Notre Dame Journal of Formal Logic
Volume 29, Number 2, Spring 1988

Epistemic Set Theory

WILLIAM N. REINHARDT

It is the purpose of this paper to formulate axioms for Gödel's modal operator B for provability (see [3], [8]) in the context of set theory. This provides a framework for consideration of the Post–Turing thesis which is more adequate than arithmetic with B, where the thesis can only be expressed as a schema. The framework also provides a new perspective on ordinal notations.

We begin with a brief discussion of the problems to be overcome in extending the arithmetic case to set theory. The relevant special fact about arithmetic is that each natural number has a canonical conceptualization. In set theory this already fails for sets of natural numbers and for real numbers. Thus there is no clear meaning to formulas $BF(x)$, where x is a set variable. We could arbitrarily confine quantification into the modal context to variables ranging over the natural numbers (or over set theoretic representations of them), but this seems excessively restrictive. In particular, if we wish to state the thesis that all intuitively decidable sets of natural numbers are recursive, this requires quantification over an arbitrary property P within the scope of B:

$$\forall P(\forall n(BP(n) \text{ or } B \text{ not } P(n)) \rightarrow \{n \mid P(n)\} \text{ is recursive}).$$

Thus we require a theory of the sort of properties which can meaningfully appear within the scope of B. We shall call such properties concepts; the variables which can occur within the scope of B will be those which range over concepts and certain combinations of concepts. Thus we shall want to extend the set of basic concepts C_b to C including combinations of concepts as well.

Now any theory of concepts must of course deal with the paradoxes. This will be done here via a theory of significant combinations, following ideas of Gödel ([4], pp. 228–229) and Kripke ([6]). (In a nutshell, conceptual combinations correspond to significant syntactic combinations.) König's paradox concerning the first ordinal which is not definable becomes particularly acute, since "definable" means "given by a concept".

We now turn to the treatment of C_b and C. Since the most natural way to proceed here would be to take both C_b and C to be urelements from the point

Received April 16, 1984; revised July 11, 1986

of view of set theory, with C generated from C_b by concept-combining functions of some sort, we expend some effort explaining why the procedure taken here (which involves only familiar combinations) should accomplish the same ends.

Let C_b be any set. We wish to introduce the domain of computable combinations C over C_b. The case of interest to us is where C_b is the set of concepts, more specifically the basic concepts (such as relations in intension) which can in principle be used by an idealized mathematician in making definitions. Thus one thinks of C_b as a set of objects which are directly available to an idealized mind. This means simply that a canonical means of reference is available for objects in C_b, in the way that we have canonical names for the natural numbers. In the latter case we have more, namely a canonical enumeration of the domain as well. Here we do not assume anything about such an enumeration. As in the case of the natural numbers, we do not ordinarily care whether we deal with the natural numbers themselves (whatever they may be) or merely with indices for or representations of them; what concerns us is primarily certain structure over them. In our case two aspects of the structure are of particular interest. The first is that of predication, the relation that holds between a concept and an object when the object falls under the concept (and variations on this theme). We shall return to this presently. The second is the structure induced by the presence of an idealized mind, specifically notions of provability. (We do not of course identify the presence of such a mind with the existence of such structures, any more than we would identify the existence of a physical object, say a pendulum, with the existence of a Hilbert space with appropriate structure.)

By C we understand C_b together with the sets which are hereditarily finite over C_b. For the kinds of quasi-syntactic combinations we are interested in, one might feel that sequences over C_b are more natural than sets. However, the two frameworks appear to be equivalent (provided we ignore issues of efficiency or feasibility) and the reduction of sequences to sets is standard and well known, so we choose to take sets over C_b as our basic kind of combination. We observe that for the Von Neuman ordinal ω, $\omega \subseteq C$, and that the elements of C_b will appear as urelements in the structure $(C, \in_C)$ (where $\in_C = \{(x,y) | y \in C \sim C_b \;\&\; x \in C \;\&\; x \in y\}$) provided that the elements of C_b happen to be either infinite or genuine urelements.

We remark that in considering combinations of concepts our interest is not primarily with mathematical combinations and permutations, but rather with combinations of the sort that form new concepts from old ones. We expect, however, that such combinations will be images of mathematical combinations, and in many cases in correspondence with mathematical combinations. Thus, rather than construct a theory of the appropriate combining functions from scratch, we shall build it on top of what are essentially ordinary syntactic combinations. As indicated above, the theory is that conceptual combinations correspond to syntactic combinations which are significant. Thus the combining functions we seek will be partial functions. This approach of course follows the previously mentioned ideas of [4] and [6]. It is perhaps worth remarking that we thus have the advantage of working in a familiar setting (syntax); the disadvantage is that it is technically tedious and messy.

