"FLEXIBLE" PREDICATES OF FORMAL NUMBER THEORY

SAUL A. KRIPKE

The famous incompleteness theorem of Gödel [1] showed that a formal system containing the usual number theory must have an "undecidable" statement whose truth-value is not determined by the formal system. The present note gives an analogous theorem constructing predicates which are "flexible" in the sense that their extensions as sets are left undetermined by the formal system. We utilize the recursive-function-theoretic approach of Kleene [2]; knowledge of his work is presupposed, and his notations and terminology will be used freely.

Consider a formal system **F** satisfying the following conditions: (1) **F** is formalized on the basis of the classical[1] predicate calculus of first order with identity; (2) for every natural number n, **F** contains, either as a primitive or by (possibly even contextual) definition, a numeral $\mathbf{n}$, which is a value of the variables (interpreted as ranging over natural numbers); (3) every partial recursive function is numeralwise representable in **F**; (4) under a Gödel numbering satisfying the usual effectiveness conditions,[2] the set of theorems of **F** is recursively enumerable; (5) if $m \neq n$, $\neg \mathbf{m} = \mathbf{n}$ is provable in **F**. These conditions are satisfied by the **F** of [2], but we do not restrict ourselves to this system. We let $\mathrm{R}_n(\mathrm{x}, \mathrm{y}_1, \cdots, \mathrm{y}_n, \mathrm{z})$ numeralwise represent Kleene's enumeration function $\Phi_n(x, y_1, \cdots, y_n)$. Noting that (3) implies that every recursive predicate is numeralwise expressible in **F**, we let $\mathrm{T}_n(\mathrm{z}, \mathrm{x}_1, \cdots, \mathrm{x}_n, \mathrm{y})$ numeralwise express $T_n(z, x_1, \cdots, x_n, y)$.

We list several definitions: (A) $\Sigma_{m,n}$ is the class of m-ary (intuitive) predicates expressible in the n-quantifier form of Kleene [2, Theorem V], with the existential quantifier first. (B) If $\mathrm{P}(\mathrm{x}_1, \cdots, \mathrm{x}_m, \mathrm{y}_1, \cdots, \mathrm{y}_n)$ numeralwise expresses $P(x_1, \cdots, x_m, y_1, \cdots, y_n)$, we shall say $\mathrm{Q}_1\mathrm{y}_1 \cdots \mathrm{Q}_n\mathrm{y}_n\mathrm{P}(\mathrm{x}_1, \cdots, \mathrm{x}_m, \mathrm{y}_1, \cdots, \mathrm{y}_n)$ is a *formalization* of $(Q_1y_1) \cdots (Q_ny_n)P(x_1, \cdots, x_m, y_1, \cdots, y_n)$ in **F** (with the formal quantifiers corresponding to their intuitive counterparts). (C) We

Received by the editors March 11, 1960 and, in revised form, November 30, 1960 and April 21, 1961.

[1] Lemma 1, Theorem 1, and Corollary 1.1 remain valid for intuitionistic systems.

[2] We shall not endeavor to make these precise, but they should imply: (1) The wffs of **F** form a recursive set; (2) the Gödel number of a partial recursive function effectively determines its numeralwise representation; (3) substitution is a recursive function; (4) the Gödel number of a compound formula effectively determines its components, quantifiers, connectives, free variables, numerals, etc., and their positions.

