Notre Dame Journal of Formal Logic
Volume 21, Number 3, July 1980

On Self-Sustenance in Systems of Epistemic Logic

ROBERT J. TITIEV

I wish to make some brief remarks about systems that are specified in terms of conditions upon model sets.* The appendix to Hintikka [2] contains thirty-nine such conditions involving epistemic and doxastic relationships; elsewhere [5], he discusses conditions on model sets in relation to alethic and deontic notions, as well as epistemic ones. I shall assume familiarity with the above works and the methodology whereby, in a system S that is specified via conditions on model sets, self-sustenance of a formula ϕ is established by showing that no model set μ in any model system meeting the conditions for S can contain $\sim\phi$. I shall write '$\vdash_S \phi$' to indicate that ϕ is self-sustaining in S; a formula will be referred to as a sentence if it has no free variables.

For Hintikka self-sustaining sentences are, with caveats, somewhat akin to valid sentences in systems of quantification theory; and, as the following passage shows, he has proposed that self-sustenance in a system of epistemic logic be understood in terms of truth in all worlds of a particular sort:

> Our results are not directly applicable to what is true or false in the actual world of ours. They tell us something definite about the truth and falsity of statements only in a world in which everybody follows the consequences of what he knows as far as they lead him. A sentence is self-sustaining if it is true in all such worlds, defensible if it is true in at least one such world, and so on. ([2], p. 36)

After mentioning his rejection of the interpretation of self-sustenance as truth in *all* possible worlds, Hintikka says:

*The author wishes to thank the referee for several helpful suggestions.

Received January 25, 1978; Revised April 11, 1979

> Instead, it is proposed that self-sustenance be interpreted as truth in every "epistemically perfect world", that is to say, in every possible world whose inhabitants all follow up the consequences of what they know far enough to see each particular consequence of what they actively know. If this interpretation of the metalogical notion of self-sustenance is adopted, together with the parallel interpretations of other basic metalogical concepts, then there is no objection to saying that the sense of knowing we are dealing with is essentially our ordinary sense of knowing. ([3], p. 2)

Given a particular system S, let us assume that $\mathscr{W}_S$ is the class of all epistemically perfect worlds, relative to S. We may then consider the quoted passages in terms of

(1) For all sentences ϕ of S, $\vdash_S \phi$ if and only if for all $w \in \mathscr{W}_S$, ϕ is true in w.

As Hintikka has pointed out, model sets may be viewed as giving descriptions (partial or full) of particular worlds. The membership of a sentence ϕ in a model set μ may be taken as indicating that ϕ is true in the world which μ describes; and, in terms of $\mathscr{W}_S$, this may be stated as follows:

(i) For every model set μ, there exists $w \in \mathscr{W}_S$ such that, for every sentence $\phi \in \mu$, ϕ is true in w.

That every epistemically perfect world is described by some model set is stated by

(ii) For every $w \in \mathscr{W}_S$, there exists a model set μ_w such that, for every sentence ϕ of S, $\phi \in \mu_w$ if and only if ϕ is true in w.

If, for every sentence ϕ and every $w \in \mathscr{W}_S$, ϕ fails to be true in w if and only if $\sim\phi$ is true in w, it then follows from (i) and (ii) that (1) holds.[1]

Assume now that the semantics involved in (1) is such that, for all sentences α, β, and all $w \in \mathscr{W}_S$, if α is true in w and $\alpha \supset \beta$ is true in w, then β is true in w. Then, if all sentences of tautologous form[2] are true in every epistemically perfect world, a consequence of (1) is

(2) Let $\phi_1, \ldots, \phi_n$, ψ be sentences of S such that the argument

$$\begin{array}{c}\phi_1\\ \cdot\\ \cdot\\ \cdot\\ \underline{\phi_n}\\ \psi\end{array}$$

is valid in sentential logic. Then $\vdash_S \psi$ whenever $\vdash_S \phi_1, \ldots, \vdash_S \phi_n$.

Let the following conditions from [2] be called *basic conditions*:

(*C*.&), (*C*.v), (*C*.~), (*C*.~&), (*C*.~v), (*C*.~~).

