Gödel's interpretation of intuitionism

W. W. Tait*

Abstract

Gödel regarded the *Dialectica* interpretation as giving constructive content to intuitionism, which otherwise failed to meet reasonable conditions of constructivity. He founded his theory of primitive recursive functions, in which the interpretation is given, on the concept of computable function of finite type. I will 1) criticize this foundation, 2) propose a quite different one, and 3) note that essentially the latter foundation also underlies the Curry-Howard type theory and hence Heyting's intuitionistic conception of logic. Thus the Dialectica interpretation (in so far as its aim was to give constructive content to intuitionism) is superfluous.

In the set of notes [1938a] for an informal lecture, Gödel refers to a hierarchy of constructive or finitary systems, the lowest level of which he called *finitary number theory* and is, in fact, primitive recursive arithmetic (PRA). He mentions three constructive extensions of this system: the intuitionism of Brouwer and Heyting (the "modal-logical route"), concerning which he is quite critical, the use of transfinite induction, inspired by Gentzen's consistency proof for PA, and the use of functions of finite or even transfinite type over the type $\mathbb{N}$ of natural numbers. The latter extension seems to be mentioned here for the first time in connection with constructive extensions of finitary number theory.[1] He ranks it the most satisfactory of these extensions

[*] Versions of this paper were presented at a joint workshop on Hilbert at Carnegie-Mellon University and the University of Pittsburgh in June 1998, at a workshop on foundations of mathematics at UCLA in May 2005 and at the IMLA '05 workshop [Intuitionistic Modal Logic and Applications, of course] in Chicago. I received helpful comments on all of these occasions.

[1] As for the history of the idea of a hierarchy of types, I know only of these examples: The hierarchy of transfinite types in which the natural numbers constitute the lowest type

of PRA; but he gives no indication there about what can be accomplished with it. However, in lectures in 1941 at Princeton and at Yale, the lecture notes for the latter of which constitute [Gödel, *1941], he introduced what came to be called the *Dialectica* or *functional interpretation* of Heyting arithmetic HA and PA (regarded as a subsystem of HA). The details of this interpretation were published first in the *Dialectica* paper [Gödel, 1958], where he presents the interpretation as a consistency proof for HA relative to his theory $\mathbf{T}$ of primitive recursive functions of finite type. But the earlier lectures emphasize rather the failure of Heyting's explanation of the logical constants to give satisfactory constructive content to intuitionistic logic and the role that Gödel believed the interpretation plays in supplying it (at least in the context of arithmetic).

My aim here is not to discuss the technical details of the interpretation, nor its applications or extensions in proof theory. These are nicely presented in A. Troelstra's introductory note to [Gödel, 1958; Gödel, 1972] in [Gödel, 1990] and in the more recent paper [Avigad and Feferman, 1998]. Rather, I want to discuss the foundation of the theory $\mathbf{T}$ itself and its relation to the foundation of HA. I believe that Gödel misconstrued both of them. His misconception of the latter led him to want to replace the notion of proof in it by the notion of function of finite type over the type $\mathbb{N}$ of natural numbers. In fact, we will see that this notion of proof itself can be understood to have type structure and, so understood, is immune to Gödel's criticism of it and, indeed, constitutes only a slight generalization of the notion of function of finite type over $\mathbb{N}$. Here of course I am referring to the Curry-Howard theory of *propositions-as-types* [Howard, 1980]. Gödel's misconception of the former consists in the belief that a constructive theory of functions of finite type requires that the higher type variables range only over functions which are provably *computable* (*berechenbaren*). But he never spelled out a

and the successor type of A is the type $A \to 2$ of all two-valued functions on A is introduced by Cantor in [Cantor, 1891], as representing an unbounded sequence of cardinal numbers. (The limit types were most likely intended to be the unions of the previous types.) In his *Principles of Mathematics*, Russell introduced the finite types built up from the type of individuals by pasing from types A and B to $A \to 2$ and the type $A \wedge B$ of sll pairs of elements from A and B respectively. (Russell called them "couples with sense".) In [Hilbert, 1926], Hilbert introduced two hierarchies of height ω_1 with lowest type M, one starting with M being the domain of natural numbers and the other with it being the second number class. In both cases, the 'successor' of the type A is the type $A \to M$ of all M-valued functions on A and the limit types are of the form $C \to M$, where C is the union of the preceding types.

satisfactory account of this which avoides the very intuitionistic logic that he was attempting to reinterpret.

1. Let FT denote the set of formulas of propositional logic built up from atomic formulas by means of implication $\rightarrow$ and conjunction $\wedge$. $FT(\mathbb{N})$ denotes the set of such propositional formulas built up from the single atomic formula $\mathbb{N}$. We are interested in type-theoretic interpretations of FT and $FT(\mathbb{N})$. Namely, regard the atomic formulas as standing for arbitrary types of objects, the atomic types. $A \rightarrow B$ is the type of all functions from A to B, and $A \wedge B$ is the type of pairs whose first term is of type A and whose second term is of type B: i.e.

$$A \rightarrow B := {}^{A}B \qquad\qquad A \wedge B := A \times B.$$

Write

$$a{:}A$$

to mean that a is of type A. We write $A \wedge B \wedge C$ for $(A \wedge B) \wedge C$, etc. Gödel included only the types built up from $\mathbb{N}$ by means of the operation of passing from types $A_1, \cdots, A_n, B$ to $A_1 \wedge \cdots \wedge A_n \rightarrow B$. Via Schönfinkel's canonical correspondence $f(a, b) = gab$ between functions f of type $A \wedge B \rightarrow C$ and functions g of type $A \rightarrow (B \rightarrow C)$, the type-forming operation $\wedge$ is dispensable in Gödel's system. Another logical equivalence, namely between $A \rightarrow B \wedge C$ and $(A \rightarrow B) \wedge (A \rightarrow C)$, signals the correspondence between objects f of the former type and (g, h) of the latter defined by $fa = (ga, ha)$. Thus, every finite type involving $\wedge$ is represented by a type of the form $A_0 \wedge \cdots \wedge A_n$ $(n \geq 0)$ where the A_i do not contain $\wedge$. So the net effect of including $\wedge$ is to allow us to represent n-tuples of objects of finite type by a single object. But it will be useful as well as instructive to include it. The extension of the finite types to include pairing is trivially conservative over **T**, and so I will not bother to distinguish it from **T**.

$\mathcal{P}$ denotes propositional logic restricted to the connectives $\rightarrow$ and $\wedge$. FT is the set of *formulas* of $\mathcal{P}$ and its *terms* are as follows:

- Each variable of type A is a term of type A.
- If $s{:}A$ and $t{:}B$, then $(s, t){:}A \wedge B$. (*Pairing*)
- If $t{:}A \wedge B$, then $tL{:}A$ and $tR{:}B$. (*Left and Right Projection*)
- If $s{:}A \rightarrow B$ and $t{:}A$, then $(s)t{:}B$. (*Evaluation*)

- If v is a variable of type A and $t(v) : B$, then $\lambda x : A.\, t(x) : A \to B$. ($\lambda-$*Abstraction*)

We may think of a pair as a function defined on $\{L, R\}$, whose first term is the value for L and whose second term is the value for R; but we do not regard$\{L, R\}$ as a type.