define a set α of formulas of $\mathbf{F}$ to be an *independent set* iff for any subset β of α, the system obtained by adding to $\mathbf{F}$ as new axioms the formulas of β, plus the negations of the formulas of $\alpha-\beta$, is consistent. To show that a set α is independent, actually it clearly suffices to prove the stated property for every *finite* subset β of α. (D) We define a predicate $\mathrm{P}(\mathrm{x}_1, \cdots, \mathrm{x}_m)$ as *flexible for* $\Sigma_{m,n}$ iff each predicate of $\Sigma_{m,n}$ has a formalization $\mathrm{Q}(\mathrm{x}_1, \cdots, \mathrm{x}_m)$ in $\mathbf{F}$ such that $\forall \mathrm{x}_1 \cdots \forall \mathrm{x}_m(\mathrm{P}(\mathrm{x}_1, \cdots, \mathrm{x}_m) \sim \mathrm{Q}(\mathrm{x}_1, \cdots, \mathrm{x}_m))$ is consistent with $\mathbf{F}$. (E) Finally, the following convention will sometimes be useful. If $\phi(x_1, \cdots, x_n)$ is a partial recursive function of n variables, then we write $\mathrm{A}(\phi(\mathbf{x}_1, \cdots, \mathbf{x}_n))$ for $\exists \mathrm{y}(\mathrm{B}(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathrm{y}) \& \mathrm{A}(\mathrm{y}))$, where $\mathrm{B}(\mathrm{x}_1, \cdots, \mathrm{x}_n, \mathrm{y})$ numeralwise represents $\phi(x_1, \cdots, x_n)$. (The quantifier $\exists \mathrm{y}$ is to have as large a scope as possible; i.e., $\mathrm{A}(\cdots)$ represents the entire displayed context in which $\phi(\mathbf{x}_1, \cdots, \mathbf{x}_n)$ occurs.) Hence if $\mathrm{B}(\mathbf{x}_1, \cdots, \mathbf{x}_n, \boldsymbol{y})$ and $\exists!\mathrm{zB}(\mathbf{x}_1, \cdots, \mathbf{x}_n, \mathrm{z})$ are provable (and thus in particular if $\phi(x_1, \cdots, x_n)=y$), then $\mathrm{A}(\phi(\mathbf{x}_1, \cdots, \mathbf{x}_n))$ is equivalent in $\mathbf{F}$ to $\mathrm{A}(\boldsymbol{y})$. In particular, we write $\mathrm{A}(\Phi_1(\mathbf{x}, \boldsymbol{y}))$ or even $\mathrm{A}(\mathbf{x}(\boldsymbol{y}))$ for $\exists \mathrm{z}(\mathrm{R}_1(\mathbf{x}, \boldsymbol{y}, \mathrm{z}) \& \mathrm{A}(\mathrm{z}))$.

LEMMA 1. *If* $\mathbf{F}$ *is consistent, then there exists a natural number e such that, for every natural number y,* $\mathrm{R}_1(\mathbf{e}, \mathbf{e}, \boldsymbol{y}) \& \exists!\mathrm{zR}_1(\mathbf{e}, \mathbf{e}, \mathrm{z})$ *is consistent with* $\mathbf{F}$.

PROOF. Let $\mathbf{F}$ be consistent. Define a partial recursive function ϕ by the stipulation that $\phi(x)=y$ if

$$\neg(\mathrm{R}_1(\mathbf{x}, \mathbf{x}, \boldsymbol{y}) \ \& \ \exists!\mathrm{zR}_1(\mathbf{x}, \mathbf{x}, \mathrm{z})) \tag{6}$$

is provable in $\mathbf{F}$. (If more than one formula of the form (6) is provable, choose the one occurring earliest in the recursive enumeration of the theorems of $\mathbf{F}$.) Let e be a Gödel number of ϕ. Then e will satisfy the requirements of the Lemma as long as no formula of the form (6) is provable for $x=e$. If some formula (6) were provable for $x=e$, then $\phi(e)=y$ would be defined. Hence, since the predicate $\mathrm{R}_1(\mathrm{x}, \mathrm{y}, \mathrm{z})$ numeralwise represents the function Φ_1, and since $\phi(e)=\Phi_1(e, e)=y$, $\mathrm{R}_1(\mathbf{e}, \mathbf{e}, \boldsymbol{y})$ and $\exists!\mathrm{zR}_1(\mathbf{e}, \mathbf{e}, \mathrm{z})$ are provable in $\mathbf{F}$. But $\phi(e)=y$ implies that $\neg(\mathrm{R}_1(\mathbf{e}, \mathbf{e}, \boldsymbol{y}) \& \exists!\mathrm{zR}_1(\mathbf{e}, \mathbf{e}, \mathrm{z}))$ is also provable, contrary to the consistency of $\mathbf{F}$. Q.E.D.

