Notre Dame Journal of Formal Logic
Volume XX, Number 2, April 1979
NDJFAM

# AN ALTERNATIVE SEMANTICS FOR KNOWLEDGE

GREGORY MELLEMA

In 1962, with the publication of *Knowledge and Belief* ([6]), Jaakko Hintikka proposed an ingenious model set theoretic semantics for knowledge, **KB**. A somewhat unusual feature of **KB**, according to which '$Kap \rightarrow p$' is valid but '$(x)KaFx \rightarrow (x)Fx$' is not, has been widely criticized.[1] It has been felt that by reading '$(x)KaFx$' as "Everything known to $a$ is known by $a$ to be $F$" (as Hintikka does) rather than "Everything is known by $a$ to be $F$" (which entails "Everything is $F$") an unnatural restriction is placed upon quantifiers ranging into epistemic contexts. Consequently, a number of modifications of **KB** have in recent years been proposed which manage to avoid this feature.[2] In what follows I shall propose still another modification **KB*** of **KB** which lacks the Restricted Range feature (as I shall call the invalidity of '$(x)KaFx \rightarrow (x)Fx$' in a situation where '$Kap \rightarrow p$' is valid). I shall argue that **KB*** avoids a number of troublesome theorems which show up in these other proposed systems and that the intuitions underlying **KB***, while slightly different from those underlying **KB**, are every bit as natural. Consider:

(1) $(x)Ka(x = x)$
(2) $(x)(\exists y)Ka(x = y)$
(3) $(x)(y)(x = y \rightarrow Ka(x = y))$
(4) $(x)Ka(Fx \vee - Fx)$.

It can easily be shown that all of these formulas are theorems in **KB**. Now in **KB** (1)-(4) are read, respectively, as:

(5) Everyone known to $a$ is known by $a$ to be self-identical,
(6) For everyone known to $a$ there is someone known to $a$ such that they are known by $a$ to be identical,
(7) Any two persons known to $a$ are, if identical, known by $a$ to be identical
(8) Everyone known to $a$ is known by $a$ to be either $F$ or not-$F$.

(Here for purposes of simplicity we assume that we are dealing with domains of *persons*.) It may be a matter of dispute whether these readings

*Received February 2, 1976*

are counter-intuitive. While it is my belief that (7) is somewhat counter-intuitive, it is not clear that all of the others are; (6), for example, seems on the surface to be entirely plausible. Now consider:

(9) Everyone is known by $a$ to be self-identical
(10) For everyone $x$ there is someone known by $a$ to be identical with $x$
(11) Any two persons are, if identical, known by $a$ to be identical
(12) Everyone is known by $a$ to be either $F$ or not-$F$.

These are, respectively, the readings of (1)-(4) in each of the proposed variations of **KB** which lack the Restricted Range feature. Clearly, (9)-(12) are far more counter-intuitive than (5)-(8). While (7) may be mildly objectionable, its implausibility is negligible in comparison with (11), which, according to Føllesdal, says that $a$ " . . . knows the right answers to all questions of identity" (Føllesdal, [4], p. 15). And it is a consequence of (9), (10), and (12) that $a$ has what philosophers have called "*de re* knowledge" regarding everyone who exists.

It would seem unfortunate, therefore, for (1)-(4) to show up as theorems in variations of **KB** which lack the Restricted Range feature. Indeed, it is open for Hintikka to argue that if the theoremhood of (1)-(4) is an unavoidable feature of a Hintikka-type semantics, then one is really better off *not* to avoid the Restricted Range feature (bearing in mind that the formulas in **KB**, which *are* the formalizations of (9)-(12), are not theorems in **KB**). Yet none of the modifications of **KB** minus Restricted Range, which have been proposed, manage to avoid the theoremhood of (1)-(4).[3]

Can a plausible system be produced after the fashion of **KB** which not only lacks Restricted Range but which fails to make theorems out of (1)-(4)? Let us begin by investigating just why systems based upon **KB** are so prone to make theorems out of (1)-(4) in the first place. Suppose we concentrate for the time being upon (1) and see how it's proven in **KB**; we utilize Hintikka's method of *reductio* proof, demonstrating that '$-(x)Ka(x = x)$' can belong to no model set $w$. Recall that for Hintikka '$Pa$' is an abbreviation for '$-Ka-$'.

| | | | |
|---|---|---|---|
| (a) | $-(x)Ka(x = x)$ | $\epsilon\ w$ | Counterassumption |
| (b) | $(\exists x)Pa(x \neq x)$ | $\epsilon\ w$ | (a), (**C**.-**U**), (**C**.-**K**) |
| (c) | $Pa(b \neq b)$ | $\epsilon\ w$ | (b), (108), for some '$b$' |
| (d) | $(\exists x)Ka(b = x)$ | $\epsilon\ w$ | (b), (108), for some '$b$' |
| (e) | $b \neq b$ | $\epsilon\ w^*$ | (c), (**C**.**P***), where $w^*$ is an alternative of $w$ for $a$ |
| (f) | $b = b$ | $\epsilon\ w^*$ | (**C**.self. =) |

The crucial move in this proof is clearly the inference from (b) to (c) and (d) by way of rule (108). Let us closely examine the rule which allows this inference. Hintikka's rule (108) is usually thought of as a license which allows one to move from formulas of the form '$(\exists x)KaFx$' to '$KaFb$' and '$(\exists x)Ka(b = x)$', for some '$b$', in some given model set $w$. But it is at the same time a license which allows one to instantiate upon formulas of the form '$(\exists x)PaFx$':

$$\frac{(\mathrm{E}x)PaFx \quad \in w}{PaFb \quad \in w}$$
$$(\mathrm{E}x)Ka(b = x) \in w, \text{ for some } `b'.$$

This latter consequence of rule (108) has been the object of criticism. Tienson, for example, writes:

