A PARADOX REGAINED

DAVID KAPLAN AND RICHARD MONTAGUE

Another attempt has recently been made (by R. Shaw) to analyze a puzzle variously known as the Hangman, the Class A Blackout, the Unexpected Egg, the Surprise Quiz, the Senior Sneak Week, the Prediction Paradox, and the Unexpected Examination. The following simple version of the paradox is sufficient to exhibit the essential features of all other versions. A judge decrees on Sunday that a prisoner shall be hanged on noon of the following Monday, Tuesday, or Wednesday, that he shall not be hanged more than once, and that he shall not know until the morning of the hanging the day on which it will occur. By familiar arguments it appears both that the decree cannot be fulfilled and that it can.

Treatments of the paradox have for the most part proceeded by explaining it away, that is, by offering formulations which can be shown not to be paradoxical. We feel, with Shaw, that the interesting problem in this domain is of a quite different character; it is to discover an exact formulation of the puzzle which is genuinely paradoxical. The Hangman might then take a place beside the Liar and the Richard paradox, and, not unthinkably, lead like them to important technical progress.

Before the appearance of Shaw's article, we had considered a form of the paradox essentially identical with his, and found it, contrary to his assertion, not to be paradoxical. At the same time we were successful in obtaining several versions which are indeed paradoxical. The present note is intended to report these observations.

It is perhaps advisable to begin with a simple treatment due to Quine. The judge's decree, $D_1$, delivered Sunday, is that one of the following three conditions will be fulfilled: (1) The prisoner $K$ is hanged on Monday noon, but not on Tuesday or Wednesday noon, and $K$ does not know on Sunday afternoon that '$K$ is hanged on Monday noon' is true; (2) $K$ is hanged on Tuesday noon, but not on Monday or Wednesday noon, and $K$ does not know on Monday afternoon that '$K$ is hanged on Tuesday noon' is true; or (3) $K$ is hanged on Wednesday noon, but not on Monday or Tuesday noon, and $K$ does not know on Tuesday afternoon that '$K$ is hanged on Wednesday noon' is true.

Received January 7, 1960.
Let $M$, $T$, and $W$ be the respective sentences 'K is hanged on Monday noon', 'K is hanged on Tuesday noon', and 'K is hanged on Wednesday noon'. Let $K_S$ be the formula 'K knows the sentence $x$ on Sunday afternoon' (regarded as synonymous with 'K knows on Sunday afternoon that the sentence $x$ is true'), and let $K_m$ and $K_t$ be analogous, but referring to Monday and Tuesday respectively, rather than Sunday. Thus, in place of the phrase 'knows that', which requires indirect discourse, we use a locution which represents knowledge as a relation between persons and sentences. Our motive is to avoid the well-known difficulties associated with indirect discourse, and to preclude the suggestion that such difficulties may be held accountable for the paradox of the Hangman.

In accordance with this usage, the variable '$x$' in $K_S$, $K_m$, and $K_t$ has names of sentences as its substituends. It is therefore desirable to introduce a system of names of expressions. Thus if $E$ is any expression, $\bar{E}$ is to be the \textit{standard name} of $E$, constructed according to one of several alternative conventions. We might, for instance, construe $\bar{E}$ as the result of enclosing $E$ in quotes. Within technical literature a more common practice is to identify $\bar{E}$ with the numeral corresponding to the Gödel-number of $E$. As a third alternative, we could regard $\bar{E}$ as the structural-descriptive name of $E$ (within some well-determined metamathematical theory). A foundation for our later arguments could be erected on the basis of any one of these conventions.