Although the crucial property of e holds only if $\mathbf{F}$ is consistent, from the constructive point of view it is noteworthy that e is constructed independently of the consistency hypothesis on $\mathbf{F}$. In fact, e depends recursively on the Gödel number of the r.e. set of (Gödel numbers of) theorems of $\mathbf{F}$. (Similar remarks apply to the existential claims made in Theorem 1 and in Corollary 1.1; these results are proved by finitary methods.) Since Lemma 1, despite its simplicity, is a powerful result, e will denote, throughout this note, the number defined in Lemma 1.

THEOREM 1. *If* **F** *is consistent, then for all non-negative integers* m *and* n, *there exists an* m-*ary predicate* $P_n(x_1, \cdots, x_m)$ *which is flexible for* $\Sigma_{m,n}$.

PROOF. The proof is an application of Lemma 1. Assume $n \geqq 1$. Define $P_n(x_1, \cdots, x_m)$ as

$$\exists y_1 \forall y_2 \exists y_3 \cdots Q y_n T_{m+n-1}(e(e), x_1, \cdots, x_m, y_1, \cdots, y_n).$$

Then if we add $R_1(e, e, y) \& \exists ! z R_1(e, e, z)$ to **F** (by Lemma 1, this cannot destroy consistency), then $P_n(x_1, \cdots, x_m)$ becomes equivalent to $\exists y_1 \forall y_2 \exists y_3 \cdots Q y_n T_{m+n-1}(y, x_1, \cdots, x_m, y_1, \cdots, y_n)$. Since every predicate in $\Sigma_{m,n}$ has (for a suitable choice of y) a formalization of this latter form, we are done. If $n = 0$,[3] define $P_0(x_1, \cdots, x_m)$ as $R_m(e(e), x_1, \cdots, x_m, 0)$. Then, if as before we add $R_1(e, e, y) \& \exists ! z R_1(e, e, z)$, $P_0(x_1, \cdots, x_m)$ reduces to $R_m(y, x_1, \cdots, x_m, 0)$; and every m-ary recursive predicate is numeralwise representable in this form, where y defines recursively the representing function of the predicate. Q.E.D.

COROLLARY 1.1. *Let* **F** *be consistent. Then there exists a predicate* $P(x)$ *in* **F** *such that the formulas* $P(n)$, *for non-negative integral* n, *form an independent set in* **F**.[4]

PROOF. Using Theorem 1, for $m = 1, n = 0$, let $P(x)$ be the "flexible" predicate given by the theorem, i.e., $R_1(e(e), x, 0)$. Let α be the set containing $P(n)$ for all non-negative integral n. To show that α is independent, let β be a finite subset of α. Then the finite set $\{n \mid P(n) \in \beta\}$ is recursive and hence is in $\Sigma_{1,0}$; let $\Phi_1(f, x)$ be a representing function for it, for fixed f. Now let $Q(x)$ abbreviate $R_1(f, x, 0)$. By the proof of Theorem 1, if **F** is consistent, it remains consistent if we adjoin $\forall x(P(x) \sim Q(x))$. But if $P(n)$ is in β, $Q(n)$ is provable in **F**, and hence we get $P(n)$ in the new system. If $P(n)$ is in $\alpha - \beta$, $\Phi_1(f, n) = 1$, so that $R_1(f, n, 1)$, and hence $\neg Q(n)$, is provable in **F**; thence $\neg P(n)$ is provable in the augmented system. Q.E.D.

