

## PREDICATIVE PROVABILITY IN SET THEORY

BY S. FEFERMAN<sup>1</sup>

Communicated by D. Scott, January 27, 1966

In [1] we gave an analysis of the notion: *predicatively provable statement of second-order arithmetic*, by means of a certain autonomous progression of theories  $H_\alpha$  formulated in a classical second-order language. The main results of [1] were: (i) Mutual intertranslatability of the progression of  $H_\alpha$  and an autonomous progression of ramified second-order theories  $R_\alpha$ . (ii) Determination of the least non-autonomous ordinal  $\Gamma_0$  of the progression of  $R_\alpha$  in terms of critical functions as:  $\Gamma_0$  is the least solution  $\alpha$  of  $\chi_0^{(\alpha)} = \alpha$  where  $\chi_\beta^{(0)} = \omega^\beta$  and (for  $\alpha \neq 0$ )  $\chi_\beta^{(\alpha)}$  is the  $\beta$ th solution  $\gamma$  of  $\chi_\gamma^{(\alpha)} = \gamma$  for all  $\alpha_1 < \alpha$ . (iii) Explicit determination of a single subsystem IR of classical analysis which has the same theorems as  $\bigcup_{\alpha < \Gamma_0} H_\alpha$ .

We now present a system PS of axioms and rules in the language of set theory which has the following strong relationship with IR:

**THEOREM.** *Under the usual interpretation of second-order arithmetic in set theory, a second-order sentence  $\phi$  is provable in PS if and only if it is provable in IR.*

Necessarily (just as with IR), the system PS does not have predicative justification as a whole: if it did we could extend it by reflection. What is claimed is that the ideal predicativist can recognize as correct each particular axiom and application of a rule of inference.

The language of PS is that of first-order predicate logic with  $=$ ; the basic nonlogical symbols are a binary relation symbol  $\epsilon$  and constant symbols  $\bar{0}$  and  $\bar{\omega}$ . New symbols may be introduced by explicit definitions under certain conditions described below. For any choice of these the notion of a formula in (generalized)  $\Sigma$ -, respectively  $\Pi$ -form is taken as explained in [2, §2]. (Operation symbols  $F(x_0, \dots, x_{m-1})$  are first to be eliminated in favor of relation symbols in a standard way.) We shall also refer to certain formulas  $\phi$  as being in provable  $\Pi \cap \Sigma$ -form on the basis of some given applications of axioms and rules of inference. As usual, a new operation symbol  $F(x_0, \dots, x_{m-1})$  (constant symbol, if  $m=0$ ) can be introduced by a definition  $\phi(x_0, \dots, x_{m-1}, y)$  free of this symbol if it has already been proved that  $\bigwedge x_0, \dots, x_{m-1} \bigvee! y \phi(x_0, \dots, x_{m-1}, y)$ . Then  $F$  is

<sup>1</sup> Part of the research was carried out while the author held a National Science Foundation Senior Postdoctoral Fellowship at the Universities of Paris and Amsterdam. The research was also supported in part by a grant from the U. S. Army Research Office (Durham), DA-ARO(D)-31-124-G655.

introduced by the new axiom,  $\wedge x_0, \dots, x_{m-1}, y [F(x_0, \dots, x_{m-1}) = y \leftrightarrow \phi(x_0, \dots, x_{m-1}, y)]$ . We shall introduce such  $F$  *only* for  $\phi$  which is in  $\Sigma$ -form in its symbols. It is easy to see that the notion of  $\Sigma$ -form is stable under this introduction, i.e., if we eliminate  $F$  from another  $\Sigma$ -formula  $\psi$  involving  $F$  in favor of  $\phi$ , the result is equivalent to a  $\Sigma$ -formula  $\psi_1$  in the former sense ( $F$ -free). The basic reasons for restricting attention to such definitions  $\phi$  will be discussed at the conclusion.

The axioms and rules of PS now follow. (We have made no special effort to find the most elegant form for these.)