> But [(108)] also says that if $(\mathrm{E}x)PaFx$ holds, then two other formulas hold, $PaFb$ and $(\mathrm{E}x)Ka(b = x)$ for some constant $b$. And this is entirely unreasonable. For it says that if there is someone who $a$ does not know to be non-$F$, then $a$ has a unique way of referring to that person. Since constant terms have descriptive content for Hintikka, this means that if there is someone of whom you are ignorant in some respect, then you know something unique about him. But, I believe, by our ordinary understanding there can be individuals of whom one knows nothing at all. ([12], pp. 5-6)

And according to Sleigh:

> What [(108)] seems to say when applied to '$(\mathrm{E}x)$-$KaFx$' is that anyone $a$ doesn't know to be $F$ is among the persons known to $a$. This seems obviously unacceptable. ([10], p. 6)

There is no doubt that the inference:

$$\frac{(\mathrm{E}x)PaFx \quad \in w}{(\mathrm{E}x)Ka(b = x) \in w}$$

is highly counter-intuitive. But notice that the proof of (1) does not rely upon this particular inference (nor the proofs of (3) and (4)). Rather, it is the other instantiation of '$(\mathrm{E}x)PaFx$' licensed by (108),

$$\frac{(\mathrm{E}x)PaFx \in w}{PaFb \quad \in w,}$$

(let us call it "(**)") which is of concern in the present context of discussion, for *all* of the formulas (1)-(4) depend upon it to be proven in **KB**. It is the crucial move by which one in enabled to get rid of the quantifier(s) in order to eventually reach the contradiction buried inside the scope of '$Pa$'. Therefore, without (**) none of the formulas in question can be proven in **KB**. In like fashion, it is only through the employment of the rules similar to (**) that the same theorems turn up in each of the other systems which have been proposed.[4]

In order, therefore, to find a way to block proofs of (1)-(4), it looks as though rules for instantiating '$(\mathrm{E}x)PaFx$' must be pretty radically revised. According to Sleigh:

> What is required is one set of instantiating conditions for formulas of the form '$(\mathrm{E}x)Ka\phi$' and a distinct set for formulas of the form '$(\mathrm{E}x)Pa\phi$'. It is possible to develop a tenable model set approach along these lines which is free of the restricted range feature of **KB** but the resulting system is surprisingly complicated. ([10], p. 6)

A way must be found to grant instantiation upon '$(\mathrm{E}x)PaFx$' only under certain specified conditions and not in the same indiscriminate fashion accorded to '$(\mathrm{E}x)KaFx$'. Can such a technique be found? And would such a technique, if it could be found, make good sense not only on a formal level but relative to an intuitive possible worlds interpretation as well?

I shall now proceed to propose a system of model sets **KB*** involving just such a technique. Consider the following quantifier rules for contexts in which variables occurring inside the scope of a *single* epistemic operator are bound from the outside by quantifiers.

$$(\mathbf{EK})\quad \begin{array}{ll} (\mathrm{E}x)KaFx & \in w \\ \hline (\mathrm{E}x)Ka(b = x) & \in w \\ KaFb & \in w, \text{ for some `}b\text{'} \end{array}$$

$$(\mathbf{UK})\quad \begin{array}{ll} (x)KaFx & \in w \\ (\mathrm{E}x)(b = x) & \in w \\ \hline (\mathrm{E}x)(b = x \;\&\; KaFx) & \in w \end{array}$$

$$(\mathbf{EPL})\quad \begin{array}{ll} (\mathrm{E}x)PaFx & \in w \\ \hline (\mathrm{E}x)(b = x) & \in w \\ (\mathrm{E}x)(b = x \;\&\; PaFx) & \in w, \text{ for some `}b\text{'} \end{array}$$

$$(\mathbf{EPS})\quad \begin{array}{ll} (\mathrm{E}x)PaFx & \in w \\ (x)KaGx & \in w \\ \hline PaFb & \in w, \text{ for some `}b\text{'} \end{array}$$

$$(\mathbf{EPM})\quad \begin{array}{ll} (\mathrm{E}x)(Gx \;\&\; PaFx) & \in w \\ (x)KaHx & \in w \\ \hline Gc \;\&\; PaFb & \in w, \text{ for some `}b\text{', `}c\text{'} \end{array}$$

$$(\mathbf{UPL})\quad \begin{array}{ll} (x)PaFx & \in w \\ (\mathrm{E}x)(b = x) & \in w \\ \hline (\mathrm{E}x)(b = x \;\&\; PaFx) & \in w \end{array}$$

$$(\mathbf{UPS})\quad \begin{array}{ll} (x)PaFx & \in w \\ (\mathrm{E}x)Ka(b = x) & \in w \\ \hline PaFb & \in w. \end{array}$$

We assume that there are no epistemic operators in '$F$' or '$G$' and that (for the sake of simplicity) negation signs have been driven through the epistemic operators. Following the tradition of others, we list the rules in a paradigmatic form and explain their use as follows. Rules (**EK**) and (**UK**) apply to *any* formula in which '$x$' occurs free within the scope of *any* '$Ka$' operator. In this way, from '$(\mathrm{E}x)(Ax \;\&\; (-Bx \vee Ka(Cx \;\&\; -Dx) \;\&\; KaEx)) \in w$', we may infer both '$(\mathrm{E}x)Ka(b = x) \in w$' and '$(Ab \;\&\; (-Bb \vee Ka(Cb \;\&\; -Db) \;\&\; KaEb)) \in w$', for some '$b$'. The other rules apply to all formulas in which '$x$' is bound by at least one '$Pa$' operator but nowhere by a '$Ka$' operator. Thus, '$(\mathrm{E}x)(Fx \vee KaGx)$', '$(\mathrm{E}x)(KaFx \;\&\; KaGx)$', and '$(\mathrm{E}x)(KaFx \;\&\; PaGx)$' fall under (**EK**) while '$(\mathrm{E}x)(Fx \;\&\; PaGx)$' and '$(\mathrm{E}x)(PaFx \;\&\; PaGx)$' fall under (**EPL**), (**EPS**), or (**EPM**). Rule (**EPM**) is a qualification to (**EPS**) and applies to formulas in which '$x$' occurs both within and without the scope of '$Pa$' operators (but nowhere within the scope of '$Ka$').