If $E$ is any expression, then $K_S(E)$ is to be the result of replacing '$x$' by $E$ in $K_S$; and analogously for $K_m(E)$ and $K_t(E)$. Thus, if we choose the first convention for forming standard names, $K_S(M)$ is the sentence 'K knows the sentence 'K is hanged on Monday noon' on Sunday afternoon'. The decree $D_1$ can now be expressed as follows:

$$[M & \sim T & \sim W & \sim K_S(\bar{M})] v$$
$$[\sim M & T & \sim W & \sim K_m(\bar{T})] v$$
$$[\sim M & \sim T & W & \sim K_t(\bar{W})].$$

A few additional conventions will be useful. We shall employ the symbol '$\vdash$' for the logical relation of derivability within elementary syntax. Thus if $S_1$ and $S_2$ are sentences, $S_1 \vdash S_2$ if and only if $S_2$ is derivable from $S_1$ in elementary syntax (or, as we shall sometimes say, $S_1$ logically implies $S_2$); similarly, we say that $\vdash S_2$ just in case $S_2$ is provable in elementary syntax. It is well known from work of Gödel that the relation of derivability within elementary syntax is itself expressible in elementary syntax. Accordingly, we let $I$ be a formula of elementary syntax, containing '$x$' and 'y' as its only free variables, which expresses in the 'natural way' that $x$ logically implies $y$. If $E_1$ and $E_2$ are any expressions, then $I(E_1, E_2)$ is to be the result of replacing '$x$' by $E_1$ and 'y' by $E_2$ in $I$. Thus the assertion that $S_1 \vdash S_2$ is expressed in elementary syntax by the sentence $K(S_1, S_2)$.

$K$ reasons in the following way that $D_1$ cannot be fulfilled. For assume that it is. First, the hanging cannot take place on Wednesday noon; for if it did, the first two disjuncts of $D_1$ would fail, and the third would hold. But then $K$ would know on Tuesday afternoon that $\sim M$ and $\sim T$ were true, and
thus since ~ M and ~ T together imply W, he would also know on Tuesday afternoon the truth of W, which contradicts ~ Kt(W).

In this part of the argument K depends on two rather plausible assumptions concerning his knowledge:

\[(A_1) \quad [\neg M \& \neg T] \supset Kt(\neg M \& \neg T)\]
\[(A_2) \quad [I(\neg M \& \neg T, W) \& Kt(\neg M \& \neg T)] \supset Kt(W).\]

\(A_1\) is a special case of the principle of knowledge by memory, and \(A_2\) of the principle of the deductive closure of knowledge, that is, the principle that whatever is implied by one's knowledge is part of one's knowledge. Both principles may appear dubious in full generality, but we can hardly deny \(K\) the cases embodied in \(A_1\) and \(A_2\), especially after he has gone through the reasoning above.

By the foregoing argument, \(A_1\) and \(A_2\) together logically imply ~ W. It is reasonable to assume that \(K\) knows \(A_1\) and \(A_2\) (again, after using them in the previous argument):

\[(A_3) \quad K_m(A_1 \& A_2).\]

Thus, by the following instance of the principle of the deductive closure of knowledge:

\[(A_4) \quad [I(A_1 \& A_2, \neg W) \& K_m(A_1 \& A_2)] \supset K_m(\neg W),\]

\(K\) is able to establish not only that he cannot be hanged on Wednesday noon, but that he knows he cannot (that is, \(K_m(\neg W)\)).

\(K\) proceeds to exclude Tuesday noon as follows. If he is to be hanged Tuesday noon, then, still assuming \(D_1\), he infers that the second disjunct of \(D_1\) must hold. It follows (by \(A_5\) below) that \(K\) would know on Monday afternoon that ~M is true. But ~M, together with ~W, implies T. Thus T is a logical consequence of \(K\)'s knowledge, and hence \(K\) knows on Monday afternoon that T is true. However, this contradicts ~Km(T).

In this part of the argument \(K\) depends on the following analogues to \(A_1\) and \(A_2\):

\[(A_5) \quad \neg M \supset K_m(\neg M),\]
\[(A_6) \quad [I(\neg M \& \neg W, T) \& K_m(\neg M) \& K_m(\neg W)] \supset K_m(T).\]

By a similar argument, employing analogous assumptions, \(K\) also excludes Monday noon as a possible time of execution and concludes that \(D_1\) cannot be fulfilled.

