

GOODMAN'S THEOREM AND BEYOND

MICHAEL BEESON

Goodman proved that if the axiom of choice $\forall x \exists y A(x, y) \rightarrow \exists f \forall x A(x, f(x))$ is added to intuitionistic arithmetic (here x, y , and f are functionals of finite type), then no new arithmetic theorems are obtained. His original proof used his theory of constructions, and was widely thought to be perhaps not the simplest proof possible. Moreover, the generalization of the result to show that the axiom of choice remains conservative even in the presence of extensionality, has remained unsolved, despite its inclusion in Friedman's list of "102 Problems in Mathematical Logic". In this paper, we give a conceptually clear proof of Goodman's theorem, and use the method to solve Friedman's problem just mentioned, as well as to obtain several other extensions of Goodman's theorem and related results.

The proof is in two steps—one step uses realizability, the other step uses forcing; in other words, two well-known tools are combined to get the result. Goodman himself has also recently given a fairly simple proof of his theorem; his proof also combines ideas related to realizability and forcing, but in a single new interpretation, instead of clearly separated. The method of this paper seems to be more easily generalized.

In [1], there is an application of the extensions of Goodman's theorem [7] proved in this paper. Friedman has given a formal theory \underline{B} for constructive mathematics which is strong in the sense that all of constructive mathematical practice can be formalized in it, but proof-theoretically weak in that it is conservative over (classical) arithmetic. We apply the results of this note to show that it is also conservative over intuitionistic arithmetic, a problem left open by Friedman.

0. Preliminaries. We assume familiarity with the system HA of intuitionistic arithmetic (see e.g., [12] for a description) and with the system HA° which has variables for functionals of finite type. If one is only interested in Goodman's theorem and the solution of Friedman's problem, this will be enough. For some theorems, we shall also assume familiarity with the systems developed by Feferman for constructive mathematics, and with Friedman's system \underline{B} . Feferman's systems are described in [1], [2], and of course the original [5]. We find the two-sorted version of these theories described in [1] the most convenient for forcing, and we assume

they are so formulated, with one sort of variables for natural numbers. Friedman has also found it more convenient to use a two-sorted theory, with special variables for natural numbers, and we also use that version of the intuitionistic set theories. (In [6], Friedman equivalently uses two unary predicates and one sort of variables.) In [3] we have described some intuitionistic set theories without the axiom of extensionality; since that paper will in any case be essential to the reader of our theorems concerning intuitionistic set theories, we do not repeat the descriptions here. By *HAS* we mean *HA* with species (set) variables and full (second order) comprehension. By *HAS^ω* we mean *HAS* with variables for functionals of finite type and for species of functionals of each type, with full comprehension.

By *AC* we mean the schema in either *HA^ω* or *HAS^ω*,

$$(AC) \quad \forall x \exists y A(x, y) \longrightarrow \exists f \forall x A(x, f(x)) .$$

By *DC* we mean the axiom of dependent choices, which can be expressed either in Feferman's theories or in Friedman's:

$$(DC) \quad \forall x \in A \exists y \in A \phi(x, y) \rightarrow \forall x \in A \exists f \in A^N (f(0) = x \text{ and } \forall n \phi(f(n), f(n')))$$

where N is the natural numbers and n' is the successor of n .

1. **How not to prove Goodman's theorem.** We give an incomplete (but simple) proof of Goodman's theorem, because it will eventually lead to a proof. We assume familiarity with Kleene's notion of recursive realizability (see [12] for explanations). If A is a formula of *HA*, then erA (" e realizes A ") is also a formula of *HA*; and we have the soundness theorem $\vdash A$ implies $\vdash \bar{e}rA$ for some e . We can extend this notion to *HA^ω* by using the hereditarily recursive operations *HRO* (see [12]); we write erA to abbreviate $er(A)_{HRO}$. Then *AC* is provably realized, so if *HA^ω* + *AC* $\vdash \phi$, we have *HA^ω* $\vdash \bar{e}r\phi$ for some e . Now, if $\vdash (er\phi \rightarrow \phi)$, we are done, for then *HA^ω* $\vdash \phi$ (and of course *HA^ω* is conservative over *HA*). Thus *HA^ω* + *AC* is conservative over *HA* for all A such that $\vdash (erA \rightarrow A)$; this holds in case A is "almost negative" (i.e., contains no disjunction and has existential quantifiers only immediately in front of atomic formulae). Now, here is the way not to prove Goodman's theorem: suppose we use, instead of recursive functions, functions recursive in some given function \underline{a} , for our realizability. If \underline{a} is "sufficiently nonrecursive", maybe we can get $erA \rightarrow A$ (for a fixed A). This will suffice. For instance, why not let \underline{a} encode the characteristic function of all the subformulae of A ? In this way one can successfully prove that *HA^ω* + *AC* is conservative over

classical arithmetic PA , but there are difficulties in replacing PA by HA ; for instance, one can not prove $\mu x A(x, y)$ is recursive in A unless A is decidable. However, there seems to be the germ of a good idea here; if only we know how to choose \underline{a} properly.