Our system **KB*** may now be formed as follows. Begin with the rules of **KB**. Replace Hintikka's (108) and (109)

$$(108)\quad \begin{array}{ll} (\exists x)KaFx & \epsilon\ w \\ \hline KaFb & \epsilon\ w \\ (\exists x)Ka(b = x) & \epsilon\ w \end{array} \qquad \begin{array}{ll} (\exists x)PaFx & \epsilon\ w \\ \hline PaFb & \epsilon\ w \\ (\exists x)Ka(b = x) & \epsilon\ w \end{array}$$

$$(109)\quad \begin{array}{ll} (x)KaFx & \epsilon\ w \\ (\exists x)Ka(b = x) & \epsilon\ w \\ \hline KaFb & \epsilon\ w \end{array} \qquad \begin{array}{ll} (x)PaFx & \epsilon\ w \\ (\exists x)Ka(b = x) & \epsilon\ w \\ \hline PaFb & \epsilon\ w \end{array}$$

by the seven rules listed above. Drop from **KB** the rules:

$$(\mathbf{C.EK.{=}*})\quad \begin{array}{l} (\exists x)Ka(b = x)\ \epsilon\ w \\ \hline Ka(\exists x)(b = x)\ \epsilon\ w \end{array}$$

$$(\mathbf{C.EK.EK.{=}*})\quad \begin{array}{l} (\exists x)Ka(b = x)\ \epsilon\ w \\ \hline (\exists x)Ka(b = x)\ \epsilon\ w^*, \text{ for every alternative } w^* \text{ of } w \text{ for } a.\text{[5]} \end{array}$$

And for purposes of convenience add the rules:

$$(\mathbf{C.E-})\quad \begin{array}{ll} (\exists x)Fx & \epsilon\ w \\ \hline -(x)-Fx & \epsilon\ w \end{array}$$

$$(\mathbf{C.U-})\quad \begin{array}{ll} (x)Fx & \epsilon\ w \\ \hline -(\exists x)-Fx & \epsilon\ w \end{array}$$

$$(\mathbf{C.K-})\quad \begin{array}{ll} Kap & \epsilon\ w \\ \hline -Pa-p & \epsilon\ w \end{array}$$

$$(\mathbf{C.P-})\quad \begin{array}{ll} Pap & \epsilon\ w \\ \hline -Ka-p & \epsilon\ w. \end{array}$$

And to all of this we add a set of rules, to be described in what follows, to govern formulas formed by quantifying in contexts two or more layers deep of epistemic operators.

Our first order of business is to establish that our system lacks the Restricted Range feature. For this it suffices to show that

$$\begin{array}{ll} (x)KaFx & \epsilon\ w \\ \hline (x)Fx & \epsilon\ w \end{array}$$

is provable in **KB***:

| | | | |
|---|---|---|---|
| (a) | $(x)KaFx$ | $\epsilon\ w$ | Counterassumption |
| (b) | $-(x)Fx$ | $\epsilon\ w$ | |
| (c) | $(\exists x)-Fx$ | $\epsilon\ w$ | (b), (**C.-U**) |
| (d) | $-Fb$ | $\epsilon\ w$ | (c), ($\mathbf{C.E_0}$), for some '$b$' |
| (e) | $(\exists x)(b = x)$ | $\epsilon\ w$ | (c), ($\mathbf{C.E_0}$) |
| (f) | $(\exists x)(b = x \ \&\ KaFx)$ | $\epsilon\ w$ | (a), (e), (**UK**) |
| (g) | $b = c \ \&\ KaFc$ | $\epsilon\ w$ | (f), (**EK**) |
| (h) | $(\exists x)Ka(c = x)$ | $\epsilon\ w$ | |
| (i) | $KaFc$ | $\epsilon\ w$ | (g), Simp |
| (j) | $Fc$ | $\epsilon\ w$ | (i), (**C.K***), Reflexivity |
| (k) | $b = c$ | $\epsilon\ w$ | (g), Simp |
| (l) | $-Fc$ | $\epsilon\ w$ | (d), (k), (**C.=**). |

Notice that a similar proof cannot be produced in **KB**; one may proceed as far as step (e), but the move from (a) and (e) to (f) is not allowed in **KB**.

As far as instantiation on '$(x)KaFx$' is concerned, our system is actually more liberal than **KB**. Our rule **(UK)** allows instantiation on '$(x)KaFx$', in a sort of large-scope fashion, with the auxiliary clause '$(\mathrm{E}x)(b = x)$'; **KB** has no such provision. At the same time, Hintikka's (small-scope) instantiation rule

$$\frac{\begin{array}{ll} (x)KaFx & \epsilon\ w \\ (\mathrm{E}x)Ka(b = x) & \epsilon\ w \end{array}}{\begin{array}{ll} KaFb & \epsilon\ w \end{array}}$$

is a consequence of our rules:

| | | | |
|---|---|---|---|
| (a) | $(x)KaFx$ | $\epsilon\ w$ | Counterassumption |
| (b) | $(\mathrm{E}x)Ka(b = x)$ | $\epsilon\ w$ | |
| (c) | $-\ KaFb$ | $\epsilon\ w$ | |
| (d) | $(\mathrm{E}x)(b = x)$ | $\epsilon\ w$ | (b), **(C.EK.=)** |
| (e) | $(\mathrm{E}x)(b = x\ \&\ KaFx)$ | $\epsilon\ w$ | (a), (d), **(UK)** |
| (f) | $b = c\ \&\ KaFc$ | $\epsilon\ w$ | (e), **(EK)** |
| (g) | $(\mathrm{E}x)Ka(c = x)$ | $\epsilon\ w$ | |
| (h) | $KaFc$ | $\epsilon\ w$ | (f), Simp |
| (i) | $b = c$ | $\epsilon\ w$ | (f), Simp |
| (j) | $Ka(b = c)$ | $\epsilon\ w$ | (b), (g), (i) |
| (k) | $-\ KaFc$ | $\epsilon\ w$ | (c), (j), (97). |

As the other end of the spectrum, however, our instantiation rules for '$Pa$' are much less liberal than in **KB**. This, of course, is all a part of our attempt to block proofs of formulas (1)-(4). If we try, for example, to prove (1) by assuming that '$(\mathrm{E}x)Pa(x \neq x)$' belongs to a model set $w$, we can infer both '$(\mathrm{E}x)(b = x)\ \epsilon\ w$' and '$(\mathrm{E}x)(b = x\ \&\ Pa(x \neq x))\ \epsilon\ w$'. But we can infer nothing that yields a contradiction; we lack the means to peel off the quantifier and get at what is inside the epistemic operator (the means of which is present in all of the other systems).

