ORDINAL LOGICS AND THE CHARACTERIZATION OF INFORMAL CONCEPTS OF PROOF

By G. KREISEL

1. Introduction

By Gödel's (first) incompleteness theorem the informal notion of *arithmetic truth* cannot be formalized in the originally intended sense of 'formalization': there is no recursive enumeration of the true formulae in the notation of classical arithmetic. All one needs here of the concept of truth is that each sentence or its negation is true, and that each theorem is true. On the other hand, mere incompleteness does not generally preclude the formalization of other informal notions, e.g. certain informal methods of proof. For, in general, it is not to be expected that every formula is either provable or refutable by such methods, even if the notation is restricted. This applies to the main subject of the present lecture, namely *finitist proof* in arithmetic as described by Hilbert–Bernays, vol. 1, and also to its subsidiary subject, namely *predicative proof*.

At first sight, another formulation of the incompleteness theorem, also due to Gödel[2], seems to prohibit such a formalization: if $P(\ulcorner A \urcorner)$ is a provability predicate (enumeration of theorems) then, for certain A, $P(\ulcorner A \urcorner) \rightarrow A$ is not provable in the system. Yet another formulation is this: for certain $A(n)$, $P(\ulcorner A(0^{(n)}) \urcorner)$ is provable in the system with free variable n, but not $A(n)$. Now consider *finitist proof*: if $P(\ulcorner A \urcorner)$ has been recognized by finitist means to be the provability predicate of a (partial) formalization, say Σ_μ, of finitist mathematics, and $P(\ulcorner A(0^{(n)}) \urcorner)$ has been established by finitist means then, on the intended meaning of free variables, $A(n)$ is finitistically established. In other words, Σ_μ is incomplete and can be extended to Σ_ν in which $A(n)$ is provable. Similar considerations apply to several other informal concepts of proof.

Thus the conclusion from the reformulation of the first incompleteness theorem is this: if the notion of finitist proof is capable of formalization at all, its proof predicate must not be recognizable as such by finitist means.

Remark. The difference between a mere extensional enumeration of the theorems and a provability predicate which can be recognized as

such, is familiar from the need for the derivability conditions in the second incompleteness theorem. Thus, if $P(m, n)$ is a proof predicate for a consistent system (S), $\overline{n}$ denotes the negation of n, and $Con\,S$ denotes $(m) \neg P(m, \ulcorner 0 = 1 \urcorner)$ then $P_1(m, n)$, i.e. $P(m, n)\ \&\ (p)[p < m \rightarrow \neg P(p, \overline{n})]$ is also a proof predicate, but $Con_1 S$ is provable in (S) itself.

To give a precise treatment of this idea of recognizing a proof predicate as such we shall consider formal systems whose constants are not only numerical terms and function symbols, but also *proof predicates*. This is independently justified by the accepted sense of *finitist* according to which finitist proofs can themselves be the subject-matter of finitist reasoning; also it is in accordance with Heyting's and Gödel's view of the place of *proof* as an object of a theory of constructivity. Just as function symbols may be introduced only after they have been shown to represent a computation procedure, so proof predicates may be introduced only after they have been shown to represent an extension procedure as above.

Roughly speaking, we propose to characterize finitist proofs by a precisely defined *class* of formal systems, namely the least class of systems Σ_μ containing a certain basic finitist apparatus and closed under the principle: if a proof predicate P_ν is recognized as such in a system Σ_μ of the class then the corresponding system Σ_ν also belongs to the class.

Our main result is that, in a precise sense, the theorems of this class are co-extensive with those of classical number theory when the latter is suitably interpreted. (Below this is shown only for formulae $(Ex)\,A(n, x)$ with primitive recursive A.) Since each of our extensions is finitist this means at least that finitist results include essentially those of classical number theory.

The important problem of establishing the converse is not discussed here. But we note at least the following point: though each extension is finitist, the general extension principle cannot be regarded as finitistically evident since it is framed in terms of the concept of finitist proof which has no place in finitist mathematics; hence at least the obvious formalization of our class is not finitist.

An important tool in this work is the use of ordinal notations for specifying formal systems, first discussed by Turing[6].