2. **Forcing.** The answer is, to choose \underline{a} to be generic with respect to a suitable set of forcing conditions. We now describe the set-up necessary to use forcing in an intuitionistic context. If T is any reasonable intuitionistic theory, let Ta be formed by adding to T a constant \underline{a} for a generic partial function from N to N . (It turns out that we have to add a partial function, not a total function; below we discuss how to talk about partial functions in HA° .) Our forcing conditions will be finite functions from N to N (which can be coded as sequence numbers if necessary, in systems which cannot speak directly about finite functions). Let C be any set of such conditions, definable in T . We use p, q, r as meta-variables ranging over C ; that is, $\forall p$ abbreviates $\forall p(p \in C) \rightarrow \dots$, etc. Following the usual (backwards) notation for forcing, we use $p \leq q$ to mean p extends q , i.e., p is defined wherever q is, and agrees with q on the domain of q , but p may have a larger domain. Now we assign to each formula A of Ta , a formula $p \Vdash A$ of T ; the free variables of $p \Vdash A$ are p together with a variable x^* for each free variable x of A . If x is a numerical variable (of any of the theories we consider), then x^* is just x ; in other words the need for starred variables arises only in connection with comprehension axioms. In HA° , too, x^* may be taken as x . For set variables, and for the operation variables in Feferman's theories, some care should be taken that the starred variables are distinct from the unstarred ones; this technicality is discussed in connection with realizability in [2].

The classes defining $p \Vdash A$ are as follows:

$$\begin{aligned}
 p \Vdash A \ \& \ B & \text{ is } p \Vdash A \ \& \ p \Vdash B \\
 p \Vdash A \ \vee \ B & \text{ is } p \Vdash A \ \vee \ p \Vdash B \\
 p \Vdash \exists x A & \text{ is } \exists x^* p \Vdash A \\
 p \Vdash A \rightarrow B & \text{ is } \forall q \leq p (q \Vdash A \rightarrow \exists r \leq q (r \Vdash B)) \\
 p \Vdash \forall x A & \text{ is } \forall x^* \forall q \leq p \exists r \leq q \ r \Vdash A .
 \end{aligned}$$

If there is more than one sort of variable, then we let the above quantifier clauses apply to each sort of variable. These clauses apply to any theory T ; to the the clauses for atomic T , we have to discuss each theory separately. We want $p \Vdash \underline{a}(\bar{n}) \cong \bar{m}$ iff $p(n) \cong m$; the object is to arrange it so this works and the rest of the axioms are forced. Since we are adding a partial function, we have to discuss how to talk about partial functions in HA° . We can do this

to a sufficient extent by talking about those partial functions whose graphs are enumerated by total functions. For instance, let $PF(f)$ be

$$\forall k, m, j (\langle k, j \rangle \varepsilon \text{ range } f \text{ and } \langle k, m \rangle \varepsilon \text{ range } f \longrightarrow j = m).$$

Then let $f(k) \cong j$ be $PF(f)$ and $\exists m (f(m) = \langle k, j \rangle)$. The theory Ta consists of T , the constant \underline{a} , and an axiom asserting that \underline{a} is a partial function from N to N ; in the case of HA^w , we use the formula PF to express this. It is easy to see that Ta is conservative over T .

Now we return to the problem of defining $p \Vdash A$ for A atomic. In HA^w and in Feferman's theories, we have an application relation $\text{App}(f, x, y)$ for an atomic formula. We set $p \Vdash \text{App}(f, x, y)$ to be $f^*(p, x) \cong y$, or more precisely, $f^*(\langle p, x \rangle) \cong y$, where $\langle p, x \rangle$ codes the pair p, x into a single object of type the same as the type of x (for HA^w ; such coding is unnecessary in Feferman's theories). Note that we are thus driven to discuss partially defined functional of finite type; but at least their arguments are total functionals, so that the same coding device as above can be used. This definition of $p \Vdash \text{App}(f, x, y)$ is supposed to apply whether f, x, y are variables or constants; so we have to give terms c^* for each constant c , in order to make the definition complete. In the case of Feferman's theories, we take \underline{a}^* so that $\underline{a}^*(p, m) = p(m)$. In the case of HA^w , we let \underline{a}^* be some term which will enumerate the graph of the finite function (coded by the sequence number) p . For the other constants c of Feferman's theories or of HA^w , we want c^* to ignore the arguments p , and do what c would do to the other arguments. In [2], where we have worked with a slightly different "uniform" version of forcing, we have given the terms c for Feferman's theory explicitly.