It should be fairly clear how to develop the syntax of a language with symbols for the concepts in C. For definiteness we indicate here a specific treatment. First, however, we indicate the primitive notions we adopt, indicate how these will suffice for defining the basic notions of the theory of concepts and relations, discuss briefly some alternative choices of primitives, and indicate the axioms we shall adopt.

Our primitives will be the usual logical ones, $\neg$, $\rightarrow$, $\forall$, $=$, as well as the provability operator B (a one-place sentential connective), together with the usual membership relation, $\in$. In addition we have the one-place predicate C (for concepts) which we have discussed above, and one-place predicates T and S for truth and significance. (Alternatively we could take predicates for truth and falsity.) We shall need, in addition to the usual variables, variables understood to range over C. We call these intensional variables. Finally, it will be convenient to have the usual set theoretic terms in our language. If we wish set theory with urelements (in particular if we wish to treat C_b as consisting of urelements), we of course also need a one-place predicate for "is a set".

We shall want variable binding terms $[\theta]_x$, $[\theta]_{x,y}$, etc. for concept formation (these are analogous to $\{x|\theta\}$, $\{(x,y)|\theta\}$, etc.); as indicated above, the concepts formed will be viewed as set theoretic combinations, and hence the terms denoting them will be set-theoretic ones, and not new primitives. Notice that $[\mathrm{red}(x)]$ corresponds almost exactly in usage to the English phrase "that x is red". Thus the brackets should perhaps be one of the basic things we axiomatize (as George Bealer has suggested in [1]). The axioms we give for truth will be written in terms of the brackets and certain operations $\dot{\&}$, $\dot{\neg}$, etc. representing conjunction, negation, etc. of propositions. These however (or something close) can be defined using truth and the brackets:

$$(x \mathbin{\dot{\&}} y) = [T(x) \mathbin{\&} T(y)]$$

$$(\dot{\neg} x) = [\neg T(x)].$$

It might appear that we still need to assume predication as a new primitive notion. Under fairly mild conditions on predication, this is not the case. It is true that if R is a relation symbol for a property r, then truth of the sentence $[R(x)]$ asserting that r holds of x is explained in terms of the predication relation, and the notion of significant applicability:

$$T[R(x)] \leftrightarrow x\eta r \mathbin{\&} S[x\eta r],$$

and similarly for falsity:

$$F[R(x)] \leftrightarrow S[x\eta r] \mathbin{\&} \neg(x\eta r).$$

But in case we assume

$$x\eta r \rightarrow S[x\eta r],$$

then these show that the notions T and F are enough: from them we may recapture η, S:

$$x\eta r \leftrightarrow T[R(x)]$$

$$S[x\eta r] \leftrightarrow T[R(x)] \text{ or } F[R(x)].$$

This of course does not give r itself given R, but it gives the structure we are concerned with; and in case the relations themselves are present, it does give predication itself.

There is some reason to view properties and predication as more basic than sets and membership; we have already remarked on the naturalness of the concept formation operator as opposed to a collection of somewhat ad hoc syntactic operations. Such a development would require allowing properties more general than those we have called concepts, but this is natural enough. In particular it is tempting to view membership as being simply predication restricted to sets. There is one way to do this which perhaps does not do great violence to the conception of a set as a pure combination whose constituents are its elements. (For this notion of set see [2], pp. 275–276). Bernays takes it for granted that we understand the relevant notion of combination in the finite case, and understand sets in the general case by analogy with this. In the famous passage in [5], pp. 262–263, on what has come to be called the iterative notion of set, Gödel uses also the phrase "combination of any number of x's" and says that "random sets are not excluded" (footnote 14), which seems to be getting at the same feature that Bernays puts in terms of "independent determinations". Neither author speaks of the constituents of a set, but both are at pains to make clear that the set is detached from any of its definitions. Sets so conceived are "obtained from" not their definitions or some mental construction or "link . . . with the reflecting subject" ([2], p. 275), but rather from their elements, elements of those, etc. In particular, sets of integers are obtained from integers, and in general "a set is something obtainable from the integers . . . by iterated application of the operation 'set of', not something obtained by [a conceptual definition which splits things into two categories]" ([5], pp. 262–263). I take it this entails that sets are not constituted by their definitions, and suggests that their constituents are their elements, etc. (For further comments on this see [9], p. 275; also compare [7], pp. 118–119. Maddy finds the first statement of this view in König.) The view of sets which I have in mind is to construe, for example, $\{x,y\}$ as $[t = x \text{ or } t = y]_t$, and more generally to take