REMARK 1. If we had desired a formal definition of "flexible predicate," the property of Corollary 1.1 would be a plausible candidate for the job: A predicate $P(x_1, \cdots, x_m)$ is *flexible* in **F** iff the set $\{P(x_1, \cdots, x_m) \mid x_1, \cdots, x_m \text{ non-negative integers}\}$ is independent

[3] Since $\sum_{m,0} \subseteq \sum_{m,1}$ a special treatment of this case is not really necessary; but it will be used for Corollary 1.1 below.

[4] According to information received from the referee and from Professor Hartley Rogers, this Corollary was first proved (using a different method) by Mostowski; following him, by Feferman and Scott.

in **F**. On this definition, even if **F** is consistent, the predicate "flexible for $\Sigma_{m,n}$" constructed in Theorem 1 is known to be "flexible" in the absolute sense only if $n=0$; for other values of n, the possibility of ω-inconsistency and even higher orders of inconsistency is a stumbling block. But by suitable manipulations this stumbling block can be removed. For example, if $m=n=1$: Let $P_1(x)$ be

$$(7)\ (\exists a T_1((e(e)_0, x, a)\ \&\ (e(e))_2 = 0) \vee (R_1((e(e))_1, x, 0)\ \&\ (e(e))_2 = 1).$$

Then if $R_1(e, e, z) \& \exists! a R_1(e, e, a)$ is added to **F**, we can replace $e(e)$ by z throughout (7). Thence if $z=2^y$, (7) would reduce to $\exists a T_1(y, x, a)$; while if $z=3^y \cdot 5$, then (7) would reduce to $R_1(y, x, 0)$. Thus the proofs of both Theorem 1 (for $m=n=1$) and Corollary 1.1 go through for $P_1(x)$ thus defined.

REMARK 2. The referee has pointed out that a weaker form of Corollary 1.1 holds in every essentially undecidable system; namely, each such system contains an infinite independent set of formulas. Since this fact appears not to be generally known, I outline a proof here: Let P be a formula undecidable in **F**, let Q_1 be a formula undecidable in $\mathbf{F}+\{P\}$, and let Q_2 be a formula undecidable in $\mathbf{F}+\{\neg P\}$. Let $R_1, \cdots, R_4$ be formulas undecidable in $\mathbf{F}+\{P\&Q_1\}$, $F+\{P\&\neg Q_1\}$, $\mathbf{F}+\{\neg P\&Q_2\}$, and $\mathbf{F}+\{\neg P\&\neg Q_2\}$, respectively. Continue in this manner; then

$$\{P, (P\ \&\ Q_1) \vee (\neg P\ \&\ Q_2), (P\ \&\ Q_1\ \&\ R_1) \vee (P\ \&\ \neg Q_1\ \&\ R_2)$$
$$\vee (\neg P\ \&\ Q_2\ \&\ R_3) \vee (\neg P\ \&\ \neg Q_2\ \&\ R_4), \cdots\}$$

form an independent set.

(*Added June* 4, 1962) REMARK 3. Mostowski's proof of Corollary 1.1. (cf. footnote 4) has appeared in [3]. It is clear that his generalization of Corollary 1.1 to r.e. families of systems [3, Theorem 4] can be proved by the present methods also; for the proof of Lemma 1 easily extends to any r.e. family of consistent systems **F**, and thence Theorem 1 and Corollary 1.1 also hold for any such family. Results on ω-complete, nonconstructive systems, like Mostowski's §6, can be proved by an extension of the present method; but we leave this matter for later work.

BIBLIOGRAPHY

1. K. Gödel, *Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme.* I, Monatsh. Math. Phys. **38** (1931), 173–198.

2. S. C. Kleene, *Introduction to metamathematics*, Van Nostrand, New York; North-Holland, Amsterdam; Noordhoff, Groningen, x+550 pp.

3. A. Mostowski, *A generalization of the incompleteness theorem*, Fund. Math. **49** (1960–61), 205–232.

HARVARD UNIVERSITY