The hangman reasons, on the other hand, that the decree can be fulfilled, and in fact on any of the days in question. Suppose, for example, that \(K\) is hanged on Tuesday noon but not on Monday or Wednesday; this is clearly a possible state of affairs. Then ~M, T, and ~W are true. Further, the sentence T is not analytic, even in the broad sense of following logically from the general epistemological principles whose instances are \(A_1 - A_6\). Appealing to intuitive epistemological principles (whose precise formulation is beyond the scope of the present paper), the hangman observes that one
cannot know a non-analytic sentence about the future. In particular, \( K \) cannot know on Monday afternoon that he will be hanged on Tuesday noon; thus we have \( \neg K_m(T) \). But the second disjunct of \( D_1 \) follows; thus \( D_1 \) is fulfilled.

As Quine points out, there is a fallacy of which \( K \) is guilty. The fallacy, repeated several times, crops up quite early in the argument, in fact, when \( K \) applies \( A_2 \). This application requires that \( \neg M \land \neg T \) logically imply \( W \), when obviously it does not. Indeed, \( \neg M \land \neg T \) together with \( D_1 \) logically implies \( W \); but to use this fact, we must replace \( A_2 \) by the following plausible analogue:

\[
(A_2') \ [ I \left( \neg M \land \neg T \land D_1, W \right) \land K_t \left( \neg M \land \neg T \right) \land K_t \left( D_1 \right)] \supset K_t \left( W \right),
\]

and add the assumption

\[
K_t \left( D_1 \right).
\]

But it is unreasonable to suppose that \( K \) knows that the decree will be fulfilled, especially in view of his attempt to prove the contrary.

As Shaw has remarked, the paradoxical flavor of the Hangman derives from a self-referential element in the decree which was not incorporated in Quine's formulation. The decree proposed by Shaw is essentially this: Either (1) \( K \) is hanged on Monday noon, but not on Tuesday or Wednesday noon, and on Sunday afternoon \( K \) does not know on the basis of the present decree that '\( K \) is hanged on Monday noon' is true, (2) \( K \) is hanged on Tuesday noon, but not on Monday or Wednesday noon, and on Monday afternoon \( K \) does not know on the basis of the present decree that '\( K \) is hanged on Tuesday noon' is true, or (3) \( K \) is hanged on Wednesday noon, but not on Monday or Tuesday noon, and on Tuesday afternoon \( K \) does not know on the basis of the present decree that '\( K \) is hanged on Wednesday noon' is true.\(^{11}\)

Two matters require clarification before a symbolic version of this decree can be given. First, we may ask what is meant by knowledge of one sentence on the basis of another. If \( A \) and \( B \) are sentences, then we understand the assertion that \( K \) knows \( B \) on the basis of \( A \) as meaning that \( K \) knows the conditional sentence whose antecedent is \( A \) and whose consequent is \( B \). Other interpretations are possible, but those known to us would not materially alter our discussion. Secondly, we may question the propriety of self-reference. How shall we treat in our symbolic version the phrase 'the present decree'? It has been shown by Gödel\(^{12}\) that to provide for self-reference we need have at our disposal only the apparatus of elementary syntax. Then, whenever we are given a formula \( F \) whose sole free variable is '\( x \)', we can find a sentence \( S \) which is provably equivalent to \( F(\overline{S}) \), that is, the result of replacing in \( F \) the variable '\( x \)' by the standard name of \( S \). The sentence \( F(\overline{S}) \) makes a certain assertion about the sentence \( S \). Since \( S \) is provably equivalent to \( F(\overline{S}) \), \( S \) makes the same assertion about \( S \), and thus is self-referential. Besides this method and its variants, no other precise ways of treating self-referential sentences are known to us.

In particular, we can find a sentence \( D_2 \) which is provably equivalent to the sentence
A PARADOX REGAINED

We may then not unreasonably identify $D_2$ with Shaw’s decree. All relevant features of $D_2$ are preserved if only two dates of execution are considered. Our analysis of Shaw’s argument will therefore be focused on a decree $D_3$ such that

\[(1) \rightarrow D_3 \equiv \left[ (\overline{M} \land \neg T \land \neg W \land \neg K_s(D_3 \supset \overline{M})) \lor \left( \overline{M} \land T \land \neg W \land \neg K_m(D_3 \supset T) \right) \lor \left( \overline{M} \land \neg T \land W \land \neg K_t(D_3 \supset W) \right) \right] \]