Thus our rules succeed very nicely in blocking the proofs of the unwanted theorems by making at least one of Hintikka's key quantifier rules drastically weaker. But by introducing a set of rules which blocks the proofs of the unwanted theorems might we not have produced a system which fails to make good sense relative to an intended possible worlds interpretation of formulas in **KB***? Perhaps the natural intuitions which underlie Hintikka's rules taken as a whole will be, once these rules have been tampered with, gone for good, and we shall be left with a new set of rules which are artificial, contrived, and unable to be backed with the authority of an intuitive model.

In order to describe in an absolutely precise fashion the intuitions underlying **KB*** and the ways in which they differ from those of **KB** we examine in the appendix a higher-order theory $W$. Here we shall discuss informally a number of these underlying intuitions. The basic idea, as described by Hintikka, is this:

> Existential generalization with respect to a term—say $b$—is admissible in (epistemic) contexts if $b$ refers to one and the same man in all the "possible worlds" we have to consider.... Now clearly $b$ refers to one and the same man in all these states of affairs if there is someone who is known by the bearer of $a$ to be referred to by $b$ .... In short, $b$ refers to one and the same man in all the "possible worlds" we have to consider in this special case if it is true to say "$a$ knows who $b$ is". But this is exactly what our solution amounts to here: existential generalization with respect to $b$ is admissible if we have as an additional premise the sentence "$a$ knows who $b$ is", formally "$(\text{E}x)Ka(b = x)$". ([6], pp. 152-153)

An individual $b$ is known by $a$ just in case '$b$' refers to the same individual in each possible world compatible with everything $a$ knows. We can imagine a line drawn connecting each of these individuals from world to world. Since such a line is said to pick out "the same" individual in each compatible world, we might call a world line which connects individuals in this fashion "rigid". It would then be natural to interpret the formulas of **KB** and **KB*** in such a way that an individual is known to $a$ in a world $w$ just in case one of $a$'s rigid world lines picks up that individual in $w$ (for simplicity let us assume $a$ is the only agent).

Given such an interpretation, the differences in the intuitions underlying **KB** and **KB*** are as follows: First, in **KB** a rigid world line picks out an individual in a world only if the individual exists in that world (Hintikka later dropped this restriction from **KB**—see footnote 5), whereas this is not so in **KB***. And, second, because **KB*** lacks Restricted Range and **KB** does not, the truth of '$(x)KaFx$' in a world does not in **KB** necessitate that every existent individual in that world is picked up by a rigid world line, as is the case in **KB***. Thus, '$(x)KaFx \in w$' is interpreted in **KB*** as

(13) Anything which exists in $w$ is picked up by a rigid world line which in every compatible world picks up something which is $F$ in that world.

and in **KB** as

(14) If anything exists in $w$ and is picked up by a rigid world line, this world line in every compatible world picks up something which exists and is $F$ in that world.

The formula '$(x)KaFx \in w$' asserts in **KB***, therefore, that

(15) For any $x$ which exists in $w$, (i) $x$ is picked up by a rigid world line, and (ii) This world line picks up something in each compatible world which is $F$ in that world.

Similarly, '$(\text{E}x)KaFx \in w$' asserts that

(16) For *some* $x$ which exists in $w$, conditions (i) and (ii) hold.

Thus '$(x)PaFx \in w$' (because it is equivalent to '$(x) - Ka - Fx \in w$') asserts

(17) For any $x$ which exists in $w$, either (i′) $x$ is not picked up by a rigid world line, or (ii′) The rigid world line which picks up $x$ does *not* in every compatible world pick up something which is not-$F$ in that world,

and '$(\mathrm{E}x)PaFx \in w$' asserts

(18) For *some* $x$ which exists in $w$, either (i′) or (ii′) holds.

The intuitions which underlie the rules of **KB*** are now easy to explain. **(EK)**, also valid in **KB**, simply licenses us to start with (16) and perform existential instantiation upon it, first upon condition (i), then upon condition (ii) to a constant '$b$' which names the object picked out by the rigid world line at every compatible world. **(UK)** tells us that (15) may be instantiated to any constant '$b$', provided $b$ exists in $w$. **(EPL)** states that existential instantiation may be performed upon (18). **(EPS)** asserts that from (18) and (15), with '$F$' in (15) replaced by any predicate constant whatever, it follows that for some '$b$', $b$ is $F$ in at least one world $w^*$ compatible with $w$ (this may be checked out as valid). **(EPM)** is merely a qualification to **(EPS)** to the effect that if $b$ is the individual which is $F$ in $w^*$, the individual picked out in $w$ by the same rigid world line might not be referred to as '$b$' in $w$. **(UPL)** licenses us to instantiate (17) to '$b$', provided $b$ exists in $w$. And **(UPS)** states that from (17), the assumption that $b$ exists in $w$, that $b$ fulfills condition (i), and that the rigid world line which picks up $b$ in $w$ picks up $b$ in every compatible world (we do not assume that this condition is automatically satisfied), we may infer that $b$ is $F$ in at least one world compatible with $w$ (this too may be checked out to be valid).

The failure of (**) to be valid in **KB*** may now be explained as follows. If we start with (18) and try to show that some individual $b$ is $F$ in at least one world compatible with $w$, we must somehow show that some individual which exists in $w$ ($c$, for example) fulfills condition (ii′). But (18) tells us only that some such individual fulfills either (i′) or (ii′). We need to know in addition that $c$ fails to fulfill condition (i′). In the case of **(EPS)** this additional information comes by way of assuming (15), which guarantees that every individual existent in $w$ fulfills (i) and hence fails (i′). But without this extra information the demonstration cannot proceed, and (**) is invalid.