We define $p \Vdash x = y$ to be $x^* = y^*$. Remember that in certain cases, for instance if these are numerical variables, x^* is just x . This completes the definition of forcing for HA^w .

The other theories for which we are defining forcing have an atomic formula $x \varepsilon y$. We take $p \Vdash x \varepsilon y$ to be $\langle p, x^* \rangle \varepsilon y^*$. In the case of Friedman's theories, the partial function \underline{a} is going to be a set, so we define \underline{a}^* to be $\{\langle \langle p, x \rangle, y \rangle : p(x) \cong y\}$. For Feferman's theories, we have to explain N^* and c_ϕ^* . We want N^* to be the set of $\langle p, n \rangle$ such that $p \Vdash n \in N$. Just take N^* to be $C \times N$; in other words, $p \Vdash n \in N$ is just $n^* \in N$. We want $c_\phi(y)$ to be $\{\langle p, x \rangle^* : p \Vdash \phi(x, y)\}$, because $c_\phi(y)$ is supposed to be $\{x : \phi(x, y)\}$. This definition may appear to be circular, but it is not, because ϕ is an elementary formula in the comprehension axiom of Feferman's

theories. See [2] for a fuller discussion of why this is not circular. Also in the case of Friedman's set theories, there need to be certain constants to denote the sets whose existence is asserted by the axioms, and certain function symbols, and we have to give c^* for terms c built up in this way. This is carried out in detail in [3]; there the version of forcing is "uniform", but the constants c^* are the same.

REMARK. Forcing for weak intuitionistic set theories offers several technical complications, which we have considered at length in [3] and with which we do not wish to become involved here. Besides these complications, forcing works only for intuitionistic set theories without extensionality, at least in the present formulation. This creates no difficulty, since we have proved in [3] that the theories with extensionality are conservative over those without, for arithmetic theorems. The "technical complications" mentioned are the same for the version of forcing used here and that used in [3]; in particular, for theories without power set, the definition must be modified by restricting the range of the starred variables, etc., as in [3]. The reader specifically interested in these theories may consult [3]; others will find the ideas well illustrated in the cases of HA and Feferman's theories, where the technical complications are not so distracting. Note that even our application of these results to the conservativeness of \underline{B} over HA uses only results of this paper for Feferman's theories, not for intuitionistic set theories.

THEOREM 2.1. (*Soundness of forcing.*) *If $Ta \vdash A$ then $T \vdash \forall p \exists q \leq p(p \Vdash A)$. Here T may be any of the following theories:*

- (i) HA^w or HAS^w .
- (ii) Feferman's theory $BEM + CA$ of [2], or the corresponding theory with restricted induction discussed in [1].
- (iii) Any of Friedman's set theories discussed in [3] (except the weakest one \underline{B}) in the version without extensionality, and with the axiom of choice DC replaced by arithmetic HA . The forcing conditions C can be any definable set of finite partial functions.

REMARKS. The theory $BEM + CA$ consists of the theory called T_0 in [5], minus the join and inductive generation axioms and with decidable equality only for integers.

Proof. By induction on the length of the proof of A . For Feferman's and Friedman's systems, the proof is nearly identical to the proofs for the "uniform" forcing given in [2] and [3] respec-

tively; we do not give any details here. The verification that the axioms of HA^ω are forced is entirely straightforward. In connection with HAS^ω , and generally in connection with any theory having a comprehension axiom, the axiom guaranteeing the existence of $\{x: \phi(x)\}$ will be forced if we can form within the theory, the set $\{\langle p, x^* \rangle: p \Vdash \phi(x)\}$. Thus it is easier to deal with an unrestricted comprehension axiom like that of HAS^ω than with a restricted one as in Feferman's theories, where one has to worry whether the restricted comprehension axiom suffices to form the set one needs.

It is worth remarking that $\forall x \exists n \underline{a}(x) = n$ will be forced if and only if $\forall n \forall p \exists q \leq p (q(n) \text{ defined})$; since we cannot guarantee this for the forcing conditions C we will need to prove Goodman's theorem, we must add a *partial* function, even if it is more complicated to do so.

LEMMA 2.1. *If A is arithmetic, then $T \vdash (A \leftrightarrow p \Vdash A)$.*

Proof. By induction on the complexity of A .

3. Realizability revisited. Now we want to use functions recursive in the generic partial function \underline{a} for realizability. First note that relative recursiveness in a partial function means that we are using Turing machines with oracles which can ask for the values of the partial function \underline{a} ; if an answer is forthcoming, the computation continues. If not, the machine "waits forever" for the oracle to answer. Thus the graph of \underline{a} may not be recursive in \underline{a} . (This is OK.)