Projection ($\wedge$-*elimination*) and evaluation ($\to$-*elimination*) are primitive operations, implicit in the notion of a function. Pairing ($\wedge$-*introduction*) is defined by

$$(s, t)L = s \qquad\qquad (s, t)R = t$$

λ-abstraction ($\to$-*introduction*) is defined by $\lambda-$*conversion*:

$$(\lambda x{:}A.\, t(x))s = t(s).$$

I will write st for $(s)t$, rst for $(rs)t$, etc., when no confusion results.

It is immediate that a formula A of $\mathcal{P}$ is a theorem of intuitionistic propositional logic iff there is a closed term of $\mathcal{P}$ of type A.

The terms of **T** are obtained by the above operations, with the atomic types restricted to $\mathbb{N}$, together with

- $0{:}\mathbb{N}$
- If $t{:}\mathbb{N}$, then $t'{:}\mathbb{N}$
- If $r{:}\mathbb{N}$, $s{:}A$ and $t{:}A \to A$, then $I_A(r, s, t){:}A$,

where, dropping the subscript, $I(r, s, t)$, the rth *iterate* of t applied to s, is defined by the equations

$$I(0, s, t) = s \qquad\qquad I(r', s, t) = tI(r, s, t).$$

The formulas of **T** are built up from equations between terms of the same type by means of the propositional connectives. Its axioms and rules of inference are those of classical propositional logic, the defining equations above for the projections, λ-abstraction and iteration, the rule of inference by mathematical induction, the axioms $(tL, tR) = t$ and the rule of inference from $\phi \to sv = tv$ to $\phi \to s = t$, when s and t are of type $B \to C$ and the variable v of type B does not occur in ϕ, s or t. (Gödel does not include this inference, which is equivalent to adding as axioms all instances of η-conversion: $\lambda x{:}A.\, tx = t$.)

In place of iteration Gödel took as primitive the more general form of primitive recursion, introducing, for $r : \mathbb{N}, s : A$, and $t : \mathbb{N} \rightarrow (A \rightarrow A)$, $R_A(r, s, t)$ of type A, where

$$R(0, s, t) = s \qquad\qquad R(r', s, t) = trR(r, s, t)$$

But in the presence of pairing, this more general form is derivable using $I = I_{\mathbb{N} \wedge A}$. Write $f = \lambda x{:}\mathbb{N}.\, R(x, s, t)$. Let $h{:}\mathbb{N} \wedge A \rightarrow \mathbb{N} \wedge A$ be defined by $h(v, u) = (v', tvu)$, where v is a numerical variable and u is a variable of type A. Define $\bar{f} = \lambda x{:}\mathbb{N}.\, I(x, (0, a), h)$ and then $f = \lambda x{:}\mathbb{N}.\, \bar{f}(x)R$. Notice that $\lambda x{:}\mathbb{N}.\, (\bar{f}x)L = \lambda x{:}\mathbb{N}.\, I(x, 0,')$ = the identity function on $\mathbb{N}$. So $f0 = s$ and $fv' = t(\bar{f}v)L(fv) = tv(fv)$.

$0 = 0' \rightarrow s = t$ is derivable for terms s and t of like type A. This is shown by induction on A. When $A = \mathbb{N}$, it follows by mathematical induction. The case $A = B \wedge C$ is trivial. If $A = B \rightarrow C$, choose a variable v of type B not to be in A. Then we have $0 = 0' \rightarrow sv = tv$ and so $0 = 0' \rightarrow s = t$. It follows that $0 = 0' \rightarrow \phi$ is derivable in $\mathbf{T}$ for all formulas ϕ. This means that negation can be defined

$$\neg\phi := \phi \rightarrow 0 = 0'.$$

Since disjunction can be contextually defined in HA by

$$A \vee B := \exists x[(x = 0 \rightarrow A) \wedge (x \neq 0 \rightarrow B)]$$

it is not needed in $\mathbf{T}$ for the functional interpretation. (See the next section.) So we may assume that the only propositional connectives are conjunction $\wedge$ and implication $\rightarrow$.

The usual axioms of 0 and successor are derivable in $\mathbf{T}$: We obtain $0 \neq t'$, i.e., $0 = t' \rightarrow 0 = 0'$, for a numerical term t by applying the sign function sgn, defined by $sgn0 = 0$ and $sgnv' = 0'$, to both sides of $0 = t'$. We obtain $s' = t' \rightarrow s = t$ by applying the predecessor function $pred$, defined by $pred0 = 0$ and $predv' = v$, to both sides of $s' = t'$.[2]

2. The functional interpretation consists in associating with each formula $\phi(\vec{v})$ of HA, with just the free numerical variables $\vec{v}$, a formula

$$(1) \qquad \exists x \forall y F(x, y, \vec{v})$$

[2] I don't understand the argument in [Gödel, 1972, fn. i] for the second of these axioms using onl;y *pred*; but the second axiom is derivable using just *pred* and mathematical induction.

where x and y each range over some finite type, and F is quantifier-free. With each proof in HA of the formula ϕ, Gödel shows how to associate a closed term f and a proof in **T** of

$$F(f\vec{v}, y, \vec{v})$$

When $\phi(v)$ is atomic, then $F(x, y, \vec{v}) = \phi(\vec{v})$ (where $\exists x$ and $\forall y$ are vacuous quantifiers over numbers), and so a proof of $0 = 1$ in HA would translate into one in **T**. Thus we have a proof (in PRA) of the consistency of HA, and so of PA, relative to **T**.

3. Gödel's foundation of the theory **T** is the notion of a *computable* object of finite type (in $FT(\mathbb{N})$). Computation concerns symbols: Given a function f of type $\mathbb{N} \to \mathbb{N}$ and a number t, however these are presented, ft is a number; but the requirement that f be computable is that we be able to effectively determine the corresponding *numeral* from the numeral for t. Thus the computable objects of type $\mathbb{N}$ are the numerals. $A \wedge B$ is again simply the cartesian product of A and B. $A \to B$ is the type of the computable functions from A to B. A computable function of this type is defined in [Gödel, 1972] to be a well-defined mathematical procedure for which it is *constructively evident* that, applied to every computable object of type A, it yields a computable object of type B.

However, this foundation is *prima facie* problematic and Gödel was well aware of this. The difficulty of course arises in the case of $A \to B$. Notice that already in the case of a function f of type $\mathbb{N} \to \mathbb{N}$ there is a problem because of the notion of constructive evidence. Suppose for definiteness that the well-defined mathematical procedure in question is given by a Turing machine. What needs to be evident is that, for every numeral input, the machine halts after a finite number of steps. This condition is of the form $\forall m \exists n C(m, n)$, where $C(m, n)$ expresses that for the input numeral $\overline{m}$ the machine halts in n steps. But the constructive meaning of this is that there is a witness, i.e. a numerical function g such that $C(m, g(m))$ is valid. Our original candidate f for a computable function is primitive recursive in this g; but what about g? If it is the constructivist contention that, to produce a function is to produce a computable function, there seems to be a circle. One may well argue that there could be good constructive grounds for admitting the function g other than having a proof that it is computable. I wholeheartedly agree with this; but then the same applies to f itself; and these grounds have nothing to do *prima facie* with f or g being computable.