If S contains the basic conditions, then every sentence of tautologous form is a theorem of S and (2) is equivalent to the following particular case:

(3) For all sentences α, β of S, if $\vdash_S \alpha$ and $\vdash_S \alpha \supset \beta$, then $\vdash_S \beta$.

Moreover, if S contains the basic conditions and (3) holds, then the existence of a sentence ϕ such that $\vdash_S \phi$ and $\vdash_S \sim\phi$ entails that, for all sentences ψ, $\vdash_S \psi$. The consistency of S is thus ensured if there is some sentence ψ such that not $\vdash_S \psi$.

If the conditions specifying a system S contain the basic ones, then $\beta \epsilon \mu$ whenever $\alpha \epsilon \mu$ and $\alpha \supset \beta \epsilon \mu$; but this alone does not suffice to show that (3) holds for S. A metaresult specific to S is required to establish (3). Further, if (3) does hold for a system S, it provides no guarantee either that (3) holds for subsystems of S or that (3) holds for extensions of S. Thus, whenever changes are proposed concerning systems S of epistemic logic, the resulting new system(s) should be shown to satisfy (3) unless one is willing to give up the relationship in (1). A failure of (1) indicates, in model-theoretic terms, that S lacks either soundness or completeness.

An example given in [7] involves formulas α, β, and γ having the forms $(Ux)(\phi_x \supset \psi_x)$, $(Ux)(\phi_x)$, and $(Ux)(\psi_x)$, respectively. For all such formulas, $(\alpha \;\&\; \beta) \supset \gamma$ is valid in quantification theory; however, by choosing ϕ_x, ψ_x so that the former contains an epistemic operator and the latter does not, one can obtain formulas such that not $\vdash_S (\alpha \;\&\; \beta) \supset \gamma$, where S is any of the systems to be considered below. This example and others like it are compatible with (1) as long as the semantics connected with $\mathscr{W}_S$ is such that there are formulas which are valid in quantification theory yet not true in all members of $\mathscr{W}_S$. But the issue of interpreting quantifiers binding variables in the scope of modal operators has nothing to do with the example to follow; it depends only upon the basic conditions along with fundamental assumptions about sentential connectives vis-à-vis truth in epistemically perfect worlds. We shall now show that (1), (2), and (3) all fail to hold in the system KB, specified below, and in several other systems of epistemic logic.

Let KB be the system which is determined by all of the conditions listed on pp. 170-173 of [2], with the exception of $(C.E)$, $(C.U)$, $(C.E_0)$, $(C.U_0)$, and $(C.{=}!)$, since these five conditions occur in connection with illustrations of undesired consequences arising from their use. Condition (109) is to be taken with an additional proviso ruling out the case where p fails to contain any epistemic operators 'K_b', 'P_b' which have 'x' in their scope; this proviso is required in order that (109) be a special case of $(C.U_{ep})$, as intended according to remarks on pp. 146-147. Finally, as mentioned on p. 13 of [3], $(C.{=}K)$ is to be supplemented so one may use it to conclude that $(Ex)K_bF \epsilon \mu$ given that $(Ex)K_aF \epsilon \mu$ and $a = b \epsilon \mu$. The above feature of $(C.{=}K)$ enables one to establish results of the forms

$$\vdash_{\overline{KB}} (a = b \;\&\; Q_1x_1 \ldots Q_nx_nK_a\phi) \supset Q_1x_1 \ldots Q_nx_nK_b\phi$$

and

$$\vdash_{\overline{KB}} (a = b \;\&\; Q_1x_1 \ldots Q_nx_nP_a\phi) \supset Q_1x_1 \ldots Q_nx_nP_b\phi,$$

where each Q_i is either an existential or a universal quantifier. Let KB^- be the system obtained from KB by deleting the two conditions $(C.EK{=})$ and $(C.EK{=}*)$; rejecting these conditions is discussed by Hintikka in [3], [4], and [5]. Now consider

α: $a = b$
β: $(Ex)K_a(x = c)$
γ: $(Ex)K_aK_a(x = c)$
δ: $(Ex)K_bK_a(x = c)$.