2. Ordinal logics

The following information is needed only for a comparison between the present approach and Turing's. Turing uses a narrow and a wide definition of logics based on ordinals.

The narrow one consists of extending some system, e.g. classical number theory Z or *Principia Mathematica* with a recursively enumerable provability predicate $P(\ulcorner A \urcorner)$ ($\ulcorner A \urcorner$ is provable) by all formulae $P(\ulcorner A \urcorner) \to A$, and iterating the procedure. To ensure recursive enumerability of the theorems of each extension, recursive ordinals are used to describe the iteration. In modern language Turing's construction can be formulated as follows: for each $n \in O$, O denoting the class of recursive ordinals, we set up recursion equations relating P_n with P_m, $m <_O n$, of a kind which have a recursively enumerable solution by results of Kleene[3]. In fact, the only consequence of $n \in O$ needed here is that these recursion equations have a unique solution, and so the full degree of undecidability of $n \in O$ is not used.

Feferman[1] has sharpened these results by giving an unambiguous method for constructing such proof predicates, and has shown that, if one starts the iteration with Z, *the class of theorems is just that obtained by adding to Z all true formulae $(x)\,A(x)$ with A primitive recursive.* In consequence: (i) the provability predicate for the whole iteration is arithmetically definable while the proof predicate is not, since, at least with the obvious coding of proofs, it would allow us to decide $n \in O$; (ii) exactly the same functions are *provably* recursive in the ordinal logics as in Z since, if $\vdash_Z [(x)\,A(x) \to (x)\,(Ey)\,B(x,y)]$ and $(x)\,A(x)$ is true,

$$\mu_y B(x,y) = \mu_y[B(x,y) \vee \neg A(y)]$$

and the latter is provably recursive in Z.

The wide definition simply associates a formal system with every recursive ordinal, e.g. for each $n \in O$, we add to Z the 'principle of transfinite induction'

$$\frac{\{y <_O n \,\&\, (x)\,[x <_O y \to A(x)]\} \to A(y)}{y <_O n \to A(y)}.$$

As shown by Wang, Shoenfield and myself[5], every true arithmetic formula is provable in one of these systems even if we require $|n| < \omega^\omega$. Conversely, for each $\alpha < \omega^\omega$, O_α is arithmetic. This is obtained by showing $\deg H_{n_0} < \deg O_{\omega^\omega}$ and then analysing the argument proof-theoretically.

The following differences between Turing's aims and consequently his methods on the one hand and ours on the other seem worth noting.

(1) Turing aimed at completeness, we do not. The results above show conclusively that on the narrow definition we do not get completeness even for two-quantifier formulae with primitive recursive scope and that on the wider definition we get it too easily. However, with respect

to the latter the following interesting problem remains open: to assign non-constructively a unique notation for each recursive ordinal and reopen the problem of completeness.

(2) For Turing the narrow definition represented simply one means of extension: perhaps the most that can be said for it is that at each stage it is stronger than merely adding consistency: e.g.

$$P(\ulcorner \neg \mathit{Con}\, Z \urcorner) \to \neg\, \mathit{Con}\, Z$$

cannot be proved in Z from $\mathit{Con}\, Z$ since otherwise $Z \cup \{\mathit{Con}\, Z\}$ could be proved consistent in itself. We have to choose the extension principle in relation to the informal proof predicate under investigation. As it happens it turns out that the extension principle is intimately related to a kind of modal interpretation of Heyting's arithmetic discussed below.

(3) By Feferman's results the choice of starting system is not important for Turing's aim of achieving completeness since no system will do it. We have to choose one which admits of an interpretation as a partial system of the informal concept under discussion.

(4) We cannot use oracles to supply ordinals or, more specifically, to ensure the unique solubility of the recursion equations for the proof predicate, but this has to be proved in one of the earlier systems.

3. Finitist proofs (general description)

The description of our class of systems, and particularly the sketches of proofs, will have to be brief. It is hoped that the full details will be published before too long.

Our variables range over the natural numbers.

We have three kinds of constants: (i) numerical terms $0, \ldots$, (ii) function symbols with one or more arguments including the successor function, relation symbols $<$ and $=$, and (iii) proof predicates $P_\mu(m, n)$.