Let me say straight off that there are two distinct ideas: one is construction and the other is computation. These have been confused in recent history, but really are distinct. 'Constructive' means that the only witnesses of existential propositions one admits are ones that can be constructed, where of course this implies some background rules of construction. From the construction of an object, a means of computing it (in cases in which this idea makes sense) may or may not be found. In the context of arithmetic and analysis, constructive does imply computable (and this may indeed be the motivation for some to proceed constructively); but this is a theorem*, it is not built into the notion of construction.*

The situation for Gödel's conception is even worse when we consider functions of higher type. He himself suggests in [Gödel, 1972, p. 275, footnote 6] that the definition of computable function of finite type may not be "sufficiently clear", but asserts that, nevertheless, there can be no doubt that the functions of finite type introduced in **T**, the primitive recursive functions, are computable. But this is at best misleading. If we accept Turing's analysis of a computation procedure at least (as indeed Gödel did), then we may regard functions of finite type A as given by (numerical codes for) certain Turing machines satisfying a certain condition $\mathcal{C}_A$, where

$$\mathcal{C}_{\mathbb{N}}(t) = t \text{ is a numeral}$$

$$\mathcal{C}_{A\to B}(f) = \forall x[\mathcal{C}_A(x) \to$$

$$\{f\} \text{ applied to } x \text{ computes to some } y \text{ such that } \mathcal{C}_B(y)].$$

$$\mathcal{C}_{A\wedge B}((a,b)) = \mathcal{C}_A(a) \wedge \mathcal{C}_B(b).$$

This seems perfectly clear. But for example in the case of $\mathbb{N} \to A$, there is a difficulty with our seeing that the primitive recursive functions are all computable. Consider the case of the iteration $fn = g^n a$, where a is a primitive recursive object of type A and g is a primitive recursive function of type $A \to A$. The proof that f is computable proceeds from the assumption that a and g are computable by showing that, for each n, fn is computable. The argument is by induction on the numerical variable v applied to the formula $\mathcal{C}_A(fv)$. But, depending on the propositional complexity of the type A, the logical complexity of the formula $\mathcal{C}_A(fv)$ can be arbitrarily high: the proof that every primitive recursive function of finite type is computable requires all of HA.In fact, given any finitely axiomatized subsystem of the axioms of HA, there is a a closed term t of $\mathbf{T}$ of type $\mathbb{N} \to \mathbb{N}$ such that $\mathcal{C}_{\mathbb{N}\to\mathbb{N}}(t)$ is not a theorem of the subsystem. Gödel is certainly right that we see immediately that f is computable, but that is because we immediately accept mathematical induction applied to the formulas $\mathcal{C}_A(fv)$. However that should not be satisfactory for Gödel, since the aim of the *Dialectica* interpretation was in part to give constructive content to such formulas. (See [Gödel, *1941].) He himself wrote in a continuation of footnote 6

> [I]f this replacement [of intuitionistic proof by computable function of finite type] is to have any epistemological significance, the concept of computable function used and the insight that these functions satisfy the axioms of $\mathbf{T}$ ... must not implicitly involve intuitionistic logic or the concept of proof used by Heyting.

He insists that his notion of computable function of type A does not encounter this difficulty; but it is clear that he himself was not entirely convinced: he continued to worry about the matter in his correspondence with Bernays and, in the end, did not release the final version [Gödel, 1972] of his interpretation for publication.

4. An alternative foundation for $\mathbf{T}$ is to introduce the so-called 'minimal term model', in which the objects of type A are just the closed terms of $\mathbf{T}$ of type A and the meaning of equations between such terms is defined as follows: First, define the notion $s \succeq t$ of a term s *reducing* to another one t as follows:

- $s \succeq s$.
- If $r \succeq s$ and $s \succeq t$, then $r \succeq t$.
- If $s = t$ is a substitution instance of a defining axiom, then $s \succeq t$.
- if $s(v) \succeq t(v)$, then $\lambda x{:}A.\, s(x) \succeq \lambda x{:}A.\, t(x)$.
- If $r \succeq s$ and $t \succeq u$, then $rt \succeq su$ and $(r, t) \succeq (s, u)$.
- If $r \succeq s$, the $r' \succeq s'$.
- If $r \succeq r_0, s \succeq s_0$ and $t \succeq t_o$, then $I_A(r, s, t) \succeq I_A(r_0, s_0, t_0)$.

The relation $s = t$ of definitional equality between terms of type A by induction on A as follows:

- If $A = \mathbb{N}$, then $s = t$ iff they reduce to a common term.
- If $A = B \wedge C$, then $s = t$ iff $sL = tL$ and $sR = tR$.
- If $A = B \to C$, then $s = t$ iff $sv = tv$, where v is a variable of type B not occurring in s or t.[3]

That this term model satisfies the axioms of Gödel's theory **T** is shown in [Tait, 1963; Tait, 1967] to be derivable in HA. Namely, apply the predicates $\mathcal{C}_A$ to the (not necessarily closed) terms of **T** of type A, let $\mathcal{C}_{\mathbb{N}}(r)$ mean that r reduces to an irreducible term, and let and let ''s computes to t" mean that s reduces to t. From $\mathcal{C}_A(r)$ it follows that r reduces to an irreducible term (necessarily unique by the Church-Rosser Theorem) and it easily follows from this that the relation $s = t$ is decidable.

But with this model, we give up founding the theory **T** on the notion of a function of finite type entirely. We are dealing instead with terms (i,e, first-order objects, codable by numbers) governed by a certain reducibility relation.

The decidability of equations is essential for Gödel's purpose, since the propositional logic in **T** is classical and so the validity of excluded middle must be justified for arbitrary equations. As a matter of fact, classical logic is needed, not for the *Dialectica* interpretation, but rather because intuitionistic propositional logic, in particular the logic of implication, is a target of Gödel's criticism that Heyting's account of the meaning of the logical constants fails to give them a satisfactory constructive sense. (We shall discuss this below.) Once we establish that excluded middle is valid for equations, however, then the propositional logic can be interpreted in terms of truth functions.[4] But as Gödel himself seems to have anticipated in notes for the lectures given in Princeton in 1941 (see Troelstra's introductory note [Gödel, 1995, p. 188] to [Gödel, *1941]), it requires the resources of all of HA to prove that, for each closed term $s{:}A = (\mathbb{N} \to \mathbb{N})$ of **T**, $\mathcal{C}_A(t)$.

[3] In the past I have defined $s = t$ simply to mean that s and t reduce to a common term; but that is unsatisfactory, since, for example, $\lambda x : A.\, fx$ ought to be definitionally equal to f and (tL, tR) ought to be definitionally equal to t.

[4] The elimination of the logical constants in terms of truth functions cannot be carried out in **T** itself, in the form Gödel presents it, since definitional equality of higher types is not expressible in **T**. However, as we note immediately below, [Spector, 1962] shows that it suffices to admit equality only between numerical terms and, with this restriction, the logical constants can be eliminated in favor of truth functions.

Moreover, *the difficulty exists even for equations between numerical terms* in this minimal term model.[5] The axiom $tv = 0 \vee tv \neq 0$, where s and t are closed terms of type $\mathbb{N} \rightarrow \mathbb{N}$, must be founded on a proof that $t\overline{n}$ reduces to a numeral for every n and, depending on t, this will require arbitrarily large subsystems of HA. So, in particular, Spector's [1962] observation that Gödel's interpretation of HA in $\mathbf{T}$ does not require equations among terms of types $\neq \mathbb{N}$ does not help here.