So far we have restricted our attention to formulas in **KB*** in which variables occur within the scope of, at most, one epistemic operator. We now present a set of rules general enough to cover all formulas in **KB*** formed by quantifying past several epistemic operators. The expressions '$Qa$' and '$Q'a$' designate arbitrary strings formed by concatenating '$Ka$' and '$Pa$'; the cases in which '$Qa$' and '$Q'a$' designate the empty string are those which involve quantifying past a single epistemic operator and yield the quantifier rules we listed earlier. The official version of our quantifier rules for **KB*** shall then be as follows:

$$\textbf{(EK)}\quad \frac{(\mathrm{E}x)KaQaFx \in w}{\begin{array}{l}(\mathrm{E}x)Ka(b = x) \in w\\ KaQaFb \in w, \text{ for some '}b\text{'}\end{array}}$$

$$\textbf{(UK)}\quad \frac{\begin{array}{l}(x)KaQaFx \in w\\ (\mathrm{E}x)(b = x) \in w\end{array}}{(\mathrm{E}x)(b = x \;\&\; KaQaFx) \in w}$$

**(EPL)**
$$\begin{array}{ll} (\mathrm{E}x)PaQaFx & \in w \\ \hline (\mathrm{E}x)(b = x) & \in w \\ (\mathrm{E}x)(b = x \ \&\ PaQaFx) & \in w, \text{ for some `}b\text{'} \end{array}$$

**(EPS)**
$$\begin{array}{ll} (\mathrm{E}x)PaQaFx & \in w \\ (x)KaGx & \in w \\ \hline PaQaFb & \in w, \text{ for some `}b\text{'} \end{array}$$

**(UPL)**
$$\begin{array}{ll} (x)PaQaFx & \in w \\ (\mathrm{E}x)(b = x) & \in w \\ \hline (\mathrm{E}x)(b = x \ \&\ PaQaFx) & \in w \end{array}$$

**(UPS)**
$$\begin{array}{ll} (x)PaQaFx & \in w \\ (\mathrm{E}x)Ka(b = x) & \in w \\ \hline PaQaFb & \in w \end{array}$$

**(EPE)**
$$\begin{array}{ll} (\mathrm{E}x)QaPaPaQ'aFx & \in w \\ \hline (\mathrm{E}x)QaPaQ'aFx & \in w \end{array}$$

**(EKE)**
$$\begin{array}{ll} (\mathrm{E}x)QaKaQ'aFx & \in w \\ \hline (\mathrm{E}x)QaQ'aFx & \in w \end{array}$$

**(EKI)**
$$\begin{array}{ll} (\mathrm{E}x)QaKaQ'aFx & \in w \\ \hline (\mathrm{E}x)QaKaKaQ'aFx & \in w \end{array}$$

**(EPI)**
$$\begin{array}{ll} (\mathrm{E}x)QaPaQ'aFx & \in w \\ \hline (\mathrm{E}x)QaPaPaQ'aFx & \in w. \end{array}$$

As before, the rules are formulated in a paradigmatic fashion for purposes of simplicity, but their range of application extends to *any* formula in **KB*** formed by quantifying past one or more epistemic operators. Thus, rule (**EK**) is not limited to formulas such as '$(\mathrm{E}x)KaKaPaKaFx$' but applies as well to something like '$(\mathrm{E}x)(\mathrm{E}y)((-Ax \vee By)\ \&\ Ka(Cx\ \&\ Pa(-By \vee Dx)))$'. Two stipulations, however, must be laid down in reference to these rules. The first is that rules (**EK**) and (**UK**) apply to *all* formulas in which '*x*' occurs free within the scope of some '*Ka*' operator itself not bound by any epistemic operators; only in cases where this condition is not met will rules (**EPL**), (**EPS**), (**UPL**), and (**UPS**) apply. Thus, '$(\mathrm{E}x)(KaFx\ \&\ PaGx)$', '$(\mathrm{E}x)(Fx\ \&\ KaPaPaGx)$', and '$(\mathrm{E}x)(PaKaFx\ \&\ KaGx)$' all fall under rule (**EK**) while '$(\mathrm{E}x)(PaKaKaFx)$' and '$(\mathrm{E}x)(PaFx\ \&\ PaKaGx)$' fall under (**EPL**) or (**EPS**).

Our second stipulation is this: In applying rule (**EPS**), a pair of free '*x*'s are to be instantiated to the same constant when and only when there is no occurrence of '*Pa*' binding one and not the other; in all other cases distinct constants must be used. Thus, while with (**EK**), '$(\mathrm{E}x)(Fx \vee KaGx)$' instantiates to '$Fb \vee KaGb$', '$(\mathrm{E}x)(Fx \vee PaGx)$' is instantiated to '$Fc \vee PaGb$' since one '*x*' is bound by an occurrence of '*Pa*' not binding the other. '$(\mathrm{E}x)(PaFx\ \&\ PaGx)$' too requires two constants (each '*x*' is bound by a '*Pa*' not binding the other), while '$(\mathrm{E}x)(Bx\ \&\ Pa(Cx\ \&\ PaDx)\ \&\ Ax)$' requires three and instantiates to something of the form '$Bc\ \&\ Pa(Cd\ \&\ PaDb)\ \&\ Ac$' (the two outside clauses take the same constant). In the case of rules (**EK**) and (**UPS**) every free occurrence of '*x*' instantiates to the same constant and no such stipulation need be observed.[6]

## APPENDIX

Let us turn our attention towards a possible worlds interpretation $\mathfrak{M}$ of the formulas in **KB***  and examine our rules in light of this interpretation. Rather than simply describe $\mathfrak{M}$, however, we shall follow Sleigh's lead and construct an interpreted, higher-order theory $W$ which talks about $\mathfrak{M}$ in that the truth-conditions for formulas of **KB*** relative to $\mathfrak{M}$ are made explicit.[7] In this way we can judge with a good deal more precision the adequacy of the rules of **KB***. Following Sleigh, we shall call $W$ the "world theory" of **KB***, and a formula in $W$ which describes the truth-conditions in $\mathfrak{M}$ of a formula '$A$' will be called the "world theory transcription" of '$A$' in $W$. Having done this, it will be very easy to judge whether **KB*** is an adequate system. Any inference

$$\frac{A \in w}{B \in w}$$

must be provable in **KB*** if and only if '$A^* \to B^*$' is provable in $W$, where '$A^*$' and '$B^*$' are the respective world theory transcriptions of '$A$' and '$B$'.