$$A = \{a \mid a \in A\} = \left[\bigvee_{a\in A} t = a\right]_t.$$

While this identifies a set A with a certain (very idealized) concept defining it, notice that a set still comes out as a certain kind of combination, and it gets the constituents *almost* right: Perhaps the notion of equality really is needed to form sets, and the disjunction can be viewed as simply an indication that the kind of combination is what we earlier called "pure". It seems quite possible, however, that this kind of combination cannot be defined, but can only be axiomatized. That is, not only are individual sets not constituted by definitions, but neither is the notion of set itself. (The explanation of a set x as something obtainable by iterated application of set formation bears this out, since "obtainable" means nothing else than the existence of an ordinal α, i.e., a set, which gives x in α iterations.) In particular, there are difficulties in defining sets as the extensions of certain properties. For example, if we try to define sets as extensions of determinate properties, it seems that complements of sets should be sets. If we restrict to hereditarily determinate properties, it is not clear why the class of sets is not a set. The case is perhaps not hopeless; a possibility in the framework to be considered here is to read determinate as "significant for every argument". Another possibility, suggested by the above discussion, is to define sets as the extensions of properties which coincide in extension with the class of their constituents. In any case we do not pursue these ideas here, but take sets as basic. We observe that as long as C is a set, this automatically gives us any equivalence classes, including equivalence classes of propositions, etc., which we may need (by the Scott method: the objects of minimal rank equivalent to x).

We are now in a position to describe our axiomatic framework.

First, we describe the set theoretic scaffold. This is quite easy to do: we adopt the usual axiom schemas of ZFC, allowing arbitrary formulas from our language in all schemas. That is, we take as axioms all properly formed universal closures of such schemas. We may add any known axioms of infinity also, but for definiteness we stop with ZFC. (If urelements are desired, then we use a suitable formulation of set theory with urelements. If the urelements do not form a set, then the axioms should include a principle of dependent choice for transfinitely many choices.)

Next we describe the special axioms for C. These say that C is the hereditarily finite sets over C_b, the basic elements x of C (x such that x is disjoint from C).

C1 $\quad \forall \overset{C}{x}(\exists \overset{C}{t}(t \in x) \to \forall t(t \in x \to C(t)))$

C2 $\quad \forall \overset{C}{x}(\forall \overset{C}{t}(t \notin x) \to \theta(x)) \;\&\; \forall \overset{C}{x}\forall \overset{C}{y}(\theta(x) \;\&\; \theta(y) \to \theta(x \cup \{y\})) \to \forall \overset{C}{x}\theta(x)$

C3 $\quad \forall \overset{C}{y}\forall \overset{C}{x}(\forall t(t \in x \to C(t)) \to C(x \cup \{y\})$

C4 $\quad \forall t(t \notin z) \to C(z)$.

We shall suppose that C is a set:

C5 $\quad \exists z \forall t(t \in z \leftrightarrow C(t))$.

This completes the axioms for C. If we work in set theory with urelements, some obvious modifications are required: C4 should express $C(\varnothing)$, C5 should say z is a set. In addition, we would need individual constants c_i, $i \in \{\neg, \&, \forall, =,$ Set, $B, \in, C, T, S\}$ for the concepts of our theory, and axioms $C(c_i)$.

Now we describe the adjoining of B. If θ is a formula *whose free variables are all intensional* (recall that these range over C) then $B\theta$ is a formula. Ignoring terms, it is now easy to describe the logical framework: We take as axioms the appropriate closures of schemas of ordinary classical logic, together with some special schemas for B. The only rule is modus ponens. Specifically, the classical schemas are

L1 $\quad$ *truth functional tautologies*

L2 $\quad \forall x(\phi \to \theta) \to (\forall x\phi \to \forall x\phi)$

L3 $\quad \phi \to \forall x\phi$, *$x$ not free in ϕ*

L4 $\quad \forall y(\forall x\theta \to \theta(x/y))$, *$x$ free for y in θ, and y is intensional if x is*

L5 $\quad \forall x(x = x)$

L6 $\quad \forall x\forall y(x = y \to (\theta(u/x) \to \theta(u/y)))$.

Here we note that in L4, if x is an intensional variable, then we require that y must be also. By the appropriate closures we mean sentences formed by prefixing universal quantifiers and (optionally) B in any order. Of course, if we put a nonintensional variable free in the scope of B, we do not get a sentence.