Notice that there are really two problems with Gödel's foundation for $\mathbf{T}$, one local and one global. The local problem is that introducing a given function f of higher type A seems to require that we prove a logically complex condition $\mathcal{C}_A(f)$. The global problem is the one just mentioned that the proof that the objects introduced in $\mathbf{T}$ really are computable, i.e. that the closed numerical terms reduce to numerals, requires all of HA.[6]

In footnote h, attached to the footnote 6 (quoted in part above) in the 1972 paper, Gödel sketches a solution, at least in connection with the minimal term model, to the local problem. The condition of being 'constructively evident' in the definition of computable function is to be replaced by the condition of being *reductively provable.*

> [The concept of reductive proof] roughly speaking, is defined by the fact that, up to certain trivial supplementations, the chain of definitions of the concept occurring in the theorem together with certain axioms about the primitive terms forms by itself a proof, i.e., an unbroken chain of immediate evidences. In special applications (as, e.g., in our case) this concept of proof can be

[5] Of course, we can prove by complete induction that $u = v \vee u \neq v$ for numerical variables u and v, and this is often expressed by saying that numerical equality is decidable. But what is proved by induction, as we normally understand it, is a statement about *numbers*, as represented by numerals in the sequence $0, 0', 0'', \ldots$, not arbitrary numerical terms.

[6] Of course, in place of HA we could use $PRA$with quantifier-free transfinite induction up to ordinals $< \epsilon_0$. But Gödel wished to provide an alternative to Gentzen's consistency proof for HA. Nevertheless, he remained interested in the problem of whether one could assign ordinals $< \epsilon_0$ to the closed terms of $\mathbf{T}$ in such a way that $s \succeq t$ and $s \neq t$ imply that the ordinal of t is less than the ordinal of s. He wrote about such an assignment in the passage from the Princeton lecture quoted by Troelstra, but it seems clear that he never actually had one. Howard [1970] produced one which works for a restricted notion of reduction (but one that nevertheless admits reduction to normal form). But in my own last contact with Gödel, a phone conversation sometime in 1973 about something quite different, he raised the question of such an assignment.

> made precise by specifying the supplementations, the axioms, and the evidences to be used.

We are to think of the definition of a function as a theorem, stating the unique existence of the object defined. The reductive proof is a proof of this theorem.[7] In our special case, the axioms are presumably the unique existence of 0 and the successor function. Moreover, it is to be decidable whether or not a proposition has a reductive proof.

I will not take the space to discuss all the alternative readings that I have tried: Let me just say that I have failed to find a coherent reading of this footnote. The idea seems to be that the predicates $\mathcal{C}_A$ are being replaced by new *decidable* predicates $\mathcal{C}'_A$, where $\mathcal{C}'_A(t)$ asserts the reductive provability of the defining equations of t (given the reductive provability of their arguments—however that is to be understood). But then $\mathcal{C}'_A(t)$ cannot imply $\mathcal{C}_A(t)$, i.e. that t is computable. So, in whatever sense it may be valid, the footnote cannot solve the global problem.

Gödel seems to have continued to believe, at least sometimes, that his notion of reductive proof solved the local problem at least. In an undated and unsent reply to a letter by F.W. Sawyer III in 1974 [**?**, pp. 210-11], he refers to footnote h again as the solution of the problem—although he states there that "to carry that out in detail is rather cumbersome, and the matter probably cannot be explained convincingly in a footnote." I don't see that he ever clearly addresses the global problem.

5. In any case, Gödel's foundation for his theory **T**, given his aims, seems to fail. I want now to consider a different approach to its foundation. Instead of Gödel's notion of *computable* function of finite type, we start with the basic concept of a function *simpliciter* of finite type over $\mathbb{N}$ and ask: what is contained in this concept? I believe that the answer to this question simply is: the principles of definition and proof in **T**. (It would in fact suffice to show merely that the concept of function of finite type over $\mathbb{N}$ at least *contains* the principles of definition and proof in **T**.) We leave out of account, as I have already argued that we should, the idea of computability. Rather, we are concerned only with what objects we can *construct* in the domain of finite types over $\mathbb{N}$, given arbitrary objects of specific such types. The notion of function is in itself *not* limited to what we can so construct. Among

[7] Gödel states that the 'type character' of the function is to be regarded as a part of its definition; but clearly this is not part of what has to be proved.

the functions of a given type, some will be constructible on the basis of the concept of finite type over $\mathbb{N}$ and some may not (but may be constructible in a domain including transfinite types over $\mathbb{N}$, for example). The proper locus of constructivity is in our *reasoning* and, in particular, in our reasoning about numbers and functions, not in the concepts of number and function.

6. We begin by separating the notion of the system FT of finite types over some ground types from that of the system $FT(\mathbb{N})$ of finite types over $\mathbb{N}$ and asking: What specific objects of type B can be constructed in the domain FT from arbitrary object of types $A_1, \ldots, A_n$, respectively, where B and the A_i are in FT? Of course, we are not really speaking of objects here, since the atomic types of FT are unspecified. Rather, we are asking what specific terms or, as I will call them, *formal objects* of type B can be constructed from variables $v_1, \ldots, v_n$ of types $A_1, \ldots, A_n$, respectively. Clearly, these objects should be closed under pairing, left and right projection, evaluation and λ-abstraction—call them the *operations of* $\mathcal{P}$. I would like to have, but don't, a satisfactory analysis of the notion of 'formal object' and be able to argue that they constitute the least class of terms containing the v_i and closed under these operations, i.e., that they are precisely the terms of $\mathcal{P}$ containing no (free) variables other than the v_i.

There is in fact another question: What types are *non-empty* for all interpretations of the atomic formulas? Reasoning classically, these are different questions. For example Peirce's law

$$[(A \to B) \to A] \to A$$

is classically non-null for every interpretation; but it seems clear that, when A is atomic for example, there is no formal object of this type.

As was first pointed out by Läuchli in [Läuchli, 1965], the choice of question is a choice of logics: The types of formally defined objects are, from either a classical or constructive point of view, the theorems of intuitionistic logic of implication and conjunction; the non-empty types are, speaking classically, the theorems of the corresponding classical logic.)[8]

Note, incidentally, that our question is also not Läuchli's, either in [Läuchli, 1965] or in [Läuchli, n.d.]. The question he answered is: for what A is there a formal object of type A? In [Läuchli, 1965], he defines the formal objects

[8] I was kindly reminded by Denis Bonnay of the relevance of Läuchli's work to the present topic.

to be certain terms in the theory $\mathcal{H}$ of an ϵ-structure consisting of infinitely many urelements and sets of finite rank over the set Π of urelements, extended by explicit definition of functions. The question of whether A has a formal object becomes a question of the validity in this structure of some formula $s \in t(A)$, where $t(A)$ codes the type A and the free variables in s and $t(A)$, representing the atomic formulas in FT, range over the sets of urelements. His result shows that there is such a term s iff A is a theorem of intuitionistic logic, i.e., iff there is a closed term of $\mathcal{P}$ of type A. When one's concern is only for logic, for what is provable, this result suffices. But our concern is with type theory, and the question is not about which types A have formal objects, but about what formal objects of type A there are. Moreover, for this purpose, the set theoretic framework of Läuchli's analysis is inappropriate, since we are asking what formal objects of type A are contained in the concept of finite type, not in the concept of $\mathcal{H}$. That his analysis is not what we want is clear from the example $s \in t(A \to (A \to A))$, where A is an atomic formulas and s is defined by

$$suv = \begin{cases} u & \text{if } A \text{ is finite} \\ v & \text{otherwise} \end{cases}$$