Consider the case T is HA^ω , which is the simplest case. Before, we wrote erA for $er(A)_{HRO}$, where the realizability on the right is Kleene's (formalized) 1945-realizability. (See [12] for this and for the definition of HRO .) Now we define $HROa$ analogously to HRO , using the functions recursive in \underline{a} , and write erA for $er(A)_{HROa}$, where now the realizability on the right also uses functions recursive in \underline{a} . That is, everything is just as before, reading "recursive in \underline{a} " for "recursive". Then we have the soundness theorem.

THEOREM 3.1. *$Ta + AC \vdash A$ implies $Ta \vdash \bar{e}rA$ for some e ; here T is HA^ω .*

Proof. By induction on the length of the proof of A ; the proof is obtained from the corresponding proof for 1945-realizability by reading "recursive in \underline{a} " for "recursive". See [12] for details.

4. How to prove Goodman's theorem. The key to Goodman's

theorem is the following lemma:

LEMMA 4.1. *Fix an arithmetic sentence A . Then there is a set C of forcing conditions (definable in HA) such that $HA^o \vdash \forall p \exists q \leq p (q \Vdash (erA \rightarrow A))$.*

Proof. If y stands for a list y_1, \dots, y_n , let \mathbf{y} be $\langle y_1, \dots, y_n \rangle$. Let C consist of all finite functions p such that for all subformulae B and D of A (including A as a subformula of itself), we have

$$(1) \quad \exists x B(x, \mathbf{y}) \text{ and } p(\langle \exists x B', \mathbf{y} \rangle) \text{ defined } \rightarrow B(p(\langle \exists x B', \mathbf{y} \rangle) \mathbf{y})$$

$$(2) \quad B(\mathbf{y}) \vee D(\mathbf{y}) \text{ and } p(\langle B \vee D', \mathbf{y} \rangle) \text{ defined } \rightarrow (p(\langle B \vee D', \mathbf{y} \rangle) = 0 \ \& \ B(\mathbf{y})) \vee (p(\langle B \vee D', \mathbf{y} \rangle) = 1 \ \& \ D(\mathbf{y})).$$

Then we define for each subformula B of A a number j_B such that

$$(i) \quad T \vdash (\forall p \exists q \leq p \ q \Vdash (B(\mathbf{y}) \rightarrow \{j_B\}^a(\mathbf{y})rB))$$

$$(ii) \quad T \vdash (\forall p \exists q \leq p \ q \Vdash (urB \rightarrow B)).$$

The definition of j_B is as follows:

j_A for A prime is an index of the identically zero function.

$$\begin{aligned} \{j_{B \rightarrow C}\}^a(\mathbf{y}) &= \{j_C\}^a(\mathbf{y}) \\ \{j_{B \ \& \ C}\}^a(\mathbf{y}) &= \langle \{j_B\}^a(\mathbf{y}), \{j_C\}^a(\mathbf{y}) \rangle \\ \{\{j_{\forall x B}\}^a(\mathbf{y})\}^a(x) &= \{j_B\}^a(x, \mathbf{y}) \\ \{j_{\exists x B}\}^a(\mathbf{y}) &= \langle \underline{a}(\langle \exists x B', \mathbf{y} \rangle), \{j_B\}^a(\underline{a}(\langle \exists x B', \mathbf{y} \rangle), \mathbf{y}) \rangle \\ \{j_{B \vee D}\}^a(\mathbf{y}) &= \begin{cases} \{j_B\}^a(\mathbf{y}) & \text{if } \underline{a}(\langle B \vee D', \mathbf{y} \rangle) = 0 \\ \{j_D\}^a(\mathbf{y}) & \text{if } \underline{a}(\langle B \vee D', \mathbf{y} \rangle) = 1. \end{cases} \end{aligned}$$

After this definition, we can prove (i) and (ii) by a straightforward simultaneous induction. Only the cases for disjunction and existential quantification involve anything new. We treat existential quantification. Suppose B is $\exists x D$. We prove (i) first. Suppose p is given; we will show $p \Vdash (B(\mathbf{y}) \rightarrow \{j_B\}^a(\mathbf{y})rB)$. Let $q \leq p$ have $q \Vdash B(\mathbf{y})$, i.e., $q \Vdash \exists x D(x, \mathbf{y})$. Then $\exists x q \Vdash D(x, \mathbf{y})$, since x^* is x for number variables. By Lemma 2.1, ($q \Vdash D(x, \mathbf{y}) \leftrightarrow D(x, \mathbf{y})$); hence $D(x, \mathbf{y})$. By (i), q has an extension r_0 with $r_0 \Vdash (D(x, \mathbf{y}) \rightarrow \{j_D\}^a(x, \mathbf{y})rD(x, \mathbf{y}))$. If this r_0 is not already defined at $\langle B', \mathbf{y} \rangle$, we can extend r_0 to a condition r by defining $r(\langle B', \mathbf{y} \rangle) = x$. (Since $D(x, \mathbf{y})$ holds, r is in C .) Otherwise replace x by $r_0(\langle B', \mathbf{y} \rangle)$; since r_0 is in C , we have $D(x, \mathbf{y})$. Thus $r \Vdash x = \underline{a}(\langle B', \mathbf{y} \rangle)$. Since $q \Vdash D(x, \mathbf{y})$, also $r \Vdash D(x, \mathbf{y})$, so by the property of r_0 , some extension r_1 of r has $r_1 \Vdash \{j_D\}^a(x, \mathbf{y})rD(x, \mathbf{y})$. Since $r \Vdash x = \underline{a}(\langle B', \mathbf{y} \rangle)$, also $r_1 \Vdash x = \underline{a}(\langle B', \mathbf{y} \rangle)$. Hence $r_1 \Vdash x = \underline{a}(\langle B', \mathbf{y} \rangle) \ \& \ \{j_D\}^a(x, \mathbf{y})rD(x, \mathbf{y})$. That is, $r_1 \Vdash \exists x (x = \underline{a}(\langle B', \mathbf{y} \rangle) \ \& \ \{j_D\}^a(x, \mathbf{y})rD(x, \mathbf{y}))$. By definition of $j_B = j_{\exists x D}$, this is $r_1 \Vdash \{j_B\}^a(\mathbf{y})rB(\mathbf{y})$, and (i) is proved for B .