$K$ is now able to show that $D_3$ cannot be fulfilled. His argument is closely analogous to the earlier fallacious argument. He excludes first Tuesday and then Monday as possible dates of execution, and he employs as assumptions on knowledge the following analogues to $A_1 - A_4$:

\begin{align*}
(B_1) & \quad \neg M \supset K_m(\neg \overline{M}) \\
(B_2) & \quad [I(\neg M, D_3 \supset T) \land K_m(\neg \overline{M})] \supset K_m(D_3 \supset T) \\
(B_3) & \quad K_s(B_1 \land B_2) \\
(B_4) & \quad [I(B_1 \land B_2, D_3 \supset \overline{M}) \land K_s(B_1 \land B_2)] \supset K_s(D_3 \supset \overline{M})
\end{align*}

The argument can be explicitly rendered as follows. First observe that, by (1) and the sentential calculus,

\begin{align*}
(2) & \quad \neg M \vdash [D_3 \supset T] \,, \\
(3) & \quad \vdash [D_3 \land T] \supset K_m(D_3 \supset \overline{T}) \\
(4) & \quad \vdash [D_3 \land T] \supset \neg M
\end{align*}

By (4),

\[(5) \quad B_1 \vdash [D_3 \land T] \supset K_m(\neg \overline{M})\]

It is known that whenever a relation of derivability holds in elementary syntax, we can prove in elementary syntax that it holds.\textsuperscript{13} Thus, by (2),

\[(6) \quad \vdash I(\neg M, D_3 \supset T)\]

and hence

\[(7) \quad B_2 \vdash K_m(\neg \overline{M}) \supset K_m(D_3 \supset \overline{T})\]

By (3), (5), and (7),

\[B_1 \land B_2 \vdash [D_3 \land T] \supset [K_m(D_3 \supset \overline{T}) \land \neg K_m(D_3 \supset \overline{T})]\]

Thus

\[(8) \quad B_1 \land B_2 \vdash D_3 \supset \neg T \,.
\]

By (1) and the sentential calculus,

\begin{align*}
(9) & \quad \vdash [D_3 \land \neg T] \supset \overline{M} \\
(10) & \quad \vdash [D_3 \land K_s(D_3 \supset \overline{M})] \supset \neg \overline{M} 
\end{align*}
By (8) and (9),
\[ B_1 \land B_2 \vdash D_3 \supset M \]
Therefore, by the principle used to obtain (6),
\[ \vdash I(B_1 \land B_2, D_3 \supset M) \]
Hence
\[ B_4 \vdash K_S(B_1 \land B_2) \supset K_S(D_3 \supset M) \]
Therefore
\[ B_3 \land B_4 \vdash K_S(D_3 \supset M) \]
Thus, by (10),
\[ B_3 \land B_4 \vdash D_3 \supset \neg M \]
Hence, by (8),
\[ B_1 \land B_2 \land B_3 \land B_4 \vdash D_3 \supset [\neg M \land \neg T] \]
But by (1) and sentential logic,
\[ \vdash D_3 \supset [M \lor T] \]
and thus, by (11),
\[ B_1 \land B_2 \land B_3 \land B_4 \vdash \neg D_3 \]
We have shown, then, that under the (quite reasonable) assumptions \( B_1 - B_4 \) the decree cannot be fulfilled.

Mr. Shaw considers his decree genuinely paradoxical, not merely incapable of fulfillment. There appears to us, however, no good reason for supposing it so. Let us attempt to show that \( D_3 \) can be fulfilled, using the hangman's earlier argument. Suppose as before that \( K \) is hanged on Tuesday noon and only then. In this possible state of affairs, \( \neg M \) and \( T \) are true. The hangman must now establish \( \neg K_m(D_3 \supset T) \). To apply his earlier line of reasoning, he must show that \( D_3 \supset T \), considered on Monday afternoon, is a non-analytic sentence about the future. But \( D_3 \supset T \) is in fact analytic; for as \( K \) has shown, \( \neg D_3 \) follows logically from general epistemological principles, and hence so does \( D_3 \supset T \).