Our formulae are prime formulae built up of the constants and variables above in the usual way, (quantifier-free) truth functional combinations of them, existential quantification, where $(Ex)\, A(n, x)$ below usually denotes a string of existential quantifiers

$$(Ex_1) \ldots (Ex_p)\, A(n, x_1, \ldots, x_p),$$

A quantifier-free, and single truth functional combinations of the latter. This process is not iterated in accordance with the finitist requirement that no premises relating to an infinite totality may be used.

The rules of proof are first those of the classical propositional calculus, the transposition rules for the existential quantifier,

$$\frac{(Ex)\,A(x) \vee (Ex)\,B(x)}{(Ex)\,[A(x) \vee B(x)]} \quad \frac{(Ex)\,A(x)\ \&\ (Ey)\,B(y)}{(Ex)\,(Ey)\,[A(x)\ \&\ B(y)]} \quad \frac{\neg(Ex)\,A(x)}{\neg A(n)}$$

$$\frac{(Ex)\,A(x) \to (Ey)\,B(y)}{(Ey)\,[A(n) \to B(y)]}$$

and conversely, with the obvious condition on n,

$$\frac{(Ex)\,A(n,x)}{A[n,\tau(n)],\ m < \tau(n) \to \neg A(n,m)},$$

if $A(n,x)$ quantifier-free, and similarly for strings of x,

$$\frac{A(\alpha) \vee B}{(Ex)\,A(x) \vee B}, \quad \frac{A(n)}{A(\alpha)}.$$

Also the schema of identity, and induction

$$\frac{A(0),\ A(n) \to A(n')}{A(n)},$$

and the axioms for the successor function. This is the basic finitist apparatus. The existential quantifier is constructive and because of the restriction on the formulae the truth functional interpretation of the logical connectives is not problematical.

Remark. It seems likely that this formulation could be considerably simplified.

In addition to induction, which is unproblematical, we use primitive recursive definitions

$$\phi(0) = \alpha, \quad \phi(n') = \psi[n, \phi(n)], \tag{1}$$

when α and ψ are already introduced (and ϕ not).

(1) can be interpreted in two ways, either as the *existence of a unique function* ϕ satisfying (1) for all n, or as a set of equations from which, for each number $0^{(n)}$, $\phi(0^{(n)}) = 0^{(m)}$ can be obtained for a unique m from (1) and the computation rules for α and ψ by a *finite number of substitutions*. The difference is clear: e.g.

$$f(n) = 2f(n+1)$$

uniquely defines the function $f(n) \equiv 0$ on the first interpretation since $f(n+k) = 2^{-k}f(n)$ and f is integer valued, but not on the second. Only the second interpretation has finitist sense. Bernays has established by finitist methods the permissibility of (1).

Remark. It would be more elegant to start off with a pure equation calculus, and an elementary semiotic, and give a formal proof for each primitive recursive definition that it uniquely defines values for $\phi(0^{(n)})$ in a pure equation calculus.

To deal with so-called definitions by transfinite induction, which are used below, we note a

Lemma. Let τ_1, τ_2 be two functions and define a tree as follows: the descendants of a node p are those values of $\tau_1(p)$, $\tau_2(p)$ which are different from 0. Then, if $\lambda(n)$ is a bound for the length of the tree with vertex n, then there is a function f, primitive recursive in λ, τ_1, τ_2 and g, which satisfies

$$f(0) = \alpha, \quad f(n) = g\{n, f[\tau_1(n)], \quad f[\tau_2(n)]\}.$$

Define $F(n, m)$ by induction w.r.t. m:

$$F(0, m) = \alpha.$$

If the tree with vertex n has length $> m$ then $F(n, m) = 0$. If the tree with vertex n has length $\leqslant m$ then

$$F(n, m) = g\{n, F[\tau_1(n), m-1], F[\tau_2(n), m-1]\}.$$

$F[n, \lambda(n)]$ is our function.

Finally, we come to the conditions to be satisfied by a proof predicate. In the definitive version it is essential to give a careful numbering of expressions, particularly of function symbols.