The special schemas for B are

L7 $\quad B(\theta \to \phi) \to (B\theta \to B\phi)$

L8 $\quad B\theta \to \theta$

L9 $\quad B\theta \to BB\theta$

L10 $\quad B\forall x\theta \to \forall xB\theta.$

Notice that in general L10 gives a sentence only when x is an intensional variable. (In adopting L10, we thereby assume $\forall xB\exists y(x = y)$, $\forall xB(x = x)$, etc.)

The remainder of the axioms for B assert the computability of elementary combinatorial functions. In stating these, we indicate that x is intensional by writing $\forall \overset{C}{x}$. The first two axioms just give the range of these variables. (It may be helpful to amplify on what is meant by computable here. If $\delta(x,y)$ is a formula which defines y as a function of x, we call it computable if $\forall \overset{C}{x} \exists yB\delta(x,y)$. For example, the formulas for the identity function and the constant function with value z are $x = y$ and $z = y$ respectively. That these are computable follows already from our logical schemas.)

B0 $\quad \forall \overset{C}{y} BC(y)$

B1 $\quad \forall x(C(x) \to \exists \overset{C}{y}(x = y)).$

The next says that the empty set is identifiable.

B2 $\quad \forall t(t \notin z) \to B\forall t(t \notin z)$

B3 $\quad \forall \overset{C}{x} \forall \overset{C}{y}(x \neq y \to Bx \neq y).$

The corresponding statement for $x = y$ is already logically valid.

B4 $\quad \forall \overset{C}{x} \forall \overset{C}{y}(x \in y \to Bx \in y).$

Note that if $y \in C_b$, and $x \in y$, this says nothing.

B5 $\quad \forall \overset{C}{y}(\exists \overset{C}{t}(t \in y) \;\&\; \forall \overset{C}{t}(t \in y \to B\theta) \to B\forall t(t \in y \to \theta)).$

This axiom asserts the surveyability of finite sets. We do not assume that $C_b(x) \to BC_b(x)$. Of course, if $z = \{x_1, \ldots, x_n\}$ happens to be $\{x | C_b(x)\}$, then $x \in z \to Bx \in z$; but of course we may not have $BC_b(x)$. There need be no proof that $x \in z$ coincides in extension with $C_b(x)$.

In introducing the special axioms for truth and significance, it will be much more perspicuous to use terms. Since terms introduce special complications in the presence of B, we begin by discussing these.

So far we have avoided any essential use of terms. In classical logic it is well known that if we can prove $\forall x\exists !y\delta(x,y)$, then we can introduce a function symbol $f(x)$ and "defining axiom" $\forall x,y(f(x) = y \leftrightarrow \delta(x,y))$, and we obtain a conservative extension. The logical axioms are the same as before, except that L4 is formulated with y replaced by an arbitrary term t:

$$\forall x\theta \to \theta(x/t)$$

with the usual provisos that x is free for t in θ, and if x is a restricted variable, say of the same sort as a new variable u, then of course we only have

L4a $\exists u(t = u)$ & $\forall x\theta \to \theta(x/t)$.

In our context, given $B\forall x\exists!y\delta(xy)$, we wish to introduce f and add

$$B\forall xy(f(x) = y \leftrightarrow \delta(x,y)),$$

and obtain a conservative extension. We can indeed adopt L4a (if x is unrestricted we can omit $\exists u(t = u)$ or regard it as itself a logical axiom), but only provided that x does not occur free in the scope of B in θ. In the latter case, we need

L4b $\exists uB(t = u)$ & $\forall x\theta \to \theta(x/t)$,

which means that we can apply universal instantiation to a term only if its value is intuitively computable. If this restriction were ignored, we would not get a conservative extension when adding terms. (In fact, this would fail spectacularly, as all functions would become computable, and hence provability would coincide with truth.)

We have dealt with the problem of free variables within the scope of B by introducing special variables of restricted range for this role. There is an alternative which is perhaps more elegant. It consists in defining $C(x)$ by the formula $B\exists y(x = y)$. To carry this out we can adopt a form for L4 suggested by L4a:

L4′ $\exists u(y = u)$ & $\forall x\theta \to \theta(x/y)$

(here θ may involve B, but y must be a variable, not any term), together with the axioms

L10′ $\forall x\exists y(x = y)$

(x,y distinct variables). From there we obtain L4 and a form of L10, namely

$$B\forall x\theta \to \forall xB(\exists y(x = y) \to \theta).$$

This gives

$$B\forall x\theta \to \overset{C}{\forall x}B\theta,$$

and one can show (using $B\forall x(x = x)$) that

$$\overset{C}{\forall x}\exists yB(x = y),$$

so that the identity function is computable.