These formulas are such that

$\vdash_{KB^-} \beta \supset \gamma$,
$\vdash_{KB^-} (\alpha \,\&\, \gamma) \supset \delta$,

yet it is not[3] the case that

$\vdash_{KB} (\alpha \,\&\, \beta) \supset \delta$.

From the above it follows that (1), (2), and (3) fail both in *KB* and also in KB^-.

A few points about systems other than *KB*, KB^- may now be appropriate. Let *IPE* be the system described in [6] and *EUP* the one given in [5], pp. 112-147, where it is to be understood that epistemic operators are involved. Some features and relationships among *KB*, KB^-, *IPE*, and *EUP* are brought out by considering

α: $(Ex)K_a(x = b) \supset (Ex)(x = b)$
β: $\sim(Ex)(x = x \,\&\, \sim K_a(Ey)(y = x))$
γ: $(b = c \,\&\, (Ex)(x = b \,\&\, K_a(x = b)) \,\&\, (Ex)(x = c \,\&\, K_a(x = c))) \supset K_a(b = c)$.

Then

$\vdash_{KB} \alpha$, not $\vdash_{KB^-} \alpha$, not $\vdash_{IPE} \alpha$, not $\vdash_{EUP} \alpha$,
$\vdash_{IPE} \beta$, $\vdash_{KB} \beta$, not $\vdash_{KB^-} \beta$, not $\vdash_{EUP} \beta$,
$\vdash_{IPE} \gamma$, $\vdash_{EUP} \gamma$, not $\vdash_{KB} \gamma$.

In [6] Sleigh discusses a number of problems connected with *IPE* and the fact that not $\vdash_{IPE} \alpha$; his remarks are also applicable to *EUP*. Formula β above is related to

$\sim(Ex)\sim K_a(Ey)(y = x)$

which, as Castañeda first pointed out in [1], is self-sustaining in *KB*. By using condition (i)b of *IPE* (see [6], p. 391) one may establish that $\vdash_{IPE} \beta$. The sentence γ is a theorem of *IPE* and *EUP* due to condition (*C.Ind*=). With the technique mentioned in Note 3 it is a straightforward matter to establish the above results of the form not $\vdash_S \phi$. Finally, since *KB* is an extension of KB^-, γ fails to be a theorem of KB^-.

Because *IPE* is presented in a context where 'K_a' and 'P_a' are the only epistemic operators occurring in formulas, our particular example showing failure of (3) in *KB* and KB^- does not carry over to *IPE*. It does, however, carry over to *EUP*, provided that (*C.*=*K*) is taken according to the remark on p. 13 of [3]. It could be avoided in *EUP* by agreeing to a version of the substitutivity principle on p. 116 of [5] which would allow '*a*' and '*b*' to be interchanged on subscripts of epistemic operators. Such a version would be strong enough to entail (*C.*=*K*), taken as above, and (*C.*=*P*).

Although the change suggested in the substitutivity principle blocks the particular method given above for showing that (1)-(3) fail, it still leaves the status of (1)-(3) unsettled. Were it to be claimed that any of the altered systems needed no additional rules, then there would be a clear burden to show that (1), (2), and (3) hold. In the case of the altered systems, however, there remains a need for further conditions to handle various loose ends.[4] And, as was noted earlier, adding conditions to a system which satisfies (1)-(3) may result in a new system where (1)-(3) no longer hold. If any of the systems discussed in this paper is to be extended satisfactorily, then the resulting system will have to meet the metatheoretical tests we have discussed.

NOTES

1. Pick any sentence ϕ of S. Assume that $\vdash_S \phi$ and $w \in \mathscr{W}_S$ such that ϕ fails to be true in w. Then $\sim\phi$ is true in w and, by (ii), there exists μ_w such that $\sim\phi \in \mu_w$. But this contradicts the assumption that $\vdash_S \phi$. Hence, $\vdash_S \phi$ only if, for all $w \in \mathscr{W}_S$, ϕ is true in w. Now assume that, for all $w \in \mathscr{W}_S$, ϕ is true in w, but not $\vdash_S \phi$. Then, for some μ, $\sim\phi \in \mu$. Then, by (i), there exists $w \in \mathscr{W}_S$ such that $\sim\phi$ is true in w. But this contradicts the assumption that ϕ is true in all members of $\mathscr{W}_S$. Hence, $\vdash_S \phi$ if, for all $w \in \mathscr{W}_S$, ϕ is true in w. Therefore, (1) holds.