$P_0(m, n)$ is the usual proof predicate for primitive recursive arithmetic, as extended above.

We use some notation for ordinals, e.g. primitive recursive orderings of primitive recursive subsets of the natural numbers.

Remark. It seems plausible that a wide range of alternative notations gives essentially the same results.

As in [1] we regard each system Σ_μ ($\mu = 2^l 3^k$) as made up of the basic finitist apparatus together with a sequence $\lambda n \pi_l(k, n)$ of axioms (the extensions), where l is a number of a primitive recursive relation which has been proved in Σ_μ to be an ordering whose first element is 0, and k is in the field of l.

If a is the successor of 0, $\pi_l(a, n)$ is an enumeration of formulae $\ulcorner A(p) \urcorner$ such that $(Ey)\, P_0(y, \ulcorner A(0^{(p)}) \urcorner)$ is provable in the basic finitist apparatus, and $P_0(m, n)$ means that m is a proof of n in this system; $P_\mu(m, n)$ means that m is a proof of n in this system from the appropriate set of axioms $\hat{n}\pi_r(s, n)$.

To establish an extension $<l, k>$ in Σ_μ we require that there be a term $\pi_l(k, n)$ in Σ_μ for which the following two statements can be proved in Σ_μ:

(i) For each k and n, a formula $\ulcorner A(p)\urcorner$ has the number $\pi_l(k, n)$ if, and only if, for some $k' <_l k$, either $\pi_l(k, n) = \pi_l(k', n')$ for some n', or it follows (in the basic finitist apparatus) from $\hat{n}\pi_l(k', n)$ that $\ulcorner A(0^{(p)})\urcorner$ follows from $\hat{n}\pi_l(k', n)$ with free variable p, i.e. it is proved in $\Sigma_{<l,k'>}$ that

$$(Ey)\, P_{<l,k'>}\ (y, \ulcorner A(0^{(p)})\urcorner).$$

Note that such a proof in Σ_μ provides a term $\tau(k, n)$ specifying k' and the axioms of $\hat{n}\pi_l(k', n)$ actually needed.

(ii) There is a proof in Σ_μ that each tree with vertex $< k, n >$ is finite if the descendants of a node N are the axioms specified by $\tau(k, n)$.

Our proposal is to identify finitist proofs in arithmetic with *the least class of systems Σ_μ containing primitive recursive arithmetic with a constructive existential quantifier, and if $P_l(k, m, n)$ is proved to be a proof predicate in Σ_μ, then $\Sigma_{<l*k>}$ also belongs to the class.* Call this class $\mathscr{H}$.

It is occasionally natural to include variables f for free choice sequences in finitist mathematics, and, if $(Ex)\, A(f, x)$ has been proved, to introduce a functional $\tau(f)$ with $A[f, \tau(f)]$. It would be desirable to develop our theory for this extended notation, but we have not done so here in order to keep the basic treatment of finitist proof strictly number theoretical. The use of variables for free choice sequences would permit a more elegant formulation of the conditions to be satisfied by a proof predicate, namely that $<_l$ be a well-ordering. Here it is necessary to distinguish between weak well-ordering when we have proved that every descending sequence is finite, i.e. $(Ex)[f(x+1) \not< f(x)]$ and strong well-ordering (needed above) where we have to prove that every descending binary tree defined by a free choice sequence is finite when $f(1)$ is the value of the vertex; $f(2), f(3)$ at the next level; $f(4), \ldots, f(7)$ at the next, etc. It is clear that in a particular finitist system an ordering may be provable to be a weak well-ordering, but not a strong one, since otherwise one would get up to an ϵ-number in each system.

4. Finitist proofs (results)

Our main results are:

(i) Every function $\tau(n)$ of the class $\mathscr{H}$ is provably recursive in classical arithmetic Z.

(ii) If $A(n, m)$ is a quantifier-free formula whose non-logical constants are provably recursive functions of Z and $(Ey)\, A(n, y)$ is provable in Z then it is also provable in some system of $\mathscr{H}$.

Thus essentially exactly the same formulae in the notation of $\mathscr{H}$ are provable in $\mathscr{H}$ and in classical arithmetic.