I. (EXTENSIONALITY).  $\wedge a, b [\wedge x(x \in a \leftrightarrow x \in b) \rightarrow a = b]$ .

II. (EMPTY SET).  $\wedge x \sim (x \in \emptyset)$ .

III. (UNORDERED PAIR).  $\wedge a, b \forall c \wedge x(x \in c \leftrightarrow x = a \vee x = b)$ .

The operation symbol  $\{a, b\}$  is now introduced by definition, as are  $\{a\} = \{a, a\}$  and  $\langle a, b \rangle = \{\{a\}, \{a, b\}\}$ .

IV. (UNION).  $\wedge a \forall b \wedge x[x \in b \leftrightarrow \forall y(y \in a \wedge x \in y)]$ .

The operation symbol  $\cup a$  is now introduced by definition, as are  $a \cup b = \cup \{a, b\}$  and  $a' = a \cup \{a\}$ .

V. (CARTESIAN PRODUCT).  $\wedge a, b \forall c \wedge x[x \in c \leftrightarrow \forall y, z(y \in a \wedge z \in b \wedge x = \langle y, z \rangle)]$ . The operation symbols  $a \times b$  and, for each  $n$ ,  $a_0 \times \dots \times a_n$  are now introduced by definitions.

VI. (PREDICATIVE REPLACEMENT RULE). *Suppose  $\phi$  is a  $\Sigma$ -formula with free variables  $a_0, \dots, a_{m-1}, x, y$  and suppose that  $\wedge a_0, \dots, a_{m-1} \wedge x \forall y! y \phi(a_0, \dots, a_{m-1}, x, y)$  has been proved. Then infer  $\wedge a_0, \dots, a_{m-1} \wedge b \forall c \wedge y[y \in c \leftrightarrow \forall x(x \in b \wedge \phi(a_0, \dots, a_{m-1}, x, y))]$ .*

To state the further axioms and rules we need some facts about those so far introduced and some further notions. First of all, the following is a derived rule of inference, given Rule VI:

(PREDICATIVE SEPARATION RULE). *Suppose  $\psi$  is a formula with free variables  $a_0, \dots, a_{m-1}, x$  and that  $\psi$  has been proved to be in  $\Pi \cap \Sigma$ -form. Then infer*

$$\wedge a_0, \dots, a_{m-1} \wedge b \forall c \wedge x[x \in c \leftrightarrow x \in b \wedge \psi(a_0, \dots, a_{m-1}, x)].$$

This is easily seen by considering a  $\Sigma$ -formula  $\phi(a_0, \dots, a_{m-1}, x, y)$  which is provably equivalent to  $[\psi(a_0, \dots, a_{m-1}, x) \wedge y = x] \vee [\sim \psi(a_0, \dots, a_{m-1}, x) \wedge y = \emptyset]$ . With this rule we can introduce for any such  $\psi$  an operation symbol  $F(a_0, \dots, a_{m-1}, b) = \{x: x \in b \wedge \psi(a_0, \dots, a_{m-1}, x)\}$ . In particular,  $b \cap a$  and  $b - a$  are introduced in this way. Note also that the predicative separation rule is a direct generalization of the second-order rule HCR (hyperarithmetical comprehension rule) used in [1].

The notion of a set  $f$  representing a function is definable by a completely restricted-quantifier formula  $F_n(f)$ , namely

$$\wedge w [w \epsilon f \rightarrow \forall x, y (x, y \epsilon U w \wedge w = \langle x, y \rangle)] \wedge \wedge x, y_1, y_2 [x, y_1, y_2 \epsilon U U f \wedge \langle x, y_1 \rangle \epsilon f \wedge \langle x, y_2 \rangle \epsilon f \rightarrow y_1 = y_2].$$

Introduce the operation  $\text{Dom}(f) = \{x: x \epsilon U U f \wedge \forall y (y \epsilon U U f \wedge \langle x, y \rangle \epsilon f)\}$ . Then it can be proved that

$$\wedge f, x \forall ! y \{ [\text{Fn}(f) \wedge \langle x, y \rangle \epsilon f] \vee [(\sim \text{Fn}(f) \vee \sim (x \epsilon \text{Dom}(f))) \wedge y = \bar{0}] \}.$$

This allows us to introduce an operation on  $f, x$  yielding this  $y$ , which we denote by  $f(x)$ . We write  $f \uparrow a$  for  $f \cap (a \times U U f)$ .

