Note on a Theorem of M. H. LÖB※

By

Kempachiro ŌHASHI

The rigoruos investigations of "truth" and "provability" in a latest quarter century began with the incomplete theorem of K. Gödel [1], where he presented a negative example, the formula asserting its own unprovability, more precisely, the formula A with q as its Gödel number expressing unprovability in S of a formula with q. Corresponding to it, Leon Henkin [2] presented the following problem :

"If S is any standard formal system adequate for recursive number theory, a formula (having a certain integer p as its Gödel number) can be constructed which expresses the proposition that the formula with Gödel number p is provable in S. Is this formula provable or not in S ? "

This problem was positively solved for a suitable "provability" predicate by Löb. [3] That is, he presented a sufficient condition of "provability" predicate for the problem.

G. Kreisel, [4] however, reviewed that the "provability" predicate used by Löb is rather strong and the one constructed by the former does not satisfy the condition by the latter. It is then natural to inquire for what kind of "provability" predicate the above problem is negatively solved.

The aim of this paper is to show that a modified "provability" predicate is essentially necessary and sufficient for the positive solution.

Acquaintance with 〔I〕 [5] is assumed.

I. Theorem of M. H. Löb

Let (S) be a standard consistent formal system containing recursive number theory and Gödel substitution function for a Gödel numbering of (S), and $\{A\}$ denote the Gödel number of a formula A, $\boldsymbol{B}(n)$ be a "provability" predicate, corresponding to the conception "provable" defined for (S), that is, a predicate that expresses the proposition that the formula with n is provable in (S), and, moreover, $S(n, m)$ be a function whose value is the Gödel number of the formula with n as its Gödel number in which is substituted a variable, corresponding to m,

※ Contribution from the Shimonoseki College of Fisheries, No. 219

in the only argument-place. By means of S (n, m) Theorem of Löb is proved. Theorem of Löb is :

If $\boldsymbol{B}$ satisfies the following conditions ;

(1) $\boldsymbol{B}(\{A \supset B\}) \supset (\boldsymbol{B}(\{A\}) \supset \boldsymbol{B}(\{B\}))$

(2) $\vdash A \rightarrow \vdash \boldsymbol{B}(\{A\})$

(3) $\vdash \boldsymbol{B}(\{A\}) \supset \boldsymbol{B}(\{\boldsymbol{B}(\{A\})\})$

then $\vdash$ A for any formula A that $\vdash \boldsymbol{B}(\{A\}) \sim A$.

By this theorem we obtain a formal proof of an incomplete theorem.

In a consistent system wth $\boldsymbol{B}$ satisfying (1), (2), (3)and (4) there is an undecidable formula.

(4) $\vdash \boldsymbol{B}(\{A\}) \rightarrow \vdash A$

Proof.

Let

1.1 $\vdash \neg A$

Suppose that

1.2 $\vdash \neg \boldsymbol{B}(\{A\})$

1.3 $\vdash \boldsymbol{B}(\{A\}) \sim A$

By Theorem of Löb

1.4 $\vdash A$

This is absurd, and therefore,

1.5 $\overline{\vdash \neg \boldsymbol{B}(\{A\})}$

Suppose that

1.6 $\vdash \boldsymbol{B}(\{A\})$

1.7 $\vdash A$

This contradicts 1.1, and therefore,

1.8 $\overline{\vdash \boldsymbol{B}(\{A\})}$

Thus we obtain an undecidable formula $\boldsymbol{B}(\{A\})$.

II. Henkin's Problem

(1), (2) and (3) are a sufficient condition for positive solution of Henkin's problem. What is necessary for it? The following theorems will answer it.

In (S) with a "provability" predicate satisfying (2) and (4), for any formula A that $\vdash \boldsymbol{B}(\{A\}) \sim A$, $\vdash A$ if and only if

(5) $\vdash \boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset \boldsymbol{B}(\{A\}))$

where φ is the formula $\boldsymbol{B}(\{\varphi\}) \supset A$.

Proof.

2.1 $\vdash \boldsymbol{B}(\{A\}) \sim A$

Suppose that

2.2 $\vdash A$

2.3 $\vdash \boldsymbol{B}(\{\varphi\}) \& \boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset A$

Conversely, 2.3 → 2.2, as is seen in the proof of Theorem of Löb, and therefore 2.2 is equivalent to 2.3.