2. A sentence is of tautologous form if it is the result of replacing the statement letters in some tautology by sentences of S. That (2) follows from (1) as claimed below may be shown by assuming that $\vdash_S \phi_1, \ldots, \vdash_S \phi_n$, where ψ follows from $\phi_1, \ldots, \phi_n$ in sentential logic. Then pick any $w \in \mathscr{W}_S$. By (1), $\phi_1, \ldots, \phi_n$ are all true in w. But $\phi_1 \supset (\phi_2 \supset (\ldots (\phi_n \supset \psi) \ldots))$ is of tautologous form and so is true in w. By repeatedly using the assumption about the semantics involved in (1), it follows that ψ is true in w. Hence, for all $w \in \mathscr{W}_S$, ψ is true in w. By (1), $\vdash_S \psi$. Therefore, (2) holds.

3. Proofs of nontheoremhood of a sentence ϕ in a system S may be given by showing the existence of a model set μ which contains $\sim\phi$ and is in a model system meeting the conditions for S. In the case of the sentence $(\alpha \,\&\, \beta) \supset \delta$ such a set μ must be infinite to take into account the doxastic rules (*C.KB*) and (*C.BK*). Let μ_1 be the set of sentences of the form $K_{i_1}B_{i_2}K_{i_3}B_{i_4} \ldots K_{i_{2n+1}}(c = c)$, where $1 \leqslant n$ and, for each j, $1 \leqslant j \leqslant 2n + 1$, $i_j \in \{a, b\}$. Let μ_2 be the set of all sentences of the form $B_{i_1}K_{i_2}B_{i_3}K_{i_4} \ldots K_{i_{2n}}(c = c)$, where $1 \leqslant n$ and, for each j, $1 \leqslant j \leqslant 2n$, $i_j \in \{a, b\}$. Let μ_3 be the set consisting of the following formulas:

   $(a = b) \,\&\, ((Ex)K_a(x = c))$, $a = b$, $(Ex)K_a(x = c)$, $\sim(Ex)K_bK_a(x = c)$, $(Ux)\sim K_bK_a(x = c)$, $K_a(c = c)$, $K_b(c = c)$, $(Ex)K_b(x = c)$, $(Ex)(x = c)$, $a = a$, $b = b$, $c = c$, and $b = a$.

   Then take $\mu = \mu_1 \cup \mu_2 \cup \mu_3$ and specify that μ is both an epistemic and a doxastic alternative to μ with respect to each of a and b. The model system $\Omega = \{\mu\}$ shows that $(\alpha \,\&\, \beta) \supset \delta$ is not a theorem of *KB*.

4. A peculiar feature of the systems *KB*, *KB*$^-$, and *EUP* is that there are formulas $\phi_{x,y}$ having 'x' and 'y' as their only free variables and such that $(Ux)(Uy)(\phi_{x,y}) \supset (Uy)(Ux)(\phi_{x,y})$ is not self-sustaining. This may be seen by letting $\phi_{x,y}$ be $K_x(y = y)$. That there remain problems with bound variables as subscripts on epistemic operators and also with certain nestings of operators has been noted by Hintikka on pp. 136-137 of [5].

REFERENCES

[1] Castañeda, H.-N., Review of [2], *The Journal of Symbolic Logic,* vol. 29 (1964), pp. 132-134.

[2] Hintikka, J., *Knowledge and Belief: An Introduction to the Logic of the Two Notions,* Cornell University Press, Ithaca, New York, 1962.

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

[4] Hintikka, J., "Individuals, possible worlds, and epistemic logic," *Noûs*, vol. 1 (1967), pp. 33-62.

[5] Hintikka, J., *Models for Modalities,* D. Reidel, Dordrecht, Holland, 1969.

[6] Sleigh, R., "On a proposed system of epistemic logic," *Noûs*, vol. 2 (1968), pp. 391-398.

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

Department of Philosophy
Wayne State University
311 Library Court
Detroit, Michigan 48202