To prove (ii) in case B is $\exists x D$, we let p be given, and prove $p \Vdash (urB \rightarrow B)$. If $q \leq p$ has $q \Vdash urB$, then $q \Vdash (u)_1 rD((u)_0, \mathbf{y})$, so by

induction hypothesis, q has an extension r with $r \Vdash D((u)_0, y)$; hence $r \Vdash \exists x D(x, y)$, that is, $r \Vdash B(x, y)$. This proves (ii) for the case of existential quantification. Disjunction may be treated similarly; this completes the proof of Lemma 4.1.

THEOREM 4.1. (*Goodman's theorem.*) *If $HA^\omega + AC \vdash A$, for A arithmetic, then $HA \vdash A$.*

Proof. Let T be HA^ω . Fix A arithmetic with $T + AC \vdash A$. By Theorem 3.1, $Ta \vdash \bar{e}rA$ for some e . By Theorem 2.1, $T \vdash \exists p(p \Vdash \bar{e}rA)$, where the forcing conditions C are chosen as above, depending on the formula A . By Lemma 4.1, $T \vdash \forall p \exists q \leq p(q \Vdash (erA \rightarrow A))$. Hence $T \vdash \exists q(q \Vdash A)$. Hence, by Lemma 2.1, $T \vdash A$. This proves the theorem.

REMARK. In [8], Goodman states a generalization of the theorem to any theory T in the language of HA^ω for which his interpretation is sound. Correspondingly, our proof clearly works in case realizability and forcing are both sound for T . In fact, in our case there is no necessity to restrict the language of T in advance. In the rest of the paper we shall exploit this situation.

5. Friedman's problem. Problem 38 on Friedman's list of "102 Problems in Mathematical Logic" asks for the extension of Goodman's theorem to the case when T is $HA^\omega + \text{extensionality}$. This theory is just HA^ω with the additional axiom

(extensionality) $\forall x(f(x) = g(x)) \longrightarrow f = g$ (at all sensible types).

Previous methods, including apparently those of [8], have not sufficed to solve this problem. Indeed, there might seem to be grounds for suspicion that the theorem might not extend to this case: for $T + AC$ refutes Church's thesis in the form $\forall f \exists e \forall x(\{e\}(x) = f(x))$ [1]. Perhaps it might also refute the weaker form, $\forall x \exists y A(x, y) \rightarrow \exists e \forall x A(x, \{e\}(x))$, with A arithmetic, which would show that Goodman's theorem does not extend. Let us begin by asking where our proof of Goodman's theorem breaks down, if we try to apply it to T . The problem is, that $(\text{ext})_{HRO}$ is not realized. We need a realizability for T such that

- (i) we can relativize it to any function \underline{a}
- (ii) we can, after choosing a suitable generic \underline{a} , get a fixed arithmetic formula to be "self-realizing"
- (iii) extensionality and AC are both realized.

It turns out that such a realizability can be found, and Goodman's theorem does extend to T . Rather than just pull this realizability out of a hat, though, it seems better to discuss the background.