In [Läuchli, n.d.] there is a better analysis of the notion of 'formal objects' of type A, according to which they are not syntactical objects, but are certain objects in $\mathcal{H}$. Let $\mathcal{D}$ be the least class containing Π and ${}^{D}E$ and $D \times E$ whenever $D, E \in \mathcal{D}$. The elements of $\bigcup \mathcal{D}$ are called *functionals*. For each $A \in FT$, $S(A)$ is defined by

- $S(A) = \Pi$, if A is atomic.
- $S(A \to B) = {}^{S(A)}S(B)$.
- $S(A \wedge B) = S(A) \times S(B)$.

A permutation σ of Π induces a permutation of each $S(A)$ as follows: For $f \in S(A \to B)$, $\sigma f = \sigma \circ f \circ \sigma^{-1}$. For $(a, b) \in S(A \wedge B)$, $\sigma(a, b) = (\sigma a, \sigma b)$. A functional $\phi \in S(A)$ is *invariant* iff $\sigma\phi = \phi$ for all permutations σ.

An assignment p of a set $p(A) \subseteq \Pi$ to each atomic A, called a *proof system*, induces an assignment of a set $p(A)$ of functionals to each type A. Namely

- $p(A \to B) = \{f \in S(A \to B) \mid \forall x \in p(A) fx \in p(B)\}$.

- $p(A \wedge B) = p(A) \times p(B)$.

Thus, S is the maximum proof system. The notion of 'formal object' of type A that we can derive from this second paper of Läuchli is that of an invariant functional that is in $p(A)$ for every proof system p, and it seems plausible that anything that we might want to count as a formal object will be represented by such a functional. They are in any case easily seen to be closed under the operations of $\mathcal{P}$.

Question. Is every invariant functional which is in $p(A)$ for each proof system p defined by a closed term of $\mathcal{P}$ of type A?

Notice that the term s defined above corresponds to the two invariant functionals f and g in $p(A \to (A \to A))$ for every proof system p, defined by $fab = a$ and $gab = b$, respectively. s denotes f in the proof system p when $p(A)$ is infinite and otherwise denotes g.

7. When we turn to the question of what is contained in the concept of a number, i.e. an object of type $\mathbb{N}$, surely the answer is it is precisely the idea of a finite iteration. A numerical variable v stands for an arbitrary finite iteration, which we may regard both as an object in itself, that can be generated from the null iteration 0 by means of the successor operation, and as a (type-free) operation which can be applied to any operation on any domain to yield its iterate.[9]

The formal objects that can be constructed in the domain of finitary types—essentially types of the form $\mathbb{N} \wedge \cdots \wedge \mathbb{N}$—from variables $v_1, \ldots, v_n$ of finitary types are exactly the terms of PRA containing only these variables. (See [Tait, 1981]. The analysis of the notion of a formal object was easy in this case.) If we are right about the formal objects of $\mathcal{P}$, then the formal objects that can be constructed in the domain $FT(\mathbb{N})$ from variables $v_1, \ldots, v_n$ of finite types over $\mathbb{N}$ are precisely the terms of **T** containing only free variables from among $v_1, \ldots, v_n$.

8. On the foundation of **T** we are considering, objects of type $A \to B$ are

[9] See [Tait, 1981]. Weyl (see [Weyl, 1921], the end of Part II §1 ("The Basic Ideas") and the first paragraph of §2a ("Functio Discreta"), and also [Weyl, 1949, 9. 33]) is in agreement with this analysis. Note also that Wittgenstein, in the *Tractatus Logico-Philosophicus* ([Wittgenstein, 1922]), and Church, in his *The Calculi of Lambda Conversion* [Church, 1941], represent the numbers as (type-free) iterations.

functions from A to B. This notion of function is open-ended: Whatever principles we introduce for defining functions of a certain type $\neq \mathbb{N}$, it will always be possible to introduce new such principles that yield more functions of that type. On the other hand, we *are* restricted to functions that are in some way *defined* and so must reject the idea of absolutely 'random' functions. For functions on our conception must be objects of proof and proofs can only refer to objects via particular representations of them.

So the notion of definitional equality makes sense for functions in our sense, but we lack Gödel's motivation to interpret equations of higher type in terms of definitional equality, namely that they are then decidable. For, even when we restrict ourselves to closed terms of **T**, the proof that definitional equality is decidable requires all of HA. We do interpret equations between closed numerical terms of **T**, $s = t$, to mean definitional equality, but there is no assumption that this relation is decidable. For s and t of type $A \to B$, $s = t$ has its usual extensional meaning:

$$s = t := \forall x{:}A.\, (fx = gx)$$

and for type $A \wedge B$

$$(s, t) = (s_0, t_0) := s = s_0 \wedge t = t_0.$$

It follows of course that the atomic formulas of **T** cannot be assumed to be decidable and that the propositional logic contained in **T** must therefore be understood as intuitionistic. Moreover, equations now have logical structure in general, and so, for the foundations of **T**, we have ended up, after all, in the domain of intuitionistic logic that it was precisely Gödel's aim to reinterpret. But his attempt to provide a foundation for **T** that avoids intuitionistic logic fails. Our aim is rather to provide a foundation for *both* **T** and HA that avoids the objections that Gödel raised concerning the latter.

9. In [Gödel, *1941], Gödel takes the fact that PA can be regarded as a subsystem of HA (namely the restriction of HA to the logical constants $\forall$ and $\to$) as a symptom that intuitionistic logic oversteps the bounds of constructivity. His diagnosis of how exactly it does so is that the intuitionistic meaning of the implication $P \to Q$, for example, requires a primitive notion of proof; for it means that every proof of P can be transformed into a proof of Q. He points out that this notion of a proof cannot be restricted to the proofs in some particular formal system, and so "proof" must be taken in an absolute sense. His contention is that

> [T]his notion of an intuitionistically correct proof or constructive proof lacks the desirable precision. In fact one may say that it furnishes itself a counterexample against its own admissibility, insofar as it is doubtful whether a proof utilizing this notion of proof is constructive or not (p. 194).

Presumably the difficulty with such a proof being constructive is that it makes at least implicit reference to all proofs: it is in this way impredicative. For example, a proof of $P \to Q$ is a proof about all proofs. So imprecision and impredicativity are the charges against intuitionistic logic as Heyting analyzed it, where the impredicativity is taken as a sign that the underlying absolute notion of proof has not been thought through sufficiently.