Notions of ordering and ordinals can be brought in, in the usual way, among these the following. Let  $\mathcal{L}(x, y)$  be any formula with just  $x, y$  free, and which we also write as  $\prec(x, y)$  or  $(x \prec y)$ . Write  $\text{Trans}_\prec(a)$  for  $\wedge x, y [x \prec y \wedge y \prec a \rightarrow x \prec a]$ ,  $\text{Trans}_\prec$  for  $\wedge z \text{Trans}_\prec(z)$ ,  $\text{Conn}_\prec(a)$  for  $x, y [x \epsilon a \wedge y \epsilon a \rightarrow x \prec y \vee x = y \vee y \prec x]$ ,  $\text{Ord}(a)$  for  $\text{Trans}_\epsilon(a) \wedge \text{Conn}_\epsilon(a)$ . Thus  $\text{Ord}(a)$  is a completely restricted-quantifier formula expressing that  $a$  is an ordinal. We express the *no-descending sequence property* of a relation by the sentence  $\sim \forall f \wedge x (x \epsilon \bar{\omega} \rightarrow f(x') \prec f(x))$ , which we designate by  $\text{NDS}_\prec$ . We can introduce  $\text{Pd}(x) = \{y: y \prec x\}$  whenever the formula  $(x \prec y)$  has been proved to be in  $\Pi \cap \Sigma$ -form and  $\wedge x \forall a \wedge y (y \prec x \rightarrow y \epsilon a)$  has been proved. We now complete the list of axioms and rules of PS with the following VII-X.

VII. ( $\omega$ ).  $\text{Ord}(\bar{\omega}) \wedge \bar{0} \epsilon \bar{\omega} \wedge \wedge x (x \epsilon \bar{\omega} \rightarrow x = \bar{0} \vee \forall y (x = y'))$ .

VIII. (NO  $\epsilon$ -DESCENDING SEQUENCES).  $\text{NDS}_\epsilon$ .

IX. (TRANSFINITE INDUCTION RULE). *Suppose  $(x \prec y)$  has been proved to be in  $\Pi \cap \Sigma$ -form and that also  $\text{NDS}_\prec$  has been proved. Then infer for any formula  $\theta(x, \dots) = \theta(x, a_0, \dots, a_{m-1})$ :*

$$\wedge a_0, \dots, a_{m-1} \{ \wedge x [ \wedge y (y \prec x \rightarrow \theta(y, \dots)) \rightarrow \theta(x, \dots) ] \rightarrow \wedge x \theta(x, \dots) \}.$$

X. (TRANSFINITE RECURSION RULE). *Suppose  $(x \prec y)$  has been proved to be in  $\Pi \cap \Sigma$ -form. Suppose further that (i)  $\text{NDS}_\prec$ , (ii)  $\text{Trans}_\prec$  and (iii)  $\wedge x \forall a \wedge y (y \prec x \rightarrow y \epsilon a)$ , have all been proved. Suppose finally that  $G(x, f, a_0, \dots, a_{m-1})$  is an already introduced operation symbol. Under these conditions infer*

$$\wedge a_0, \dots, a_{m-1} \wedge x \forall f \wedge y [y \prec x \rightarrow y \epsilon \text{Dom}(f) \wedge f(y) = G(y, f \uparrow \text{Pd}(y), a_0, \dots, a_{m-1})].$$

Using VIII and X we get definition by transfinite recursion on the ordinals in the usual way. We can also get  $\epsilon$ -recursion (e.g., definition of the rank function) by introducing the transitive closure  $\prec$  of  $\epsilon$ . By applying IX and X to certain relations  $\prec$  in the natural numbers  $\bar{\omega}$ , it is a straightforward matter to see that every theorem of IR is a theorem of PS. The proof of the converse involves some new ideas which will be presented elsewhere.