$\vdash \boldsymbol{B}(\{A\}) \sim A$ and $\vdash (\boldsymbol{B}(\{\varphi\}) \supset A) \sim \varphi$.

2.3 is then equivalent to

2.4 $\vdash (\boldsymbol{B}(\{A\}) \sim A) \supset ((\boldsymbol{B}(\{\varphi\}) \supset A) \sim \varphi) \supset (\boldsymbol{B}(\{\varphi\})\ \&\ \boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset A))$

Take A', φ' and K for $\boldsymbol{B}(\{A\})$, $\boldsymbol{B}(\{\varphi\})$ and $\boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\})$ respectively.

2.5 $\vdash (A' \sim A) \supset ((\varphi' \supset A) \sim \varphi) \supset (\varphi\ \&\ K \supset A))$

2.6 $\vdash \neg(\neg A' \vee A)\ \&\ (A' \vee \neg A)) \vee \neg(\neg(\neg\varphi' \vee A) \vee \varphi)\ \&\ (\neg\varphi \vee (\neg\varphi' \vee A))$
$\vee (\neg\varphi' \vee \neg K \vee A)$

2.7 $\vdash (A'\ \&\ \neg A) \vee (\neg A'\ \&\ A) \vee ((\neg\varphi' \vee A)\ \&\ \neg\varphi) \vee (\varphi\ \&\ \varphi'\ \&\ \neg A)$
$\vee \neg\varphi' \vee \neg K \vee A$

2.8 $\vdash (A' \vee A) \vee ((\neg\varphi' \vee A) \& \neg\varphi) \vee (\varphi \& \varphi' \& \neg A) \vee \neg\varphi' \vee \neg K$

2.9 $\vdash A' \vee A \vee (\neg\varphi' \& \neg\varphi) \vee (A \& \neg\varphi) \vee ((\varphi \vee \neg\varphi') \& (\neg A \vee \neg\varphi')) \vee \neg K$

2.10 $\vdash A' \vee (\neg\varphi' \& \neg\varphi) \vee (A \& \neg\varphi) \vee A \vee \varphi \vee \neg\varphi' \vee K$

2.11 $\vdash \neg\varphi' \vee \varphi \vee A' \vee A \vee \neg K$

2.12 $\vdash (\neg\varphi' \vee \varphi) \vee (\neg K \vee (\neg\varphi' \vee A)) \vee (\neg K \vee (\neg\varphi' \vee A'))$

2.13 $\vdash (\boldsymbol{B}(\{\varphi\}) \supset \varphi) \vee (\boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset A))$
$\vee (\boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset \boldsymbol{B}(\{A\}))$

The first term and the second in the disjunction 2.13 are of the same formula, and moreover, are equivalent to the third. For, suppose

2.14 $\vdash \boldsymbol{B}(\{\varphi\}) \supset \varphi$

2.15 $\vdash \boldsymbol{B}(\{\varphi\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset A)$

2.16 $\vdash \boldsymbol{B}(\{\varphi\}) \supset A$

2.17 $\vdash \varphi$

2.18 $\vdash \boldsymbol{B}(\{\varphi\})$

2.19 $\vdash A$

2.20 $\vdash \boldsymbol{B}(\{A\})$

2.21 $\vdash \boldsymbol{B}(\{\boldsymbol{B}(\{\varphi\}) \supset A\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset \boldsymbol{B}(\{A\}))$

Conversely, suppose 2.21

2.22 $\vdash \boldsymbol{B}(\{\varphi\}) \supset (\boldsymbol{B}(\{\varphi\}) \supset \boldsymbol{B}(\{A\}))$

2.23 $\vdash \boldsymbol{B}(\{\varphi\}) \supset A$

2.24 $\vdash \varphi$

2.25 $\vdash \boldsymbol{B}(\{\varphi\})$

2.14 $\vdash \boldsymbol{B}(\{\varphi\}) \supset \varphi$

Thus 2.21 is equivalent to 2.14 and then to 2.13. Since replacements from 2.4 to 2.13 are equivalent, the theorem is obtained.

In (S) with $\boldsymbol{B}$ satisfying (6), $\vdash A$ for any formula A such that $\vdash \boldsymbol{B}(\{A\}) \sim A$ if H, which is the formula in the problem of H. Henkin, is formally obtained by (2), (6) and rules of inference.

(6) $\vdash \boldsymbol{B}(\{X \supset Y\}) \rightarrow \vdash \boldsymbol{B}(\{X\}) \supset \boldsymbol{B}(\{Y\})$

Proof.