Now Mr. Shaw's judge, if it were suggested to him that \( K \) might be able to show his original decree \( (D_3) \) incapable of fulfillment, might attempt to avoid official embarrassment by reformulating his decree with an added stipulation, as follows. Unless \( K \) knows on Sunday afternoon that the present decree is false, one of the following conditions will be fulfilled: (1) \( K \) is hanged on Monday noon but not on Tuesday noon, and on Sunday afternoon \( K \) does not know on the basis of the present decree that '\( K \) is hanged on Monday noon' is true, or (2) \( K \) is hanged on Tuesday noon but not on Monday noon, and on Monday afternoon \( K \) does not know on the basis of the present decree that '\( K \) is hanged on Tuesday noon' is true.

But in avoiding official embarrassment the judge has plunged himself
into contradiction. Now we have a genuinely paradoxical decree! To demonstrate this, it is best to give a symbolic version, and in so doing to treat self-reference just as before; that is, we find a sentence $D_4$ (regarded as expressing the decree) such that

$$
(1) \quad \neg D_4 \equiv [K_S(\neg D_4) \lor [M \land T \land K_S(D_4 \supset M)] \lor [\neg M \land T \land \neg K_m(D_4 \supset T)]].
$$

We shall employ the following plausible assumptions, of which $C_1$ is an instance of the principle that whatever is known is true, and $C_2 - C_8$ are analogues to $B_1 - B_6$:

$$(C_1) \quad K_S(\neg D_4) \supset \neg D_4$$

$$(C_2) \quad \neg M \supset K_m(\neg M)$$

$$(C_3) \quad K_m(C_1)$$

$$(C_4) \quad [I (C_1 \land \neg M, D_4 \supset T) \land K_m(C_1) \land K_m(\neg M)] \supset K_m(D_4 \supset T)$$

$$(C_5) \quad K_S(C_1 \land C_2 \land C_3 \land C_4)$$

$$(C_6) \quad [I (C_1 \land \ldots \land C_4, D_4 \supset M) \land K_S(C_1 \land \ldots \land C_4)] \supset K_S(D_4 \supset M)$$

$$(C_7) \quad K_S(C_1 \land \ldots \land C_6)$$

$$(C_8) \quad [I (C_1 \land \ldots \land C_6, \neg D_4) \land K_S(C_1 \land \ldots \land C_6)] \supset K_S(\neg D_4).$$

First observe that, by (1),

$$
(2) \quad C_1 \vdash D_4 \supset K_S(\neg D_4).
$$

By (1) and (2),

$$
(3) \quad C_1 \land M \vdash D_4 \supset T,$$

$$
(4) \quad C_1 \vdash [D_4 \land T] \supset \neg K_m(D_4 \supset T),$$

$$
(5) \quad C_1 \vdash [D_4 \land T] \supset \neg M.$$

By (5),

$$
(6) \quad C_1 \land C_2 \vdash [D_4 \land T] \supset K_m(\neg M).$$

By (3) and the fact that whenever a relation of derivability holds, we can prove that it holds, we obtain:

$$
(7) \quad \vdash I (C_1 \land \neg M, D_4 \supset T).$$

Hence

$$
C_4 \vdash [K_m(C_1) \land K_m(\neg M)] \supset K_m(D_4 \supset T).$$

Therefore, by (6),

$$
C_1 \land \ldots \land C_4 \vdash [D_4 \land T] \supset K_m(D_4 \supset T).$$

Thus, by (4),

$$
C_1 \land \ldots \land C_4 \vdash [D_4 \land T] \supset [K_m(D_4 \supset T) \land \neg K_m(D_4 \supset T)].$$
and therefore

(8) \[ C_1 \land \ldots \land C_4 \vdash D_4 \supset \neg T \, . \]

By (1) and (2),

(9) \[ C_1 \vdash [D_4 \land \neg T] \supset M \, , \]

(10) \[ C_1 \vdash [D_4 \land K]\{D_4 \lor M]\} \supset \neg M \, . \]

By (8) and (9),

\[ C_1 \land \ldots \land C_4 \vdash D_4 \supset M \, , \]

and hence, by the principle invoked in connection with (7),

\[ \vdash I(C_1 \land \ldots \land C_4, D_4 \lor M) \, . \]