The reason he gives for why the notion of proof underlying intuitionistic logic in Heyting's sense cannot be proof in some particular formal system is that "For this notion the axioms of intuitionistic logic would not hold." The grounds for this remark most likely come from his paper "An interpretation of the intuitionistic propositional logic" [Gödel, 1933f] in which, reading the necessity operator $\mathcal{B}$ as meaning "it is provable that", he gives a translation of Heyting's propositional calculus into the modal system $S4$ such that every theorem of Heyting's system translates into a theorem of $S4$ and, for example, the translation of $P \vee \neg P$ is not a theorem.

$$\mathcal{B}[\mathcal{B}P \implies P]$$

is a theorem of $S4$, and so when he writes that $\mathcal{B}$ cannot be understood as provability in a formal system, he is referring to the fact that, if the formal system is consistent and includes HA, say, then the above theorem is false when we take P to be the formula $0 = 1$. That this is what Gödel had in mind with his remark that the relevant notion of proof cannot be proof in some fixed formal system is further corroborated by his reference to intuitionism in [Gödel, 1938a] as the "modal-logical route" to extending constructive mathematics beyond finitary number theory. [10] Noting there the non-constructive

[10] This undoubtedly has a source in Heyting's discussion of intuitionistic logic in [Heyting, 1930b], where he too introduced $\mathcal{B}P$ as a meaningful proposition which, at least in some cases, is distinct from P itself. But in his later treatments of intuitionistic logic, he drops this: Every proposition P requires a 'construction' or proof as its warrant. To assert that P is provable is no more than to assert that P. If $\mathcal{B}P$ were a proposition distinct from P, what would *its* proofs be, other than proofs of P? See [Tait, 2001, p. 118] for further discussion of this.

character of $S4$ as he interpreted it (since $\mathcal{B}P$ is a disguised existential proposition), Gödel goes on to say that this non-constructivity can be avoided by replacing provability as the basic notion by the proof *relation* '$z \vdash P$' or '$z \vdash P, Q$', meaning that z is a proof of Q from P. Gödel did not develop the treatment of intuitionism in terms of the provability relation, and in particular he neither explicitly identifies the range of the variable z in "$z \vdash P$" nor tries to define the meanings of the logical constants in terms of the provability relation. In the 1960's Kreisel took the variables to range over constructions and undertook to create a general theory of type-free constructions in which the logical constants can be defined. Kreisel [Kreisel, 1962; Kreisel, 1965], Nicolas Goodman, e.g., [Goodman, 1970], and others obtained limited results in this direction, but they are of an *ad hoc* character: given a part of logic or arithmetic to be interpreted, one finds axioms for the theory of constructions which suffices for it; and with the development of the Curry-Howard theory, which introduces type structure into the notion of proof, the theory of constructions was more or less abandoned.[11] Gödel's doubts about the constructive character of intuitionistic logic seem to me to have been tied to this conception of Heyting's interpretation of the logical constants, according to which it involves quantification over a type-free domain of constructions. As far as I know, there is no mention in Gödel's papers of the propositions-as-types conception.

10. The Curry-Howard theory of propositions-as-types provides a foundation for Heyting's logic and is, at the same time, a simple generalization of Gödel's primitive recursive functions of finite type. Namely the sentences (i.e. closed formulas) of HA are themselves to be understood as types of objects. Likewise, the Curry-Howard theory completes our foundation for the theory $\mathbf{T}$ by interpreting the universal closures of formulas of $\mathbf{T}$ and, in particular, of the equations of higher type, as types. In fact, we will combine the two by considering the theory $\mathbf{HA}^{\omega}$, which extends $\mathbf{T}$ by adding intuitionistic quantification theory over each type in $FT(\mathbb{N})$.[12]

[11] The situation is quite analogous to the type-free (i.e. rank-free) theory of sets. There is no internal ground for the choice of axioms, only the external motive of wanting to derive this or that particular result. Once one adopts the cumulative conception of sets, the axioms (of ZF and much more) suggest themselves readily, as following from the conception.

[12] The embedding of $\mathbf{HA}^{\omega}$ in the Curry-Howard theory is carried out in [Martin-Löf, 1998], the manuscript for which was apparently distributed to a number of people in 1972.

As in the case of the primitive recursive functions of finite type, where we separated the theory of finite types over a set of arbitrary atomic types from the contribution of $\mathbb{N}$ when we take it to be the only atomic type, in the present case, we could separate intuitionistic or *Heyting logic* $\mathbf{HL}^{\omega}$ of finite type from its special application $\mathbf{HA}^{\omega}$ to arithmetic and ask what formal objects of type ϕ are there, when ϕ is a sentence of $\mathbf{HL}^{\omega}$. But what I could say about that can be gleaned from what I have had to say about formal objects of types in FT (namely, not enough), and in the interests of brevity, I will leave this question aside. As in the earlier case, it suffices anyway simply to see that the terms we actually introduce should count as formal objects.

We have already specified the types expressed by equations between closed numerical terms. We also know what types $\phi \rightarrow \psi$ and $\phi \wedge \psi$ are, once we know what types the sentences ϕ and ψ express. It remains only to explain the meaning of the sentences $\forall x : A.\, \phi(x)$ and $\exists x : A.\, \phi(x)$, where $\phi(v)$ is a formula containing only the variable v of type A. But these are simply generalizations of $A \rightarrow B$ and $A \wedge B$, respectively: Namely, corresponding to

$$A \rightarrow B :=\ {}^{A}B \qquad\qquad A \wedge B := A \times B$$

we have

$$\forall x{:}A.\, \phi(x) := \Pi_{x:A}\phi(x) \qquad\qquad \exists x{:}A.\phi(x) := \Sigma_{x:A}\, \phi(x)$$

In other words, $\forall x{:}A.\, \phi(x)$ is the type of functions f defined on A such that, for $a{:}A$, $fa{:}\phi(a)$, and $\exists x{:}A.\, \phi(x)$ is the type of pairs (a, b) such that $b{:}\phi(a)$. Note that we have now specified the type of all equations $s = t$ between closed terms of $\mathbf{T}$ and so of all sentences of $\mathbf{HA}^{\omega}$.

Evaluation extends to the case in which s is of type $\forall x : A.\, \phi(x)$ and $t : A$. Namely st is of type $\phi(t)$. So in this case, evaluation corresponds to $\forall$-*instantiation* or -*elimination*. Similarly, pairing extends to forming pairs (s, t) of type $\exists x{:}A.\, \phi(x)$, where $s{:}A$ and $t{:}\phi(s)$. It corresponds, in this case, to $\exists$-*introduction*.

The defined operations of λ-abstraction and projection also extend to these new function and pairing types, too, as follows:

If v is a variable of type A and $t(v) : \phi(v)$, then $\lambda x : A.\, \phi(x)$ is of type $\forall x{:}A.\, \phi(x)$. This corresponds to *universal quantifier introduction.* However, we need to make a familiar restriction on λ-abstraction in the context of $\mathbf{HL}^{\omega}$. Suppose that $t(v)$ contains a variable u of some type $\psi(v)$, which contains one or more occurrences of v. Then $\lambda x\!:\!A.\, \phi(x)$ cannot be formed. To understand this in the more usual logical context, note that the variable u, as a term of type $\psi(v)$ corresponds to a hypothesis of $\psi(v)$, so that $t(v)$ is a proof of $\phi(v)$ from the hypothesis $\psi(v)$. So our restriction coincides with that of not generalizing on a variable that occurs in a hypothesis of a proof.