We are now ready to turn to the axioms for truth and significance. We state these now: the special quasi-syntactic notations should be fairly perspicuous after our earlier discussion; they are explained in detail immediately following the axioms.

We first give the axioms for significance. As usual, we intend the appropriate closures of the indicated formula. For the classical atomic formulas, these say that the atomic formula is significant for all arguments (for example, that $\forall x \forall y S[x = y]$):

S1 $S[\alpha]$

where α is any of $x = y$, $x \in y$, $C(x)$. For T and S themselves, the significance is conditional:

S2 $S[T(x)] \leftrightarrow S(x)$
$S[S(x)] \leftrightarrow S(x)$

S3 $S(\dot{\neg} x] \leftrightarrow S(x)$.

For implication, we could require $S(x \dot{\rightarrow} y) \leftrightarrow S(x)$ & $S(y)$. However, we adopt the more liberal policy suggested by Kripke:

S4 $S(x \dot{\rightarrow} y) \leftrightarrow (S(x)$ & $S(y))$ or $(S(x)$ & $\neg T(x))$ or $T(y)$.

Similarly with universal quantification:

S5 $S(\dot{\forall} ux) \leftrightarrow \forall y S(A(y,u,x))$
or $\exists y(S(A(y,y,x))$ & $\neg T(A(y,u,x)))$.

Here $A(y,u,x)$ is the result of applying the object y to the concept x at the place u.

Finally, for B we assume

S6 $S(\dot{B}x) \leftrightarrow S(x)$ & $C(x)$.

The schemas for truth are simply the usual truth conditions, conditional on the significance of the relevant formula. The axioms are suitable closures of these.

T0 $T(x) \rightarrow S(x)$
T1 $S[\alpha] \rightarrow (T[\alpha] \leftrightarrow \alpha)$, α atomic
T2 $S(\dot{\neg} x) \rightarrow (T(\dot{\neg} x) \leftrightarrow \neg Tx)$
T3 $S(x \dot{\rightarrow} y) \rightarrow (T(x \dot{\rightarrow} y) \leftrightarrow (Tx \rightarrow Ty))$
T4 $S(\dot{\forall} ux) \rightarrow (T(\dot{\forall} ux) \leftrightarrow \forall y TA(y,u,x))$
T5 $S(\dot{B}x) \rightarrow (T(\dot{B}x) \leftrightarrow BTx)$.

Note that in T5, since x occurs within the scope of B, x must be an intensional variable.

Let us represent linguistic symbols by triples $\langle t,m,x \rangle$ where t indicates the type of symbol (sentential connective, predicate symbol, variable, quantifier, or function symbol; 0, 1, 2, 3, 4 respectively), m is the number of places (or the sort of variable), and x is an index for the particular concept to be symbolized. In case r is for example a binary relation, we may suppose that $R = \langle 2,2,r \rangle$ is the predicate symbol for r; similarly that $\langle 3,1,\forall \rangle$ is the universal quantifier, etc. If we do not have the concepts r and $\forall$ themselves as objects of our theory, but only indices $i_r, i_\forall$ for them, these of course become $\langle 2,2,i_r \rangle$, $\langle 3,1,i_\forall \rangle$.

We may now define the operations $\dot{\neg}$, etc. We set

$$(\dot{\neg} x) = \langle\langle 0,1,i_\neg \rangle, x \rangle$$

$$(x \dot{\rightarrow} y) = \langle\langle 0,2,i_\rightarrow \rangle, x, y \rangle$$

$$(\dot{\forall} ux) = \langle\langle 3,1,i \rangle, u, x \rangle$$

$$(\dot{B}x) = \langle\langle 0,1,i_B \rangle, x \rangle.$$

We also put

$$c_x = \langle 4,0,\langle 0,x\rangle\rangle;$$

this provides us with canonical names as will be convenient. This gives names for arbitrary x, but notice that only when x is in C do we have c_x in C; this has the effect of allowing so-called singular propositions while preventing undesirable features which usually accompany their introduction. We can define substitution of terms for free variables in one of the usual ways now, and set

$$\begin{aligned} A(y,u,x) &= \text{the result of substituting } c_y \text{ for} \\ &\quad\ \text{all the free occurrances of } u \text{ in } x \\ A(y_1,u_1;y_2,u_2;x) &\text{ is } A(y_1,u,A(y_2,u_2,x)). \end{aligned}$$