Therefore

\[ C_6 \vdash K\{C_1 \land \ldots \land C_4\} \supset K\{D_4 \lor M\} \, , \]

and thus

\[ C_5 \land C_6 \vdash K\{D_4 \lor M\} \, . \]

Hence, by (2),

\[ C_1 \land C_5 \land C_6 \vdash D_4 \lor \neg M \, . \]

Therefore, by (2) and (8),

(11) \[ C_1 \land \ldots \land C_6 \vdash D_4 \supset \neg K\{\neg D_4 \lor \neg M \land \neg T\} \, . \]

But by (1),

\[ \vdash D_4 \supset [K\{\neg D_4\} \lor M \lor T] \, , \]

and thus, by (11),

(12) \[ C_1 \land \ldots \land C_6 \vdash \neg D_4 \, . \]

We have shown, then, that under our assumptions the decree cannot be fulfilled.

But using (12) and the principle used to obtain (7), we obtain:

\[ \vdash I(C_1 \land \ldots \land C_6, \neg D_4) \, . \]

Hence

\[ C_8 \vdash K\{C_1 \land \ldots \land C_6\} \supset K\{\neg D_4\} \, . \]

Therefore

(13) \[ C_7 \land C_8 \vdash K\{\neg D_4\} \, . \]

But by (1),

\[ \vdash K\{\neg D_4\} \supset D_4 \, , \]

and thus, by (13),

(14) \[ C_7 \land C_8 \vdash D_4 \, . \]
Under our assumptions, then, the decree necessarily will be fulfilled. Thus if the formulation $D_4$ is adopted, both $K$ and the hangman are correct!

What we have shown is that the assumptions $C_1 - C_8$ are incompatible with the principles of elementary syntax. The interest of the Hangman stems from this fact, together with the intuitive plausibility of the assumptions. Indeed, before discovering the present paradox, we should certainly have demanded of an adequate formalization of epistemology that it render the conjunction of $C_1 - C_8$, if not necessary, at least not impossible. Thus the Hangman has certain philosophic consequences; but these can be made sharper by consideration of a simpler paradox, to which we were led by the Hangman.

First, it should be observed that if we consider only one possible date of execution, rather than two, a paradox can still be obtained. In this case the decree is formulated as follows. Unless $K$ knows on Sunday afternoon that the present decree is false, the following condition will be fulfilled: $K$ will be hanged on Monday noon, but on Sunday afternoon he will not know on the basis of the present decree that he will be hanged on Monday afternoon.

What is more important, however, is that the number of possible dates of execution can be reduced to zero. The judge's decree is now taken as asserting that the following single condition will be fulfilled: $K$ knows on Sunday afternoon that the present decree is false. Thus we consider a sentence $D_5$ (regarded as expressing the decree) such that

\[(1) \vdash D_5 \equiv K_s \left( \neg D_5 \right) .\]

The paradox rests on three simple assumptions which are analogous to $C_1$, $C_3$, and $C_4$:

\begin{align*}
(E_1) & \quad K_s \left( \neg D_5 \right) \supset \neg D_5 \\
(E_2) & \quad K_s \left( \neg E_1 \right) \\
(E_3) & \quad \left[ \left( E_1 \wedge \neg D_5 \right) \wedge K_s \left( \neg E_1 \right) \right] \supset K_s \left( \neg D_5 \right) .
\end{align*}

By (1),

\[\vdash D_5 \supset K_s \left( \neg D_5 \right) .\]

Hence

\[E_1 \vdash D_5 \supset \neg D_5 ,\]

and therefore

\[(2) \quad E_1 \vdash \neg D_5 .\]

By (2) and the fact that whenever a relation of derivability holds, we can prove that it holds, we obtain:

\[\vdash E_1 \wedge \neg D_5 .\]

Thus

\[E_3 \vdash K_s \left( \neg E_1 \right) \supset K_s \left( \neg D_5 \right) ,\]
and therefore

\[ E_2 \& E_3 \vdash K_s (\neg D_5) \]

But then, by (1), we obtain:

(3) \[ E_2 \& E_3 \vdash D_5 \]

We have shown, in (2) and (3), that the assumptions \( E_1 \rightarrow E_3 \) are incompatible with the principles of elementary syntax. But \( E_1 \rightarrow E_3 \) are even more plausible than \( G_1 \rightarrow G_8 \). Not only are \( E_1 \rightarrow E_3 \) simpler than their earlier counterparts, but they have the added advantage of containing no instance of the principle of knowledge by memory.