(i) is seen easily by constructing a truth definition in Z for each system Σ_μ of $\mathscr{H}$, and observing that, if P_ν can be proved to be a proof predicate in Σ_μ, a truth definition can be defined for Σ_ν in Z too.

Two steps are used to establish (ii).

First, we use the result[4] that exactly the same functions are provably recursive in Z and in Heyting's arithmetic.

Secondly, we show that Heyting's arithmetic may be regarded as the metamathematics of $\mathscr{H}$ in the following precise sense: Gödel's interpretation of Heyting's propositional calculus in modal logic[2] can be extended to Heyting's arithmetic with the additional restriction that the modal operator B ('is provable') is replaced by: is provable in $\mathscr{H}$.

In detail: we leave prime formulae unchanged; if A^*, B^* are the translations of A and B, $(A \,\&\, B)^*$ is $A^* \,\&\, B^*$, $(A \overrightarrow{\vee} B)^*$ is replaced by $(E\mu)$ [P_μ is a proof predicate and $P_\mu(\ulcorner A_1^{*\urcorner})] \overrightarrow{\vee} (E\mu)$ [P_μ is a proof predicate and $P_\mu(\ulcorner B_1^{*\urcorner})$], negation is not needed, $[(x)\,A(x)]^*$ is $(E\mu)$ [P_μ is a proof predicate and $P_\mu(\ulcorner A_1^*(x)^\urcorner)$], $[(Ex)\,A(x)]^*$ is $(Ex)\,A^*(x)$, where A_1^* is obtained from A^* by replacing each *free* variable n in A^* by $0^{(n)}$. Note that however complicated the logical structure of A may be, e.g. however many iterated implications it may contain, A^* will be an assertion of the form: a certain concretely specified formula is provable in some Σ_μ. Also note that an assertion $P_\nu\{(E\mu)$ [P_μ is a proof predicate and $P_\mu(\ulcorner A^{*\urcorner})$]$\}$ means that P_μ is *proved* to be a proof predicate in Σ_ν.

As in Gödel's original work[2] the axioms of the propositional calculus go into true assertions on this interpretation, the rules for quantifiers are straightforward and so are the equality axioms and those for constant functions. The induction axiom requires essentially that the union of systems $\Sigma_{\nu(n)}$ is again a system, namely, suppose

$$P_\tau(\ulcorner A^*(0)^\urcorner), \quad P_\tau\{\ulcorner (E\mu)\,[P_\mu \text{ is a proof predicate and } P_\mu(\ulcorner A^*(0^{(n)})^\urcorner)] \rightarrow$$

$$(E\mu')\,[P_{\mu'} \text{ is a proof predicate and } P_{\mu'}(\ulcorner A^*(0^{(n+1)})^\urcorner)]^\urcorner\}$$

then μ' may be replaced by $\mu'(n+1)$, where $\mu'(n)$ is a term of Σ, $\mu'(0) = \tau$, and there is a proof that each $\mu'(n)$ is a number of a proof predicate permissible in Σ_τ. What we need now is the union of $\Sigma_{\mu'(n)}$, say Σ_{μ^*}, when we have

$$P_\tau(\ulcorner P_{\mu^*} \text{ is a proof predicate and } P_{\mu^*}(\ulcorner A^*(n)^\urcorner)^\urcorner).$$

Recalling the condition on proof predicates this is essentially equivalent

to saying that the union of well-orderings is well-ordered, which can certainly be established in each of our systems.

The conditions on Gödel's Bp are

$$Bp \to BBp, \quad [Bp \,\&\, B(p \to q)] \to Bq \quad \text{and} \quad Bp \to p.$$

The first two conditions are clearly true for our interpretation, and so is the third by our closure condition.

The translation of $(Ey)\,A(n,y)$ asserts the existence of a proof of $(Ey)\,A(n,y)$ in $\mathscr{H}$.

Note incidentally that the application of the classical propositional calculus is much more reasonable in our translation than in general modal logic. For it is not at all clear that it always makes sense to say of a proposition that it is provable without further restriction, while our provability statements are purely existential assertions about decidable relations, just the kind of statements in our basic finitist apparatus.