To handle the atomic formulas, we put

$$\begin{aligned} (u \dot{=} v) &= \langle\langle 1,2,i_=\rangle,u,v\rangle \\ (u \dot{\in} v) &= \langle\langle 1,2,i_\in\rangle,u,v\rangle \\ (\dot{C}u) &= \langle\langle 1,1,i_C\rangle,u\rangle \\ (\dot{T}u) &= \langle\langle 1,1,i_T\rangle,u\rangle \\ (\dot{S}u &= \langle\langle 1,1,i_S\rangle,u\rangle, \end{aligned}$$

and for example,

$$\begin{aligned} [Cx] &= \langle\langle 1,1,i_C\rangle,c_x\rangle \\ &= A(c_x,u,\dot{C}u) \quad (u \text{ any variable}) \\ [x=y] &= \langle\langle 1,2,i_=\rangle,c_x,c_y\rangle \\ &= A(c_x,u;c_y,v;(\dot{u}=v)) \quad (u,v \text{ any distinct variables}). \end{aligned}$$

The bracket notation can be defined more generally, but this is all that is needed for understanding our axioms.

The T-S axioms may now be understood as asserting the existence of nine concepts in C (corresponding to $\neg$, $\rightarrow$, $\forall$, $=$, $\in$, C, T, S, B) of which the universal closures of the T-S axioms are provable. That is, replacing $i_\neg, i_\rightarrow \ldots$ by variables $u_0, u_1, \ldots$ this has the form

$$\exists u_0, u_1 \ldots B\forall x,y \ldots (\ldots (S(\dot{\neg}x) \leftrightarrow Sx) \ldots).$$

Alternatively we can expand our official language to include individual constants $c_\neg$ for $i_\neg$, etc., with axioms

$$\exists uB(u = c_\neg).$$

Notice that the choice of the nine indices $i_\neg$, etc. here is not as arbitrary as one might suppose. First of all, i_r must be in C; but where? To the extent that we are using the structure of C to represent conceptual combinations, and are ignorant of the exact conceptual structure of r (e.g., is it defined or basic?), we cannot arbitrarily pick an index without imposing or excluding structure on r unintentionally. Thus we have refrained from specifying a particular indexing.

This concludes the presentation of the axioms. This presentation is inelegant in the following way. We have used terms in our presentation of the axioms, yet have given axioms for T and S which give information only about formulas which do not involve terms (except for canonical names c_x). Even without B,

$$x = t \to (T[\theta(x)] \to T[\theta(t)])$$

is not (contrary to appearances) a logical validity; it and a similar principle for S must be explicitly assumed if terms are desired.

A few observations on the intended interpretation of theories of significance. For a more detailed discussion, see [10]. If Γ is an ordinary theory, we take $\Gamma \vdash \sigma$ as a warrant for the truth (or at least assertability) of σ. Here, however, $\Gamma \vdash T[\sigma]$ provides such a warrant but $\Gamma \vdash \sigma$ does not unless $\Gamma \vdash S[\sigma]$. Because of the liar type paradoxes there will in fact be sentences such that $\Gamma \vdash \sigma$ but σ does not satisfy the sufficient condition we have for significance, namely $\Gamma \vdash S[\sigma]$. Some σ will thus have a purely formal meaning; this is because T and S are partial predicates only. Such sentences may however contain information about genuinely significant sentences; this can be expressed by significant statements such as

$$\forall n(\Gamma \vdash \sigma \to T[\theta_{n,\sigma}]),$$

but not (in general) by asserting $T[\sigma]$.

It should now be clear how to formulate the Post–Turing thesis in our framework. The properties we quantify over are just the formulas in C with one free variable, say u. Because we have imposed no restrictions on the relations which may be indexed, this gives a very general notion of property. (In particular it is more general than those expressible by first-order sentences of arithmetic.) Notice that if P is such a property we express $P(x)$ by $T(A(x,u,P))$. Since u is determined by P here, we may write $x\eta P$. Thus the strong form of the Post–Turing thesis becomes

$$\forall P(\forall \overset{\omega}{n}(n\eta P \to Bn\eta P) \to \exists \overset{\omega}{e} \forall \overset{\omega}{n}(n\eta P \leftrightarrow n \in W_e)).$$

While this is more adequate than arithmetical versions, I do not wish to make exaggerated claims about its adequacy. While it in some way avoids ramified types in favor of arbitrary properties, it is formulated in a theory of significance which is inadequate to assign it significance.