In view of certain obvious analogies with the well-known paradox of the Liar, we call the paradox connected with \( D_5 \) the Knower.

Let us now examine the epistemological consequences of the Knower. There are a number of restrictions which might be imposed on a formalized theory of knowledge in order to avoid the contradiction above. Of these, the simplest intuitively satisfactory course is to distinguish here as in semantics between an object language and a metalanguage, the first of which would be a proper part of the second. In particular, the predicate 'knows' would occur only in the metalanguage, and would significantly apply only to sentences of the object language. According to this proposal, a sentence like 'he knows 'Snow is white'' or 'Socrates knows 'there are things which Socrates does not know'' would be construed as meaningless. A less restrictive course would involve a sequence of metalinguages, each containing a distinctive predicate of knowledge, which would meaningfully apply only to sentences of languages earlier in the sequence. A more drastic measure (which seems to us distinctly unreasonable) is to reject some part of elementary syntax, perhaps by denying the existence of self-referential sentences.

The assumptions \((E_1) \rightarrow (E_3)\) are instances of the following schemata:

\[(S_1) \quad K_S (\phi) \supset \phi ,\]

\[(S_2) \quad K_S (K_S (\phi) \supset \phi) ,\]

\[(S_3) \quad [I (\phi) \& K_S (\phi) ] \supset K_S (\psi) ,\]

where \( \phi \) and \( \psi \) are arbitrary sentences. Using the Knower, we can show that any formal system containing the apparatus of elementary syntax, and including among its theorems all instances of \((S_1) \rightarrow (S_3)\), is inconsistent. Using the Liar, Tarski has obtained a similar result: any formal system containing the apparatus of elementary syntax, and including among its theorems all sentences

\[ T (\phi) \equiv \phi ,\]

where \( \phi \) is a sentence of the formal system, is inconsistent.\(^{14}\) The precise relations between Tarski's result and ours are not at present clear, but would appear to constitute an interesting subject of research.

It should be mentioned that if any one of \( S_1 \rightarrow S_3 \) is removed, it can be
shown that the remaining schemata are compatible with the principles of elementary syntax.

NOTES

1. The content of this article was presented before the Philosophy Club of the University of California at Los Angeles on March 14, 1958.


5. Abraham Kaplan, in conversation.


8. In connection with this treatment of knowledge, see Carnap, 'On belief sentences', in *Meaning and Necessity*, enlarged edition (Chicago 1956) pp. 230-232, in which, however, only the relation of belief is considered explicitly. We should mention as well another departure from ordinary usage. Judicial decrees would ordinarily not be construed as indicative sentences. To those who are bothered by our practice of identifying decrees with indicative sentences, we suggest reading 'the indicative corresponding to the decree' for 'the decree'.


10. By *elementary syntax* we understand a first-order theory containing—in addition to the special formulas $K_S$, $K_m$, $K_t$, $M$, $T$, and $W$—all standard names (of expressions), means for expressing syntactical relations between, and operations on, expressions, and appropriate axioms involving these notions. The form of such a theory will of course depend on the convention adopted for the assignment of standard names. If the second convention is adopted, we could identify elementary syntax with Peano's arithmetic (the theory $P$ of Tarski, Mostowski, Robinson, *Undecidable Theories*) or even with the much weaker theory $Q$ (of the same work)—in either case, however, supplemented by the special formulas mentioned above.
11. The only significant respect in which this version differs from Shaw's is in saying 'K does not know on the basis of the present decree' where Shaw would say 'K cannot deduce from the present decree'. But the latter version cannot be taken in its usual sense. On Tuesday afternoon, for instance, K's deduction will involve as premises not only the decree but also the mnemonic knowledge of not having been hanged on Monday or Tuesday noon.


*University of California*
*Los Angeles, California*