The case of projections is less straightforward vis-a-vis the usual formalization of logic. Let $b\!:\!\exists x\!:\!A.\, \phi(x)$. Then bL is of type A and bR is of type $\phi(bL)$. But bL is not a term in the sense of $\mathbf{T}$, and so $\phi(bL)$ is not a formula of $\mathbf{HL}^\omega$ when the variable v of type A actually occurs in $\phi(v)$. Gentzen avoided this problem by formulating a form of $\exists$-elimination which stays within $\mathbf{HL}^\omega$; namely for any formula ψ (meaning in particular that it does not contain the symbol x)

$$\exists x\!:\!A.\, \phi(x),\ \forall x\!:\!A.\, [\phi(x) \rightarrow \psi]\ \Rightarrow\ \psi$$

This is easily derivable using projections: if $b\!:\!\exists x\!:\!A.\, \phi(x)$ and $f\!:\!\forall x\!:\!A[\phi(x) \rightarrow \psi]$, then $f(bL)(bR)\!:\!\psi$. So the type-theoretic conception of intuitionistic logic includes all the logical inferences of $\mathbf{HL}^\omega$. However, it also leads out of the class of formulas of $\mathbf{HL}^\omega$ to formulas that contain terms of type A which are not terms of $\mathbf{T}$. But this in no way detracts from the point we are making that $\mathbf{HL}^\omega$ can be founded on a notion of proof which is a natural generalization of the (impredicative) notion of a primitive recursive function of finite type over $\mathbb{N}$: it only indicates a certain incompleteness in the logical framework of $\mathbf{HL}^\omega$.[13]

11. The only remaining types of inference in $\mathbf{HA}^\omega$ that we have not covered are the principle of iteration, which we have already discussed, and the principle of proof by mathematical induction

$$\phi(0),\ \forall x\!:\!\mathbb{N}[\phi(x) \rightarrow \phi(x')]\ \ \Rightarrow\ \ \forall x\!:\!\mathbb{N}.\, \phi(x)$$

But, given s of type $\phi(0)$, t of type $\forall x : \mathbb{N}[\phi(x) \rightarrow \phi(x')]$, and r of type $\mathbb{N}$, $f = \lambda x : \mathbb{N}.R(x, s, t)$ is of type $\forall x : \mathbb{N}.\phi(x)$. The reduction of this primitive recursion to iteration simply generalizes the reduction in the case discussed in §1, where the types are in $FT(\mathbb{N})$: Let $h : \exists x : \mathbb{N}.\, \phi(x)$ be defined by

[13] In the case of first-order logic, this incompleteness is inessential in the sense that every first-order sentence which has a proof in the sense of type theory (i.e. using the projections for $\exists$-formulas) from a set of first-order premises together with the assumption that the domain of individuals is non-empty is already provable in HA from those premises. See [Tait, 2003].

$h(v,u) = (v', tvu)$. If $\bar{f} = \lambda x : \mathbb{N}.\, I_{\psi}(x, (0,a), h)$, where $\psi = \exists x : \mathbb{N}.\, \phi(x)$, then $f = \lambda x{:}\mathbb{N}.\, \bar{f}(x)R$. For $\lambda x{:}\mathbb{N}.\, (\bar{f}x)L = \lambda x{:}\mathbb{N}.\, I_{\mathbb{N}}(x, 0, ')$ = the identity function on $\mathbb{N}$. So $f0 = s$ and $fv' = t(\bar{f}v)L(fv) = tv(fv)$.

12. We have exhibited Heyting's underlying proof theoretical foundation of intuitionistic arithmetic as consisting of a hierarchy of types over $\mathbb{N}$ which is is a simple generalization of $FT(\mathbb{N})$. Viewed in this way, it has the same claim to constructivity as the 'proof theory' of $FT(\mathbb{N})$, namely **T**. It is difficult to see why one should regard the notion of proof in intuitionistic arithmetic as in any sense more problematic than the primitive recursive functions of finite type. In particular, the fact that the types ϕ of HA involve quantifiers is of no consequence, since the use of bound variables is in principle eliminable. (See [Weyl, 1918] and, more generally, [Tait, 1998b].)

The circularity that Gödel felt is involved in proving propositions, say, of the form $(A \to B) \to C$, which seemed to him to concern proofs referring to 'all proofs', disappears once we see that proofs are stratified into types, and that the types of the proofs associated with each of the distinct subformulas of this formula are all distinct. Certainly this foundation for intuitionistic arithmetic still involves impredicativity: If we prove the induction step $\forall x[\phi(x) \to \phi(x')]$ itself by induction, then this is a case of impredicative definition, but *in exactly the same sense as in* **T** when we define $g : (\mathbb{N} \to \mathbb{N}) \to (\mathbb{N} \to \mathbb{N})$ using primitive recursion and then define $fn = g^n a$ for some $a{:}\mathbb{N} \to \mathbb{N}$. Gödel's further complaint about intuitionistic logic that the notion of proof involved is 'absolute' and cannot be confined to the proofs in some definite formal system now appears as perfectly analogous to the situation regarding the finite types over $\mathbb{N}$. The primitive recursive functions of finite type A themselves form an incomplete totality of objects of type A. By going to transfinite types we obtain, as Gödel, himself, observed, new functions of type A. Similarly, by going to types 'higher' than the first-order types in the extended sense, we would obtain new proofs of sentences of HA. In both cases, function types $A \to B$ or $\forall x F(x)$ are open-ended.

Our foundation for intuitionistic logic is not *prima facie* entirely faithful to Heyting's. Heyting defines a proof of $\phi \to \psi$, for example, to be a method of transforming any proof of ϕ into a proof of ψ, whereas on our account, a proof of $\phi \to \psi$ is a function from the objects of type ϕ to the objects of type ψ. What counts as a proof in this sense is always relative to the system one is considering—the finite types over $\mathbb{N}$, for example, or the sentences of HA or of $\mathbf{HA}^{\omega}$. However, it seems reasonable to take the view that any adequately defined function of type $\phi \to \psi$ is constructed within some extension of $\mathbf{HA}^\omega$, and so the notion of a proof of some sentence χ of $\mathbf{HA}^\omega$ would coincide with that of an object of type χ, where the notion of proof here is indeed open-ended.

Given that Gödel's interpretation of HA requires a distortion of the intuitionistic meaning of the logical constants—for example, ignoring that, in a proof $p{:}A \to \exists x{:}B.\ \phi(x)$, when $a{:}A$, then the witness $x = (pa)L$ in general depends on a—clearly the *Dialectica* interpretation offers no advantage in understanding intuitionistic logic.

13. For my final remark, I would like to repeat that, in what seems to me to be the only viable foundation for Gödel's theory of primitive recursive functions and indeed for $\mathbf{HA}^\omega$, nothing is said about computability: The objects introduced are the objects whose existence follows from the concepts involved, i.e., which can be defined or constructed from the operations implicit in these concepts. In the context in question, computability follows from contructablity; but that is a *theorem.* It is not built into the concepts themselves.