The language of $W$ can be described as follows:

(i) Every symbol in **KB*** except '$Ka$' and '$Pa$' belongs to the language of $W$;

(ii) In addition:

'$I^*$' is a one-place predicate constant in $W$
'$E$' is a two-place predicate constant in $W$
'$R$' is a two-place predicate constant in $W$
'$B_1$', '$B_2$', etc., are two-place predicate constants in $W$
'$P_1$', '$P_2$', etc., are two-place predicate variables in $W$
'$w_i$', '$w_j$', etc., are variables in $W$
'$w_r$' is a constant in $W$;

(iii) Atomic well-formed formulas are as follows:

(a) If '$A$' is a wff in **KB***, then if '$A$' contains no occurrences of '$Ka$', '$Pa$', or individual constant terms, '$(A)w_i$' is a wff in $w$, for any $i$
(b) '$I^*P_i$' is a wff in $W$, for any $i$
(c) '$Exw_i$' is a wff in $W$, for any $i$ and any variable or constant $x$ in $W$
(d) '$Rw_iw_j$' is a wff in $W$, for any $i,j$
(e) '$P_ixw_j$' is a wff in $W$, for any $i,j$ and any variable or constant $x$ in $W$
(f) '$B_ixw_j$' is a wff in $W$, for any $i,j$ and any variable or constant $x$ in $W$.

Intuitively, atomic formulas in $W$ are to be understood in the following way. '$(A)w_i$' asserts that a formula '$A$' of **KB*** is true in world $w_i$ (relative to $\mathfrak{M}$). '$I^*P_1$' asserts that a world line $P_1$ is "privileged" in the sense of "rigid" described earlier.[8] '$Exw_i$' is understood to say that $x$ exists in world $w_i$. '$P_1xw_j$' asserts that world line $P_1$ picks up object $x$ at world $w_j$. '$Rw_iw_j$' asserts that world $w_j$ is an alternative to $w_i$ (relative to $a$). And '$B_ixw_j$' is understood to say simply that $x$ is $B_i$ at $w_j$. In what follows we shall refer to the '$B_i$' predicates as "Badge Predicates".

I believe we now have the machinery to express in the language of $W$ the explicit truth-conditions of any formula in **KB*** relative to our intended model $\mathfrak{M}$. We now give a recursive procedure to show how to derive, for any formula '$A$' in **KB***, the world theory transcription '$A^*$' of '$A$' in $W$. We begin by defining the notion 'Immediate $W$-translation' ('I$W$-translation' for short). For purposes of abbreviation we shall shorten '$w_j$' to '$j$' throughout and for convenience we shall assume that '&' and '-' are the only truth-connectives in '$A$'. First, we suppose that '$A$' contains no individual constant terms from the language of **KB***.

(i) If $F$ is atomic, the I$W$-translation of $F$ is '$(F)i$'
(ii) *If $F$ is of the form '$B$ & $C$', the I$W$-translation of $F$ is '$(B)i$ & $(C)i$'*
(iii) If $F$ is of the form '$-B$', the I$W$-translation of $F$ is '$-((B)i)$'
(iv) If $F$ is of the form '$KaB$', then

(a) If every free variable in $B$ is bound *in* $B$ by at least one epistemic operator, the I$W$-translation of $F$ is '$(j)(Rij \rightarrow (B)j)$'
(b) If $x_1, \ldots, x_n$ occur free in $B$ outside the scope of all epistemic operators in $B$, the I$W$-translation of $F$ is '$(j)(Rij \rightarrow (\mathrm{E}y_1) \ldots (\mathrm{E}y_n)(P_1 y_1 j \ \& \ldots \& \ P_n y_n j \ \& \ (B(y_k/x_k))j))$', where $F$ occurs within the context '$\ldots (I^*P_1 \ \& \ P_1 x_1 i) \ldots \& \ldots . (I^*P_n \ \& \ P_n x_n i) \ldots \& \ldots F \ldots$'

(v) If $F$ is of the form '$PaB$', the I$W$-translation of $F$ is '$-Ka - B$'
(vi) If $F$ is of the form '$(\mathrm{E}x)B$', then

(a) If '$x$' does not occur free in $B$, the I$W$-translation of $F$ is '$(B)i$'
(b) If '$x$' occurs free in $B$ but never within the scope of an epistemic operator, the I$W$-translation of $F$ is '$(\mathrm{E}x)(Exi \ \& \ (B)i)$'
(c) If '$x$' occurs free in $B$ and for at least one occurrence of '$Ka$' not directly preceded by an odd number of negation signs, '$x$' occurs within its scope and it (the '$Ka$') occurs within the scope of no epistemic operator, then

(i) If $B$ is of the form '$-(C \ \& \ D)$' and '$x$' occurs *only* within the scope of epistemic operators, the I$W$-translation of $F$ is '$(\mathrm{E}x)(Exi \ \& \ -((\mathrm{E}P_1)(I^*P_1 \ \& \ P_1 xi \ \& \ (C)i) \ \& \ (\mathrm{E}P_2)(I^*P_2 \ \& \ P_2 xi \ \& \ (D)i)))$'
(ii) Otherwise the I$W$-translation of $F$ is '$(\mathrm{E}x)(Exi \ \& \ (\mathrm{E}P_1)(I^*P_1 \ \& \ P_1 xi \ \& \ (B)i))$'

(d) In all other cases the I$W$-translation of $F$ is '$(\mathrm{E}x)(Exi \ \& \ (P_1)((I^*P_1 \ \& \ P_1 xi) \rightarrow (B)i))$'

(viii) If $F$ is of the form '$(x)B$', the I$W$ translation of $F$ is '$-(\mathrm{E}x) - B$'.