I conclude with a brief discussion of ordinal notations. The point is not to develop the theory, but to give an example of an interesting notion which involves in an essential way both constructive and nonconstructive elements. One of the main points about the theory I have presented in this paper is that it incorporates ordinary set theory as understood classically, rather than attempting to reinterpret the formal language constructively, either by replacing the classical theory with something more amenable to constructive ideas, or by some interpretation of classical logic in intuitionistic logic. But the theory also incorporates an epistemic element, which gives rise to the possibility of a significantly different point of view on ordinal notations: namely, they may be viewed as notations (a notion with an epistemic or constructivistic content) for ordinals (conceived in a purely classical nonconstructive way). It seems important to bring out this possibility, because (a) the classical view ordinarily replaces essentially epistemic notions with related extensional notions (such as that of a recursive set), while (b) on constructive views any reference to a set ordinarily presupposes having a conception of the set.

I have emphasized that the notion of an ordinal notation involves in an essential way both constructive and nonconstructive notions. Let me try to explain the point by analogy with the notion of logical validity. This notion can be viewed in both finitary and infinitary ways, or rather, there are really two different notions depending on which way we choose. The analogy I want to suggest is

$$\text{finitary} : \text{infinitary} :: \text{constructive} : \text{classical} + \text{epistemic}.$$

The completeness theorem of first-order logic relates the two notions. This theorem involves in an essential way both finitary and infinitary notions: the notion of logical validity, conceived as truth in all structures, is essentially infinitary, while the notion of a formally valid argument is finitary. The completeness theorem shows that the two are extensionally equivalent; thus, in a certain sense, after the theorem is proved one can always avoid the infinitary notion and use the finitary notion. If one does not allow infinitary notions in the first place, any special importance of the notion will have to arise in some other way, but if one accepts infinitary notions then the completeness theorem gives it a particular interest. Now, in the case of the notion of ordinal notation, the most natural understanding of the notion involves the notion of an ordinal, which is quite nonconstructive, and the notion of a notation, which in its most natural interpretation involves an essentially constructive element. We can, of course, make constructive analogues of the notion of ordinal, and thus conceive of ordinal notations in a purely constructive way. It is possible that we could prove something analogous to the completeness theorem; it would say that if we accept the nonconstructive notions, then there is a constructive notion that is extensionally equivalent to the mixed notion of a notation for a nonconstructively conceived ordinal. (Another possibility is that a pure classical notion could be extensionally equivalent to the mixed notion.) Thus it is of interest as an application of the epistemic set theory here proposed, which mixes classical nonconstructive ideas with the element of constructivity in the notion of provability, to give a definition of a system of notations for an ordinal.

There are various conditions one may impose on an adequate set of notations. In the present framework these can be expressed directly in terms of intensional relations and functions and their relation to ordinals, without assumptions about Turing machine representability. Note that the latter would be required in any classical treatment, and that in constructivistic treatments, it does not make sense to refer to ordinals apart from a conception of them, so that there is no possibility of relating abstractly given ordinals to our intensional relations and functions in the way intended here. This point will be made clearer as we proceed. Notice that avoiding assumptions about Turing machine representability is of interest even if we assume the Post–Turing thesis: there may be concepts which are intuitively decidable (or weakly decidable) and hence recursive (or r.e.), say by some machine e, but of e one cannot show some property that may be shown working from the original concept.

We note first that we can express "the concept P defines the ordinal α" by

$$\forall\nu(\nu \in \alpha \leftrightarrow \nu\eta P).$$

(Here we use ν, α, etc. for ordinals. Let us write $D(P,\alpha)$ for this. This gives us the notion of a definable ordinal: $T[\exists PD(P,\alpha)]$ and of a concept defining an ordinal: $T[\exists\alpha D(P,\alpha)]$. The potential ordinal notations will thus be P for which $BT[\exists\alpha D(P,\alpha)]$.

We can now clarify the remark made above that this makes a use of classical notions not available to the constructivist. Specifically, the notion

$$BT[\exists\alpha D(P,\alpha)]$$

uses the classical quantifier $\exists\alpha$. Although α cannot go free within the scope of B (so that $BD(P,\alpha)$, for example, is not an intelligible notion), it can be quantified out to leave P, which *can* go in the scope of B. Thus we may speak of knowing that P defines an ordinal without having presupposed direct access to the ordinal. In effect this gives us the notion of a concept P of an arbitrary ordinal as opposed to a constructively given ordinal. We are now in a position to explain what a system of ordinal notations is.