A proof generally can be represented by such a figure as presented in Fig. 1, that is, a proof is regarded as a tree, and each top of it is an axiom of (S), and a sequence (of formula) under a line is derived from a seqence on the line by the rule of inference. For example, $\triangle_3(\triangle_5)$ in Fig. 1 is derived from $\triangle_2$ and $\triangle_2 \supset \triangle_3(\triangle_4)$.

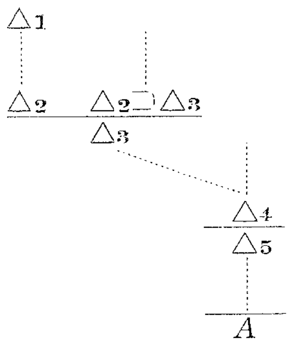

Fig. 1.

where rules of inference are as follows;

(a) $$\frac{A \quad A\supset B}{B}$$

(b) $$\frac{C\supset A\ (t)}{C\supset \forall x A(x)}$$

(c) $$\frac{A(t)\supset C}{\exists x A\ (x)\ \supset C},$$

where t is not included free in C.

Now, from the proof of the formula H, we will construct PA, which is a proof of A, as follows. Replace H in PH by A, and

$$\dfrac{A(\cdots\{\cdots H\cdots\},\cdots) \qquad \dfrac{\{H\}=\{\boldsymbol{B}(\{H\}) \quad \{H\}=\{\boldsymbol{B}(\{H\})\}\supset(A(\cdots\{\cdots H\cdots\}\cdots)\supset A(\cdots\{\cdots \boldsymbol{B}(\{H\})\cdots\}\cdots)\cdots)}{A(\cdots\{\cdots H\cdots\}\cdots)\supset A(\cdots\{\cdots \boldsymbol{B}(\{H\})\cdots\},\cdots)}}{A(\cdots\{\cdots \boldsymbol{B}(\{H\})\cdots\},\cdots)}$$

is replaced by a new figure defined below.

H and $\boldsymbol{B}(\{H\})$ are wffs without any free variables, and then the fact that not only $\vdash H \sim \boldsymbol{B}(\{H\})$ but H is $\boldsymbol{B}\ (\{H\})$, is used only in the forms $\{\cdots H\cdots\}$ and $\{\cdots \boldsymbol{B}(\{H\})\cdots\}$, $(\{H\}=\{\boldsymbol{B}(\{H\})\}\ )$.

i. e. in a schema as follows;

(h) $$\frac{\{H\}=\{\boldsymbol{B}(\{H\})\} \quad A(\cdots\{\cdots H\}\cdots)}{A(\cdots\{\cdots \boldsymbol{B}(\{H\})\cdots\}\cdots)}$$

It is sufficient to show how to construct PA in the following figure.

$$\frac{\{H\}=\{\boldsymbol{B}(\{H\})\} \quad \times(\cdots \boldsymbol{B}(\{\cdots Y(\cdots H\cdots)\cdots\}\cdots)\cdots)}{\times(\cdots \boldsymbol{B}(\cdots\{\cdots Y(\cdots \boldsymbol{B}(\{H\})\cdots)\cdots\})\cdots)},$$

where X and Y do not contain $\boldsymbol{B}$, or it is divided into the above cases.

Take A_0 for $Y\ (\cdots H\cdots)$

B_0 〃 $Y\ (\cdots \boldsymbol{B})(\{H\})\cdots)$

A_1 〃 $\mathbf{B}(\{A_\circ\})$
B_1 〃 $\mathbf{B}(\{B_\circ\})$
A_2 〃 $\times(\cdots A\cdots)$
B_2 〃 $\times(\cdots B\cdots)$

$$\dfrac{\dfrac{A\sim\mathbf{B}(\{A\}) \qquad \overset{\vdots}{(A\sim\mathbf{B}(\{A\}))\sim(A_\circ\sim B_\circ)}}{A_\circ\sim B_\circ} \qquad \overset{\vdots}{(A_\circ\sim B_\circ)\sim(B_\circ\supset A_\circ)}}{B_\circ\supset A_\circ}$$
$$\vdots$$
$$\mathbf{B}(\{B_\circ\supset A_\circ\})$$
$$\vdots$$
$$B_1\supset A_1$$

Fig. 2.