5. Predicativity

Wang[7] gives a reasonable indication for treating predicative *proof*. The following modifications are needed: (i) To obtain the ordinals needed for the extension it seems best to add free function variables to his systems and extend the notion of term; a number of an ordering of a system is said to be a notation for a (provable) well-ordering $<$, if $(Ex)\,[f(x+1) \not< f(x)]$ can be proved in the system. (ii) The extension principle is now: if $<$ is proved to be a well-ordering in a system Σ_μ, then the system with types indexed by $<$ is said to be proved in Σ_μ to be permissible (as a predicative proof predicate). Here, too, though each extension is predicative provided $<$ has been recognized by predicative means to be a well-ordering, the general extension principle is not since the concept of predicative proof has no place in predicative mathematics in the present sense.

It should be noted that the new extension principle includes the previous one since now a truth definition can be introduced, and by means of a truth definition: from $P_\mu(\ulcorner A(0^{(n)})\urcorner)$ follows $A(n)$. For, by induction

$$(n)\,P_\mu(\ulcorner A(0^{(n)})\urcorner) \to (n)\,T_\mu(\ulcorner A(0^{(n)})\urcorner)$$

and

$$(n)\,T_\mu(\ulcorner A(0^{(n)})\urcorner) \leftrightarrow T_\mu(\ulcorner (n)\,A(n)\urcorner).$$

I have no information about the least ordinal not obtained by these extensions starting from Z.

The notion of predicative *definability* does not seem to come into the

present scheme at all. To fix ideas we consider number theory and a theory with two types of variables, for natural numbers and functions (of the natural numbers into the natural numbers). We call number theory predicative because it has a *unique* minimal model, i.e. a model no subset of which is also a model. We call predicative those theories of second order which have a unique minimal model in which the individuals are the natural numbers.

Evidently hyperarithmetic set theories are predicative in this sense. Non-trivial group theory (with at least two elements) clearly is not predicative in the present sense because there exist infinitely many non-isomorphic minimal models. It seems possible that there are set theories whose unique minimal model properly includes all hyperarithmetic sets. Naturally, this definition is itself quite impredicative.

6. General remarks

The present study is at the same time a theory of extensions of formal systems and of a single concept, namely the totality of such extensions.

In its former role it seems to be complementary to the current tendencies of invoking oracular identities or even statistical principles proposed for machines which are to 'accept' a rule of inference if it has been successful a certain number of times. Instead we have here mathematical principles governing extensions, e.g. for getting from a fragment of finitist mathematics to a larger one. While this principle is devoid of finitist sense it is, for example, constructive. At least as far as the actual work involved in extending the whole body of mathematics is concerned, the extensions of finitist mathematics as seen by a finitist seem to be a closer model than the proposals for learning machines. Naturally, if they are a good model it is also understandable why there is trouble about a mathematical theory of the principles governing extensions of the whole body of mathematics.

In its latter role the present theory is closer to old-fashioned number theory and analysis than to modern algebra. For we do not consider all models of a formal system, but minimal models (satisfying certain closure conditions).

REFERENCES

[1] Feferman, S. Ordinal logics re-examined, and on the strength of ordinal logics. *J. Symbol. Logic* (*Abstr.*), 23 (1958).

[2] Gödel, K. Eine Interpretation des intuitionistischen Aussagenkalküls. *Ergebnisse eines mathematischen Kolloquiums*, 4, 39–40 (1931–32).

[3] Kleene, S. C. On the form of predicates in the theory of constructive ordinals. *Amer. J. Math.* 66, 41–58 (1944).

[4] Kreisel, G. Mathematical significance of consistency proofs. *J. Symbol. Logic*, 23, 155–182 (1958).

[5] Kreisel, G., Schoenfield, J. R. and Hao Wang, *H*, *O* and *W* (to be published).

[6] Turing, A. M. Systems of logic based on ordinals. *Proc. Lond. Math. Soc.* (2), 45, 161–228 (1939).

[7] Hao Wang. The formalization of mathematics. *J. Symbol. Logic*, 19, 241–266 (1954).