Here we let '$i$' be a meta-linguistic variable whose value is determined as follows. If '$F$' is '$A$', then '$i$' stands for '$r$'. And if '$F$' is '$E$' in the procedure prescribed below, then '$i$' stands for the variable '$j$' in the quantifier '$(Ej)$' or '$(j)$' most directly binding '$E$'. If '$F$' is '$E$' and there is no such quantifier, then '$i$' stands for '$r$'.

We now say that a formula '$G$' is the '$W$-translation' of a formula '$F$' if and only if (i) There is a finite sequence of formulas $F_1, \ldots, F_n$ such that $F_1$ is the I$W$-translation of $F$, $F_{i+1}$ is the I$W$-translation of $F_i$, for all $i$, and $G$ is the I$W$-translation of $F_n$, and (ii) For every subformula in $G$ of the form '$(A)i$', '$A$' is atomic. On the basis of all of this we now calculate '$A^*$' in the following way. Let '$A_1$' be the Immediate $W$-translation of '$A$'. Now let '$D$' be the leftmost subformula in '$A_j$' which is of the form '$(E)i$' for some $i$ and non-atomic formula '$E$'. Then '$A_{j+1}$' is the formula obtained by replacing '$D$' by the $W$-translation of '$E$', if it has one, and by its Immediate $W$-translation otherwise. When a formula '$A_j$' is reached in which there is no such subformula '$D$', then '$A_j$' is '$A^*$'.

If '$A$' contains constant singular terms from the language of **KB***, we derive it as follows. Since **KB*** is a system of our own making there is no problem in assuming that for each constant singular term '$b$' in **KB*** there is a badge predicate in **KB*** which holds uniquely of $b$. Replace every subformula '$(\mathrm{E}x)C$' in '$B$' by '$(\mathrm{E}x)(Ex \ \& \ C)$', every subformula '$(x)C$' in '$B$' by '$(x)(Ex \rightarrow C)$', and every *atomic* subformula '$D$' in '$B$' which contains a constant singular term '$b$' by '$(\mathrm{E}x)(D(b/x) \ \& \ B_j x)$', where '$B_j$' is the badge predicate of '$b$' in **KB*** and '$E$' is not a symbol in **KB***. When all singular terms have been eliminated the result is '$A$'. Now go to the recursive procedure outlined above with the following revision of 'I$W$-translation'. Add to the first line of (vi), "... if $B$ is of the form '$Ex$ & $C$', then", and add "(vii) If $F$ is of the form '$(\mathrm{E}x)B$' and $B$ is not of the form '$Ex$ & $C$', then ..." followed by (a)-(d) of (vi) with the clause '$Exi$' deleted everywhere it occurs.

Let us now look at some examples of basic formulas and their respective world theory transcriptions in $W$. From now on we shall often abbreviate '$(B)i$' as '$Bi$'.

(i) $KaB$:
$(i)(Rri \to (B)i)$
(ii) $(\mathsf{E}x)KaBx$:
$(\mathsf{E}x)(Exr \,\&\, (\mathsf{E}P_1)((I^*P_1 \,\&\, P_1xr) \,\&\, (i)(Rri \to (\mathsf{E}y)(P_1yi \,\&\, Byi))))$
(iii) $(x)KaBx$:
$(x)(Exr \to (\mathsf{E}P_1)((I^*P_1 \,\&\, P_1xr) \,\&\, (i)(Rri \to (\mathsf{E}y)(P_1yi \,\&\, Byi))))$
(iv) $(\mathsf{E}x)(Cx \,\&\, KaBx)$:
$(\mathsf{E}x)(Exr \,\&\, (\mathsf{E}P_1)((I^*P_1 \,\&\, P_1xr) \,\&\, Cxr \,\&\, (i)(Rri \to (\mathsf{E}y)(P_1yi \,\&\, Byi))))$
(v) $(\mathsf{E}x)PaBx$:
$(\mathsf{E}x)(Exr \,\&\, (P_1)((I^*P_1 \,\&\, P_1xr) \to (\mathsf{E}i)(Rri \,\&\, (\mathsf{E}y)(P_1yi \,\&\, Byi))))$
(vi) $(\mathsf{E}x)(KaBx \,\&\, PaCx)$:
$(\mathsf{E}x)(Exr \,\&\, (\mathsf{E}P_1)((I^*P \,\&\, P_1xr) \,\&\, ((i)(Rri \to (\mathsf{E}y)(P_1yi \,\&\, Byi)) \,\&\, (\mathsf{E}i)(Rri \,\&\, (\mathsf{E}y)(P_1yi \,\&\, Cyi)))))$
(vii) $(\mathsf{E}x)KaKaBx$:
$(\mathsf{E}x)(Exr \,\&\, (\mathsf{E}P_1)((I^*P_1 \,\&\, P_1xr) \,\&\, (i)(Rri \to (j)(Rij \to (\mathsf{E}y)(P_1yj \,\&\, (Byj))))))$
(viii) $(\mathsf{E}x)(Ka(Bx \,\&\, KaCx))$:
$(\mathsf{E}x)(Exr \,\&\, (\mathsf{E}P_1)((I^*P_1 \,\&\, P_1xr) \,\&\, (i)(Rri \to ((\mathsf{E}y)(P_1yi \,\&\, Byi) \,\&\, (j)(Rij \to (\mathsf{E}y)(P_1yj \,\&\, Cyj))))))$
(ix) $(\mathsf{E}x)PaKaBx$:
$(\mathsf{E}x)(Exr \,\&\, (P_1)((I^*P_1 \,\&\, P_1xr) \to (\mathsf{E}i)(Rri \,\&\, (j)(Rij \to (\mathsf{E}y)(P_1yj \,\&\, Byj))))).$