$$\dfrac{\dfrac{A\sim\mathbf{B}(\{A\}) \qquad \overset{\vdots}{(A\sim\mathbf{B}(\{A\}))\supset(A_\circ\supset B_\circ)}}{A_\circ\sim B_\circ} \qquad \overset{\vdots}{(A_\circ\sim B_\circ)\supset(A_\circ\sim B_\circ)}}{A_\circ\supset B_\circ}$$
$$\vdots$$
$$\mathbf{B}(\{A_\circ\supset B_\circ\})$$
$$\vdots$$
$$\dfrac{A_1\supset B_1 \qquad (A_1\supset B_1)\supset((B_1\supset A_1)\supset(A_1\sim B_1))}{(B_1\supset A_1)\supset(A_1\supset B_1)}$$

Fig. 3.

$$\dfrac{A_2 \qquad \dfrac{\dfrac{\dfrac{B_1\supset A_1 \qquad (B_1\supset A_1)\supset(A_1\supset B_1)}{A_1\sim B_1} \qquad \overset{\vdots}{(A_1\sim B_1)\supset(A_2\sim B_2)}}{A_2\sim B_2} \qquad \overset{\vdots}{(A_2\sim B_2)\supset(A_2\supset B_2)}}{A_2\supset B_2}}{B_2}$$

Fig. 4.

If parts (h) in PH are replaced by a figure presented in Figs 2, 3, 4, we have a new figure PA. It is easily seen that PA constructed here is a proof of A such that $\vdash \mathbf{B}(\{A\})\sim A$. Thus theorem is proved.

Applying the above replacement to 2.4, we have,

A formula A, such that $\vdash \mathbf{B}(\{A\})\sim A$, is formally derived in (S) satisfying Deduction theorem and (2), $\vdash A$ if and only if for any formula X and Y

(7) $\vdash \mathbf{B}(\{\mathbf{B}(\{X\})\supset Y\})\supset(\mathbf{B}(\{X\})\supset \mathbf{B}(\{Y\}))$

Remarks.

Asssume

(7) $\vdash \mathbf{B}(\{\mathbf{B}(\{X\})\supset Y\})\supset(\mathbf{B}(\{X\})\supset \mathbf{B}(\{Y\}))$

Take $B(\{X\})$ for Y,

3.1 $\vdash \mathbf{B}(\{\mathbf{B}(\{X\})\supset \mathbf{B}(\{X\})\})\supset(\mathbf{B}(\{X\})\supset \mathbf{B}(\{\mathbf{B}(\{X\})\}))$

3.2 $\vdash \mathbf{B}(\{X\})\supset \mathbf{B}(\{\mathbf{B}(\{X\})\})$

Therefore, (7) $\rightarrow$ (2).

On the other hand, assume (3), (1),

3.3 $\vdash \boldsymbol{B}(\{\boldsymbol{B}(\{X\})\supset Y\})\supset(\boldsymbol{B}(\{\boldsymbol{B}(\{X\})\})\supset\boldsymbol{B}(\{Y\}))$

3.4 $\vdash \boldsymbol{B}(\{\boldsymbol{B}(\{X\})\supset Y\})\supset(\boldsymbol{B}(\{X\})\supset\boldsymbol{B}(\{Y\}))$.

Therefore (1) and (2) → (7).

By the theorems proved above, we have the following chief proposition.

In (S) with $\boldsymbol{B}$ satisfying (2), (4), and (6), $\vdash A$ is formally derivable for any formnla A, provided that $\vdash \boldsymbol{B}(\{A\})\sim A$, if and only if for any formula X and Y

$$\vdash \boldsymbol{B}(\{\boldsymbol{B}(\{X\})\supset Y\})\supset(\boldsymbol{B}(\{X\})\supset\boldsymbol{B}(\{Y\}))$$

The theorems obtained above are only the results on general investigations without entering into precise considerations of "provability", and then they do not answer the question. For it, particular invesitgations of "provability" of each axiom system have to be made.

References

1) GÖDEL, K. : 1931. Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme I. *Monatshefte fur Math. Physik.*, **38** pp. 173—198.

2) Problem 3. *J. Sŷmb. Logic*, **17** (1952) p. 160.

3) LÖB, M. H. : 1955. Solution of a problem of Leon Henkin. *J. Symb. Logic*, **20** pp. 115—118.

4) KREISEL, G. : 1956. *Math. Rev.*, vol. **17**. no. pp. 5—6.

5) KLEENE, S. C. : 1952. Introduction into metamathematics. cited [1].