It is also seen that PS is a subsystem of Zermelo-Fraenkel set theory excluding the power-set axiom but including the axiom of choice (the axiom of dependent choices seems necessary to verify IX). Therefore PS can be consistently extended in many natural ways to stronger theories T, for example by adding (i) the axiom of constructibility [3] and/or (ii) the hypothesis that all sets are countable. Some of these T will continue to be conservative extensions of IR with respect to arbitrary second-order statements as in our main theorem; this is the case for PS+(ii). Others will be conservative only with respect to purely arithmetical statements and some small classes of analytic statements; this is the case for PS+(i). It would be interesting to find out which are the extensions of the former kind.

From the point of view of a model-theoretic approach to predicativity, the essential characteristic of PS is that with any theorem  $\phi$  of PS can be associated a finite sequence  $G_i(x_0, \dots, x_{k_i-1})$  of operations such that any transitive collection  $\mathfrak{M}$  of sets which is closed under the  $G_i$  is a model of  $\phi$  and further the definitions of the  $G_i$  in  $\mathfrak{M}$  are *absolute*. More precisely, each  $G_i(x_0, \dots, x_{k_i-1})$  has been introduced by a  $\Sigma$ -formula  $\phi_i$  for which  $\bigwedge x_0, \dots, x_{k_i-1} \forall y \phi_i(x_0, \dots, x_{k_i-1}, y)$  has been proved. Owing to the  $\leq$ -persistence of  $\Sigma$ -formulas [2], if  $\mathfrak{M}, \mathfrak{M}^*$  are transitive and closed under the operations in  $\phi_i$  and  $\mathfrak{M} \subseteq \mathfrak{M}^*$ , then the value ascribed by  $\phi_i$  in  $\mathfrak{M}$  to  $G_i(a_0, \dots, a_{k_i-1})$  for  $a_0, \dots, a_{k_i-1} \in \mathfrak{M}$  is the same as its value in  $\mathfrak{M}^*$ . It would be especially interesting to know what extensions T of PS also have as models those  $\mathfrak{M}$  satisfying such absolute closure conditions.

One can also give a notion of a *provably  $\leq$ -persistent* formula  $\psi$  relative to any set  $\mathcal{A}$  in the symbols used here, namely by extending the language to include a new sort of variable  $x^*, y^*, \dots$  and demanding that  $\mathcal{A} \cup \mathcal{A}^* \cup \{ \bigwedge x \forall x^* (x = x^*) \} \cup \{ \bigwedge x, y^* [y^* \epsilon x \rightarrow \forall z (y^* = z)] \} \vdash (\psi \rightarrow \psi^*)$ . It follows by the argument for Theorem 1 of [2] that then for some  $\Sigma$ -formula  $\theta$ ,  $\mathcal{A} \vdash (\psi \leftrightarrow \theta)$ . The case is similar for a notion of *provably  $\leq$ -invariant formula*. This shows that no new theorems would be obtained from PS if throughout in the statement of the rules we used provably  $\leq$ -persistent and provably  $\leq$ -invariant formulas instead of  $\Sigma$ -formulas and provably  $\Pi \cap \Sigma$ -formulas, respectively.

#### BIBLIOGRAPHY

1. S. Feferman, *Systems of predicative analysis*, J. Symbolic Logic 29 (1964), 1-30.
2. S. Feferman and G. Kreisel, *Persistent and invariant formulas relative to theories of higher type*, Bull. Amer. Math. Soc. 72 (1966), 480-485.
3. K. Gödel, *The consistency of the axiom of choice and the generalized continuum hypothesis with the axioms of set theory* (2nd printing), Annals of Mathematics Studies No. 3, Princeton Univ. Press, Princeton, N. J., 1951.