The confusion between what is computable and what is constructive (i.e. can be constructed) has led to the debate, for example, about whether Markov's principle, that for a decidable numerical predicate $\phi(v)$,

$$\neg\forall x{:}\mathbb{N}.\ \phi(x) \to \exists x{:}\mathbb{N}.\ \neg\phi(x),$$

is constructive. This principle holds under the *Dialectica* interpretation[14] and, indeed, if the antecedent is true, a witness for the conclusion can be computed. But the principle is not valid intuitionistically. Indeed, it is not constructive: From the assumption of the antecedent $\neg\forall x{:}\mathbb{N}.\ \phi(x)$, i.e. from a variable v of this type, no proof of the conclusion can be constructed. On might be tempted to argue that we can construct a witness by iterating a search procedure; but no iteration (i.e. bound on the search) is given us by v. I would rather view Markov's principle as an example of why, if one is looking for methods of proof which automatically yield algorithms for computing a witness for existential theorems, intuitionistic logic is too narrow.

[14] Gödel cites this at the end of footnote h in [Gödel, 1972] as evidence for the "higher degree of constructivity" of his interpretation of the intuitionistic logical constants over Heyting's.

References

Avigad, J. and Feferman, S. [1998]. Gööödel's functional ("Dialectica") interpretation, *[Buss, 1998]* pp. 337–405.

Buss, S. (ed.) [1998]. *Handbook of Proof Theory*, Amsterdam: Elsevier.

Cantor, G. [1891]. Über eine elementare Frage der Mannigfaltigkeitslehre, *Jahresbericht der deutschen Mathematiker-Vereiningung* **1**: 75–78. In [Cantor, 1932]. Translated in [Ewald, 1996], volume 2.

Cantor, G. [1932]. *Gesammelte Abhandlungen mathematischen und philosophischen Inhalts*, Berlin:Springer. ed. E. Zermelo.

Church, A. [1941]. *The Calculi of Lambda-Conversion*, Princeton: Princeton University Press.

Dekker, J. (ed.) [1962]. *Recursive Function Theory, Proceedings of Symposia in Pure Mathematics*, Vol. 5, Providence: American Mathematical Society.

Ewald, W. [1996]. *From Kant to Hilbert: A Source Book in the Foundations of Mathematics. Two Volumes*, Oxford: Oxford University Press.

Gödel, K. [1933f]. Eine Interpretation des intuitionistischen Aussagenkalküls, *Ergebnisse eines mathematischen Kolloquiums* **4**: 39–40. Reprinted with an Englsh translation in [Gödel, 1986].

Gödel, K. [1938a]. Lecture at Zilsel's, *In [Gödel, 1995]*, pp. 87–113.

Gödel, K. [*1941]. In what sense is intuitionistic logic constructive?, *[Gödel, 1995]*, pp. 189–201.

Gödel, K. [1958]. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, *Dialectica* **12**: 280–287. Reprinted with an Englsh translation in [Gödel, 1990, 240-252]. [Gödel, 1972] is a revised version.

Gödel, K. [1972]. On an extension of finitary mathematics which has not yet been used, *[Gödel, 1990]*, pp. 271–280. Revised version of [Gödel, 1958].

Gödel, K. [1986]. *Collected Works, Vol. I*, Oxford: Oxford University Press.

Gödel, K. [1990]. *Collected Works, Vol. II*, Oxford: Oxford University Press.

Gödel, K. [1995]. *Collected Works, Vol. III*, Oxford: Oxford University Press.

Goodman, N. [1970]. A theory of constructions equivalent to arithmetic, *[Kino, Myhill and Vesley, 1970]*, pp. 100–120.

Heyting, A. [1930b]. Sur la logique intuitionniste, *Académie Royale de Belgique, Bulletin 16* pp. 957–963.

Hilbert, D. [1926]. Über das Unendliche, *Mathematische Annalen* **95**: 161–90. Translated by Stefan Bauer-Mengelberg in [van Heijenoort, 1967, 367-92].

Hindley, J. and Seldon, J. (eds) [1980]. *To H.B. Curry: Essays on Combinatorial Logic, Lambda Calculus and Formalism*, London: Academic Press.

Howard, W. [1970]. Assignment of ordinals to terms for primitive recursive functionals of finite type, *In [Kino et al., 1970]* pp. 443–458.

Howard, W. [1980]. The formula-as-types notion of construction, *in* J. Hindley and J. Seldin (eds), *[Hindley and Seldon, 1980]*, pp. 479–490.

Kino, A., Myhill, J. and Vesley, R. (eds) [1970]. *Intuitionism and Proof Theory*, Amsterdam: North-Holland.

Kreisel, G. [1962]. Foundations of intuitionistic logic, *[Nagel, Suppes and Tarski, 1962]*, pp. 198–210.

Kreisel, G. [1965]. Mathematical logic, *[Saaty, 1965]*, pp. 95–195.

Läuchli, H. [1965]. Intuitionistic propositional calculus and definably non-empty terms. (abstract), *The Journal of Symbolic Logic* **30**: 263.

Läuchli, H. [n.d.]. An abstract notion of realizability for which intuitionistic predicate calculus is complete, *In [Kino et al., 1970]* pp. 227–234.

Mancosu, P. (ed.) [1998]. *From Brouwer to Hilbert: The Debate on the Foundations of Mathematics in the 1920's*, Oxford: Oxford University Press.

Martin-Löf, P. [1998]. An intuitionistic theory of types, *In [Sambin and Smith, 1998, 221-244]* .

Nagel, E., Suppes, P. and Tarski, A. (eds) [1962]. *Logic, Methodology and the Philosophy of Science*, Stanford: Stanford University Press.

Saaty, T. (ed.) [1965]. *Lectures on Modern Mathematics*, New York: Wiley.

Sambin, G. and Smith, J. (eds) [1998]. *Twenty-Five Years of Constructive Type Theory*, Oxford: Oxford University Press.

Spector, C. [1962]. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of the principles formulated in current intuitionistc mathematics, *[Dekker, 1962]* pp. 1–27.

Tait, W. [1963]. A second order theory of functionals of higher type, with two appendices. Appendix A: Intensional functionals. Appendix B. An interpretation of functionals by convertible terms., *Stanford Seminar Report* pp. 171–206. A published version is [Tait, 1967].

Tait, W. [1967]. Intensional interpretations of functionals of finite type I, *Journal of Symbolic Logic* **32**: 198–212.

Tait, W. [1981]. Finitism, *Journal of Philosophy* **78**: 524–556.

Tait, W. [1998b]. Variable-free formalization of the Curry-Howard type theory, *[Sambin and Smith, 1998, 265–274]* .

Tait, W. [2001]. Gödel's unpublished papers on foundations of mathematics, *Philosophia Mathematica* **9**: 87–126.

Tait, W. [2003]. The completeness of Heyting first-order logic, *The Journal of Symbolic Logic* **68**(3): 751–763.

van Heijenoort, J. (ed.) [1967]. *From Frege to Gödel: A Source Book in Mathematical Logic*, Cambridge: Harvard University Press.

Weyl, H. [1918]. *Das Kontinuum: Kritische Untersuchungen über die Grundlagen der Analysis*, Leipzig: Veit.

Weyl, H. [1921]. Über die neue Grundlagenkrise der Mathematik, *Mathematische Zeitschrift* **10**: 39–79. Translated by P. Mancosu in [Mancosu, 1998].

Weyl, H. [1949]. *Philosophy of Mathematics and Natural Science*, Princeton: Princeton University Press.

Wittgenstein, L. [1922]. *Tractatus Logico-Philosophicus*, London: Routledge & Kegan Paul. German/English. Tr. D. F. Pears and B. F. McGuinness, 1961 (second edition, 1963).