Consider the following realizabilities: For HA we have Kleene's recursive realizability, where the realizing objects are numbers, representing partial recursive functions. For HA° , we have Kreisel's modified realizability, which is not a single notion, but an abstract interpretation, assigning a formula $A_0(f, x)$ to each $A(x)$; one can think of $A_0(f, x)$ as " f realizes $A(x)$ ". But to obtain a specific notion of realizability, one must follow Kreisel's interpretation with an interpretation of HA° in HA , such as HRO or HEO (see [12]). Alternately, one could first take the HRO or HEO interpretation, then use Kleene's realizability. (This was what we used for Goodman's theorem.) The main difference between Kreisel's realizabilities and Kleene's is that Kreisel's realizing objects are functionals of finite type (or their interpretations in HRO or HEO); that is, they are hereditarily total (on the proper type). Examination of the proof of Lemma 4.1 will convince one that this feature renders them unsuitable for our purposes: requirement (ii) will never be met. On the other hand, if we first interpret HA° in HA , then use Kleene's realizability, we will have trouble with (iii). If we use HRO , we will not get extensionality. If we use HEO , we will get extensionality, but $(AC)_{HEO}$ is not realized. It seems, if we want to get AC realized, we have to do realizability directly on HA° . But in order to meet (ii), we have to use partial objects. Which brings us to a gap in the literature that could have been filled long ago: solve the proportion,

abstract modified realizability:	
	specific "total realizabilities" in HRO, HEO
	as
X : Kleene's recursive realizability.	

Clearly " X " is an abstract interpretation of HA° in a suitable theory of partial functionals; then Kleene's realizability will result by an interpretation analogous to HRO . And, the very realizability which we need will result by an interpretation analogous to HEO .

The details of this program are so similar to what is known already that we give only a sketch. First we describe a suitable theory HA^p of partial functionals. These are to be partial functionals of partial arguments, not only partial functionals of total arguments. We have the same type structure as for HA° . We have an application relation $\text{App}(f, x, y)$ at each sensible triple of types, which is meant to stand for $f(x) \cong y$. So we need the axiom $\text{App}(f, x, y)$ and $\text{App}(f, x, z) \rightarrow x = z$. Exactly as in Feferman's theories, we introduce "application terms", or for short, terms, which are not officially part of the language. These terms are built up from variables and constants; we include the same constants as in

HA^ω . A list of abbreviations for use in connection with application terms is in [5] or [1]; in particular, we can make sense of $\phi(t)$ where t is a term and ϕ a formula. The axioms of HA^P include the full schema of induction, and the defining axioms for the constants, including a constant for recursion.

Next we note that the total types can be defined in HA^P . Namely, for each total type σ there is a formula $TOT_\sigma(f)$, defined by, $TOT_0(x)$ is $x=x$, where x is a type-0 variable; generally $TOT_\sigma(x)$ has x a free type σ variable. $TOT_{(\sigma,\varepsilon)}(f)$ is $\forall x(TOT_\sigma(x) \rightarrow TOT_\zeta(f(x)))$. (So in accordance with our abbreviations, $f(x)$ is defined and is total of type ζ for each x of type σ). Thus in a natural sense HA^P contains HA^ω (one has to prove the constants of HA^ω are hereditarily total). In other words, there is an interpretation of HA^ω in HA^P , such that the interpretation A^* of each theorem A of HA^ω is a theorem of HA^P .

Next we give an abstract realizability interpretation of HA^P in itself. We assign to each formula A of HA^P , a formula $A^0 = \exists f A_0(f, x)$, where x are the free variables of A , and f may be a finite list of variables. (Think of $A_0(f, x)$ as “ f realizes $A(x)$ ”.) The clauses defining A^0 are the same as for Kreisel’s modified realizability, only now the meaning is different. For instance, if A^0 is $\exists f A_0(f, x)$ and B^0 is $\exists g B_0(g, x)$, then $(A \rightarrow B)^0$ is $\exists F \forall f (A_0(f, x) \rightarrow B_0(F(f), x))$; but now, according to our conventions about application terms, $F(f)$ does not have to be defined unless $A_0(f, x)$; this imparts the desired Kleene flavor to our interpretation. Composing this interpretation with the interpretation of HA^ω in HA^P , we get a realizability interpretation of HA^ω in HA^P . We summarize all this in a theorem:

THEOREM 5.1. *If $HA^P \vdash A$, then for some terms t , $HA^P \vdash A_0(t, x)$. Moreover, if $HA^\omega + AC \vdash A$, then for some terms t , $HA^P \vdash ((A^*)_0(t, x)$.*

Proof. The first part is standard. The second part follows, if we show that $(AC)^*$ is provably realized. Now an instance of $(AC)^*$ is

$$\begin{aligned} \forall x(TOT_\sigma(x) \longrightarrow \exists y(TOT_\zeta(y) \ \& \ \phi(x, y))) \longrightarrow \\ \exists f(TOT_{(\sigma,\varepsilon)}(f) \ \& \ \forall x(TOT_\sigma(x) \longrightarrow \phi(x, f(x)))) . \end{aligned}$$

It is easy to see that we can get this realized, if we show that for each type σ , TOT_σ is a “self-realizing” formula. We say a formula B is self-realizing if there is a term t_B such that

- (1) $HA^P \vdash (B(x) \rightarrow B_0(t_B(x), x))$
- (2) $HA^P \vdash (B_0(f, x) \rightarrow B(x))$.