It is a simple matter, incidentally, to assign world theory transcriptions in $W$ to formulas in Hintikka's **KB**. We need to make only two alterations to our definition of '$IW$-translation' as follows: Replace '$(j)(Rij \to (\mathsf{E}y_1)\ldots(\mathsf{E}y_n)(P_1y_1j \,\&\ldots\&\, P_ny_nj \,\&\, (B(y_k/x_k))i))$' by '$(j)(Rij \to (\mathsf{E}y_1)\ldots(\mathsf{E}y_n)((P_1y_1j \,\&\ldots\&\, P_ny_nj) \,\&\, (Ey_1j \,\&\ldots\&\, Ey_nj) \,\&\, (B(y_k/x_k))i)))$', and delete part (vi)-(c). The differences in the intuitions underlying the rules of **KB** and the rules of **KB***, therefore, are twofold. Privileged world lines in **KB** pick out individuals in worlds only if they happen to exist in such worlds (Hintikka actually has now dropped this restriction by dropping his rule (**C.EK.**=*)). And the truth of '$(x)KaFx$' in a world does not in **KB** necessitate that every existent individual in that world is picked up by a privileged world line, as is the case in **KB***.

Finally, axioms for $W$ are as follows:

(i) $(i)(Rii)$
(ii) $(i)(j)(k)((Rij \,\&\, Rjk) \to Rik)$
(iii) $(P_1)(i)(\mathsf{E}x)(P_1xi \,\&\, (y)(P_1yi \equiv P_1xi))$
(iv) $(P_1)(P_2)(x)(y)(i)(j)((P_1xi \,\&\, P_2xi \,\&\, Rij) \to (P_1yj \equiv P_2yj))$
(v) $\ulcorner(\mathsf{E}P_1)(i)(x)(P_1xi \equiv B_jxi)\urcorner$ is an axiom, for any badge predicate '$B_j$'.

With these axioms we can prove world theory transcriptions of the theorems of **KB***.[9]

## NOTES

1. See for example Castañeda [1], p. 134, Castañeda [2], pp. 9, ff., Clark [3], pp. 177, ff., Sellars [8], pp. 191, 195-198, and Stine [11], pp. 127-129. For an excellent discussion of these criticisms see Sleigh [9].

2. Føllesdal [4], pp. 11-13, Sleigh [9], pp. 69-71 (the system **K**), and Tienson [12], pp. 7-8. Still a fourth system lacking Restricted Range has been proposed by Stine in [11], pp. 131, ff., but it is a vastly different kind of system from **KB** and all of the others, and hence we shall not regard it as a modification of **KB** in what follows.

3. In Føllesdal's system the theorems in question are actually

   (1′) $(x) \text{ - } Pa \text{ - } (x = x)$
   (2′) $(x)(\mathsf{E}y) \text{ - } Pa \text{ - } (x = y)$

(3′) $(x)(y)(x = y \rightarrow - Pa - (x = y))$
(4′) $(x) - Pa - (Fx \vee - Fx)$,

which in his system are not equivalent, respectively, with (1)-(4). As Stine suggests in [11], p. 137, however, the English readings corresponding to formulas such as these are really not much better off than (9)-(12).

4. A detailed discussion of this point appears in [7], Chapter Four.

5. Rule (**C.EK.**=*) is dropped by Hintikka in [5], p. 4, along with his rule (**C.EK.**=) for reasons he does not make clear. Because (**C.EK**=*) and (**C.EK.EK**=*) are both unnecessary and may be argued to be mildly implausible ("Knowing who *b* is entails knowing that *b* exists" and "Knowing who *b* is entails knowing that one knows who *b* is", respectively), we do not include them in **KB***. We do, however, retain (**C.EK.**=): From '$(Ex)Ka(b = x) \in w$' infer '$(Ex)(b = x) \in w$'.

6. I wish to thank Edmund L. Gettier III for numerous discussions concerning the subject matter of this paper. Also, I am indebted to Robert C. Sleigh, Jr. for criticisms of an earlier version of this paper.

7. Sleigh [9], pp. 71-73. In doing so we shall in effect specify only the features of $\mathfrak{M}$ which are relevant to the task of semantically evaluating our rules.

8. From a formal point of view we leave ourselves uncommitted as to whether the privileged world lines in $\mathfrak{M}$ are all and only the rigid world lines. Thus, we leave ourselves open to the future possibility of regarding privileged world lines as some particular subclass of all the rigid world lines.

9. A number of such proofs may be found in [7], Chapter Five.

## REFERENCES

[1] Casteñeda, H., "Review of *Knowledge and Belief*," *The Journal of Symbolic Logic*, vol. 29 (1964), pp. 132-134.

[2] Castañeda, H., "On the logic of self-knowledge," *Noûs*, vol. 1 (1967), pp. 9-21.

[3] Clark, R., "Comments," in *Perception and Personal Identity*, N. Care and R. Grimm (eds.), Case Western Reserve Press, Cleveland (1969), pp. 174-187.

[4] Føllesdal, D., "Knowledge, identity, and existence," *Theoria*, vol. 33 (1967), pp. 1-27.

[5] Hintikka, J., "'Knowing oneself' and other problems in epistemic logic," *Theoria*, vol. 32 (1966), pp. 1-13.

[6] Hintikka, J., *Knowledge and Belief*, Cornell University Press, Ithaca (1962).

[7] Mellema, G., Ph.D. Dissertation (unpublished), University of Massachusetts, 1974.

[8] Sellars, W., "Some problems about belief," in *Words and Objections*, D. Davidson and J. Hintikka (eds.), D. Reidel, Dordrecht (1969), pp. 186-205.

[9] Sleigh, R., "Restricted range in epistemic logic," *Journal of Philosophy*, vol. 69 (1972), pp. 67-77.

[10] Sleigh, R., "Epistemic logic, again," mimeographed, University of Massachusetts (1972).

[11] Stine, G., "Quantified logic for knowledge statements," *Journal of Philosophy*, vol. 71 (1974), pp. 127-140.

[12] Tienson, J., "The 'Basic restriction' in Hintikka's *Quantified Epistemic Logic*," mimeographed, Indiana University (1972).

*Calvin College*
*Grand Rapids, Michigan*