Let us suppose that Q is a property of the sort just discussed, that is, $BT[\exists\alpha D(Q,\alpha)]$. Let us also suppose that N is a property such that $B\forall^C xS[x\eta N]$. We call N a system of notations for Q in case we have B of the following:

1. $(\forall P\colon P\eta N)BT[\exists\alpha D(P,\alpha)]$
2. $(\forall P\colon P\eta N)BP\eta N$
3. $(\forall P\colon P\eta N)(\forall\alpha\colon D(P,\alpha))(\alpha\eta Q)$
4. $(\forall\lambda\colon D(Q,\lambda))(\forall\alpha\colon \alpha < \lambda)\exists!P(P\eta N \;\&\; D(P,\alpha)$
5. $(\forall\lambda\colon D(Q,\lambda))(\forall\alpha,\beta\colon \alpha,\beta < \lambda)(\forall P_1,P_2\colon D(P_1,\alpha),D(P_2,\alpha))$
   $\alpha < \beta \rightarrow B\forall\ \alpha,\beta(D(P_1,\alpha)\ \&\ D(P_2\beta) \rightarrow \alpha < \beta)$
   $\lim(\alpha) \rightarrow B\forall\alpha(D(P_1,\alpha) \rightarrow \lim(\alpha))$
   $\beta = \alpha + 1 \rightarrow B\forall\ \alpha,\beta(D(P_1,\alpha)\ \&\ D(P_2,\beta) \rightarrow \beta = \alpha + 1)$
   $\alpha = 0 \rightarrow B\forall\alpha(D(P_1\alpha) \rightarrow \alpha = 0)$.

One could in a similar way impose stronger conditions; for example, one could require that if α is a limit ordinal, then for some $\beta < \alpha$ we have a cofinal sequence $h(\nu)$, $\nu < \beta$ converging to α. Such conditions must be stated relative to N of course; the requirement would be for a function $f\colon N \rightarrow N$ (which would give the notation for β given that for α) and a function g in $\prod\limits_{P\eta N}^{(fP)} N$, where (fP) is the notation in N for ordinals below the ordinal β defined by fP. Both f and g are to be intensional functions, of course; g is to give for each P_2 in (fP) defining ν, a notation in N for $h(\nu)$.

REFERENCES

[1] Bealer, G., *Quality and Concept*, Clarendon Press, Oxford, 1982.

[2] Bernays, P., "On platonism in mathematics," *L'enseignement Mathématique*, 1st ser. vol. 34 (1935), pp. 52–69; translated by C. D. Parsons and reprinted as pp.

274–286 in *Philosophy of Mathematics*, ed., P. Benacerraf and H. Putnam, Prentice-Hall, Englewood Cliffs, New Jersey, 1964. Page references in this paper are to the reprinted version.

[3] Gödel, K., "Eine Interpretation des intuitionistischen Aussagenkalküs," *Ergebnisse eines Mathematischen Kolloquiums*, vol. 4 (1933), pp. 39–40.

[4] Gödel, K., "Russell's mathematical logic," pp. 125–153 in *The Philosophy of Bertrand Russell*, ed., P. Schilpp, Tudor Publishing Company, New York, 1944; reprinted as pp. 211–232 in *Philosophy of Mathematics*, ed., P. Benacerraf and H. Putnam, Prentice-Hall, Englewood Cliffs, New Jersey, 1964. Page references in this paper are to the reprinted version.

[5] Gödel, K., "What is Cantor's continuum problem?" *American Mathematical Monthly*, vol. 54 (1947), pp. 515–525; revised and reprinted as pp. 258–273 in *Philosophy of Mathematics*, ed., P. Benacerraf and H. Putnam, Prentice-Hall, Englewood Cliffs, New Jersey, 1964. Page references in this paper are to the reprinted version.

[6] Kripke, S., "Outline of a theory of truth," *Journal of Philosophy*, vol. 72 (1975), pp. 690–716.

[7] Maddy, P., "Proper classes," *The Journal of Symbolic Logic*, vol. 48 (1983), pp. 113–139.

[8] Myhill, J., "Some remarks on the notion of proof," *Journal of Philosophy*, vol. 57 (1960), pp. 461–471.

[9] Reinhardt, W., "Satisfaction definitions and axioms of infinity in a theory of properties with necessity operator," pp. 267–303 in *Mathematical Logic in Latin America*, ed., A. I. Arruda, R. Chuaqui, and N. C. A. da Costa, North Holland, Amsterdam, 1980.

[10] Reinhardt, W., "Some remarks on extending and interpreting theories with a partial predicate for truth," *Journal of Philosophical Logic*, vol. 15 (1986), pp. 219–251.

Department of Mathematics
University of Colorado
Boulder, Colorado 80309