It is easy to prove by induction on B that if B is negative (written

with application terms allowed) then B is self-realizing. Now $TOT_{(\sigma, \zeta)}$ can be written with application term $f(x)$ as $\forall x(TOT_{\sigma}(x) \rightarrow TOT_{\zeta}(f(x)))$, and so it is negative, by induction on the types, when application terms are allowed. Hence $(AC)^*$ is provably realized, and our sketch of the proof of Theorem 5.1 is complete.

Now we indicate how to recover Kleene's 1945-realizability. Interpret all the variables, of whatever type, to range over the integers; and let all the App relations be interpreted as $\{e\}(x) \cong y$. If A is a formula of arithmetic, then $(A)_0$, so interpreted, will be provably equivalent to erA , where r is 1945-realizability.

The next part of our program is to give a model of HA^P in HA , in which $(ext)^*$ is satisfied. This turns out to be easier suggested than done. Let us examine the obvious approach and see what is wrong with it. The "obvious approach" is to define for each type σ , a set M_{σ} of partial effective operations, and an extensional-equality relation on M_{σ} , and then interpret the type σ variables as ranging over M_{σ} , and type σ equality as the extensional equality relation. Then we try to check $(ext)^*$: suppose F and G are of type 2 in this model, i.e., are in M_2 . What happens if $F(f) = G(f)$ for every total, type 1 f ; but F and G disagree on some partial f ? Then F and G will not be set extensionally equal, but they should be. Now as a matter of fact, Myhill-Shepherdson's theorem [11, p. 359] shows that this cannot happen; but this theorem is not provable in HA , as shown in [4].

The solution to this problem is as follows: First we define the total effective operations E_{σ} ; any integer is in E_0 , and $n \sim_0 m$ iff $n = m$; then e is in $E_{(\sigma, \zeta)}$ if $\forall x \in E_{\sigma}(\{e\}(x)$ is defined and in $E_{\zeta})$; and $a \sim_{(\sigma, \zeta)} b$ if whenever x and y are in E_{σ} , with $x \sim_{\sigma} y$, we have $\{a\}(x) \sim_{\zeta} \{b\}(y)$. (This is just a version of *HEO*.) Next, we define our model of HA^P by, $M_0(n)$ is $n = n$; $M_{(\sigma, \zeta)}(e)$ iff $\forall x(M_{\sigma}(x)$ and $\{e\}(x)$ defined $\rightarrow M_{\zeta}(\{e\}(x))$) and $\forall x, y(E_{\sigma}(x)$ and $E_{\sigma}(y)$ and $x \sim_{\sigma} y$ and $\{e\}(x)$ defined $\rightarrow (\{e\}(y)$ defined and $\{e\}(x) \sim_{\zeta} \{e\}(y))$. Now it is easy to prove that the interpretation $*$ from HA^{ω} to HA^P , followed by this model, takes the type σ objects onto E_{σ} . That is, the total objects in M_{σ} are just E_{σ} . Note carefully that the relation $a \sim_{(\sigma, \zeta)} b$ is defined even if a or b is not in $E_{(\sigma, \zeta)}$; in that case the definition requires (according to our conventions) that for x in E_{σ} , if $\{a\}(x)$ is defined, so is $\{b\}(x)$. Thus the second part of the definition of $M_{(\sigma, \zeta)}$ makes good sense. (To understand the definition; consider a few examples at low types.) A straightforward use of the recursion theorem (to get a model of the recursion operator in HA^P) shows that the M_{σ} do indeed form a model of $HA^P + (ext)^*$, when equality at higher types is interpreted by means of the relations \sim_{σ} .

For the rest of this section, T will be $HA^{\omega} + ext$. Let Ta be

formed as before, by adding a constant \underline{a} for a partial function from integers to integers. We can form the theory $HA^P a$ even more naturally, letting \underline{a} stand for a partial function (without coding). Ta has a natural interpretation in $HA^P a$, which we denote by $*$ since it extends the old $*$. Next we relativize the model M to \underline{a} , by replacing $\{e\}(x)$ by $\{e\}^{\underline{a}}(x)$ throughout. Of course, now the interpretation goes, not into HA , but back into $HA^{\omega} a$ (which is conservative over HA .) Now we write erB to abbreviate $((B^*)_o)_M(e)$.

