi\{d^{\alpha}_{\beta}\} \vee \sim\beta=\beta) \rightarrow \forall\alpha\phi]$ is in Γ'.

(2) Let f be a one-one mapping of the free variables of Γ' onto a set of constants new to Γ'.

(3) Let L be the set of all biconditionals of the form $[\phi \leftrightarrow \phi \times \times \{d^{\alpha_0 \ldots \alpha_{n-1}}_{f(\alpha_0) \ldots f(\alpha_{n-1})}\}]$ where ϕ is a formula of Γ' and $\alpha_0, \ldots, \alpha_{n-1}$ are distinct variables comprising some or all of the variables which are free in ϕ.

(4) By Lemma 9, the union of Γ' and L is consistent.

(5) By Lemma 10, there is an extension Γ^* of the union of Γ' and L such that Γ^* is maximally consistent with respect to the formulas of that union.

(6) Let $c_0, c_1, \ldots$ be an enumeration of all constants in the range of f.

(7) For every term τ, let $\bar{\tau}$ = the least indexed constant c_i such that $[c_i = \tau]$ is in Γ^* (if there is such a constant; and 0 otherwise).

(8) Let $M = \langle T, B, K, I, A, H, R\rangle$ satisfy the following conditions:
 (a) $T = \Gamma^*$,
 (b) R is that function whose domain is the set of all variables and constants τ such that $[\tau = \tau]$ is in Γ^*, and for every such term τ, $R(\tau) = \bar{\tau}$,
 (c) B = the set of all $\phi\{R\}$, where Bϕ is in Γ^*,
 (d) K = the set of all $\phi\{R\}$, where Kϕ is in Γ^*,
 (e) I = the set of all sequences $\langle\phi_0\{R\}, \ldots, \phi_{n-1}\{R\}\rangle$, where $\mathrm{I}(\phi_0, \ldots, \phi_{n-2}; \phi_{n-1})$ is in Γ^*,
 (f) A = the set of all $\zeta\{R\}$ such that for some term τ, $[\mathrm{K}(\tau = \zeta)\ \&\ \mathrm{A}\tau]$ is in Γ^*, and
 (g) H = the set of all sentences $\pi\overline{\tau_0} \ldots \overline{\tau_{n-1}}$, where π is an n-place predicate and $\pi\tau_0 \ldots \tau_{n-1}$ is in Γ^*.

(9) For every term τ, if $[\tau = \tau]$ is in Γ^*, then there is a constant c_i such that $[c_i = \tau]$ is in Γ^*. For: by Lemma 6, $\exists\alpha\tau = \alpha$ is in Γ^*, by (1)

$[\tau = \beta]$ is in Γ^* for some new variable β, by (3) and (6), $[\tau = \beta \leftrightarrow \tau = = c_i]$ is in Γ^* for some c_i, and hence $[c_i = \tau]$ is in Γ^*.

(10) Claim: for every variable α, if $[\alpha = \alpha]$ is in Γ^* then $[\alpha = f(\alpha)]$ is in Γ^*. The proof is easy by appealing to (3) and the transitivity of identity.

(11) Claim: if α is a free variable in ϕ and α is in the domain of R, then $[\phi \leftrightarrow \phi\{d^{\alpha}_{R(\alpha)}\}]$ is in Γ^*. The claim can be proved by repeated use of (3), (10), and (Ax12).

(12) Claim: M is a model. The proof of this claim is lengthy and appeals to most of the axioms. Note especially that due to (11) we can freely interchange free variables for their representative names in Γ^*, and that the items in B and K are indeed (closed) sentences due to (Ax15).

(13) Claim: if θ is a term of Γ^* then $\mathrm{Val}(M, \theta) = \bar{\theta}$, and if θ is a formula of Γ^* then M sat θ just in case θ is in Γ^*. The proof of this claim proceeds by induction on the complexity of θ in more or less the usual manner.

These hints may suffice to indicate how the theorem is proved.

It appears that our system is semantically adequate and in intuitive agreement with the informal remarks made in Section 1.

Department of Philosophy, University of Rochester.

BIBLIOGRAPHY

[1] R. A. Eberle, 'Denotationless Terms and Predicates Expressive of Positive Qualities', in *Theoria* **35**, part 2 (1969), 104–123.

[2] J. Hintikka, *Knowledge and Belief*, Cornell University Press, 1962.

[3] J. Hintikka, 'Semantics for Propositional Attitudes', in *Models for Modalities*, D. Reidel Publishing Co., Dordrecht-Holland, 1969, 87–111.

[4] J. Hintikka, 'The Semantics of Modal Notions and the Indeterminacy of Ontology', in *Synthese* **21** (1970), 408–424.

[5] J. Hintikka, 'Knowledge, Belief, and Logical Consequence', in *Ajatus* **32** (1970), 32–47.

[6] D. Kaplan, 'Quantifying In', from *Words and Objections: Essays on the Work of W. V. Quine* (ed. by D. Davidson and J. Hintikka), D. Reidel Publishing Co., Dordrecht-Holland, 1969, 178–214.

[7] E. J. Lemmon, 'New Foundations for Lewis Modal Systems', in the *Journal of Symbolic Logic* **22**, No. 2 (1957), 176–186.

[8] R. Montague, 'Pragmatics and Intensional Logic', in *Synthese* **22**, Nos. 1/2 (1970), 68–94.

[9] R. L. Purtill, 'Believing the Impossible', in *Ajatus* **32** (1970), 18–24.

[10] W. V. Quine, 'Quantifiers and Propositional Attitudes', in *The Ways of Paradox*, Random House, New York, 1966, 183–194.