LEMMA 5.1. *If $T + AC \vdash B$, then $HA^{\omega} a \vdash \bar{e}rB$ for some e .*

Proof. Suppose $T + AC \vdash B$. By Theorem 5.1, HA^P proves $(E \rightarrow B)_o^*(t)$ for some application term t ; here E is a conjunction of instances of ext . But E is equivalent to its realizability interpretation, since E is negative. (To be precise: E^* is negative, when written with application terms. Such a formula is equivalent to its interpretation.) Hence, for some term s , HA^P proves $E \rightarrow (B_o^*(s))$. Next we note, by induction on the complexity of application terms s , that there is for each term s a number e such that $HA^{\omega} a$ proves that \bar{e} plays the role of s in the model M (which is now relativized to \underline{a} , remember.) That is, the number which represents s in the model does not depend on \underline{a} . Hence $(B_o^*(s))_M$ is just $(B_o^*)_M(\bar{e})$. Since E_M^* is provable, we have $HA^{\omega} a \vdash (B_o^*)_M(\bar{e})$, that is, $HA^{\omega} a \vdash erB$. This proves the lemma.

LEMMA 5.2. *(The analog of Lemma 4.1.) Fix an arithmetic A . Then there is a set C of forcing conditions (definable in HA^{ω}) such that $HA^{\omega} a \perp \forall p \exists q \leq pq \Vdash (erA \rightarrow A)$.*

Proof. Take the same C as in Lemma 4.1. Moreover, take the same numbers j_B . Then we claim (i) and (ii) in the proof of Lemma 4.1 still hold, with the new meaning of realizability. For this only one thing needs to be added to the proof of Lemma 4.1: we have to check that these numbers belong to M_s of the appropriate type, i.e., that they are extensional. This is an easy induction, noting for the basis case that only the values of \underline{a} are used in the definitions of the j_B .

Now we are ready to solve Friedman's problem.

THEOREM 5.2. *Let A be arithmetic. If $HA^{\omega} + AC + ext \vdash A$, then $HA^{\omega} \vdash A$.*

Proof. Exactly as for Theorem 4.1, using our new extensional

realizability: Fix A , with $T + AC \vdash A$, where as above T is $HA^\omega +$ extensionality. By Lemma 5.1, for some e , we have $HA^\omega a \vdash \bar{e}rA$. By Lemma 5.2, we can choose forcing conditions so that $HA^\omega a \vdash \forall p \exists q \leq p (q \Vdash (erA \rightarrow A))$. By Theorem 2.1, $HA^\omega \vdash \exists q (q \Vdash A)$. So, by Lemma 2.1, $HA^\omega \vdash A$. Since HA^ω is conservative over HA , $HA \vdash A$. This completes the proof.

REMARKS. There are several other interesting models of HA^r , including the “effective operation” version of Platek’s “hereditarily consistent” functionals. In Platek’s dissertation, the fundamental role of partial functionals in recursion theory is made clear — it seems an historical accident that total functionals have had the limelight.

6. Results for stronger theories. We first consider the extension of Goodman’s theorem to the theory of species, and then some Goodman-style theorems for Feferman’s and Friedman’s theories.

THEOREM 6.1. *$HAS^\omega + AC$ is conservative over HAS for arithmetic theorems.*

Proof. All we have to do is show that realizability and forcing are sound for HAS^ω . First, we extend the HRO model from HA^ω to HAS^ω . Species of type σ objects get interpreted as species of numbers in HRO_σ ; thus the HRO interpretation goes from HAS^ω to HAS . Now Troelstra has given a realizability for HAS in [12]; so we can write erA for $er(A)_{HRO}$, where the realizability on the right is Troelstra’s. Again, $(AC)_{HRO}$ is realized; the essential point here is that the finite types in HRO are defined by self-realizing formulae. And, as before, this realizability can be relativized to any partial function \underline{a} . We have already proved the soundness of forcing for HAS^ω ; and now the proof can be completed exactly like the proof of Theorem 4.1.

We next give Goodman-style theorems for Feferman’s and Friedman’s theories. These theorems say that the axiom of dependent choices does not enable us to prove any more arithmetic theorems than we can prove without it. In [1], this result is applied to show that certain of these theories are actually conservative over HA , a result of some philosophical significance considering the vast amount of mathematical practice that can be formalized in these systems.

Modulo the work on forcing and realizability which has already been done for these systems in [2] and [3], these theorems are easy. This is one of the main advantages of the present method of proof: we use familiar tools instead of inventing a new one.

We now discuss the extension of Goodman's theorem to Friedman's theories. A little care is necessary here: if we delete *DC* from Friedman's theories, we cannot prove the existence of any functions, not even plus and times on the integers. The proper result is as follows:

THEOREM 6.2. *Let T be any of the intuitionistic set theories discussed in [3], except the weakest one \underline{B} . If we delete *DC* from T , but add back the axioms of *HA*, then no arithmetic theorems are lost.*

REMARK. Thus T can be Friedman's T_1, T_2 (which is essentially Myhill's *CST*), T_3, T_4 , Zermelo set theory, or intuitionistic *ZF*. The theorem is true for \underline{B} too, in view of the theorem of [1] that the arithmetic theorems of \underline{B} are exactly those of *HA*. The present proof does not apply to \underline{B} directly since we do not know that forcing is sound for \underline{B} . The proof of [1] makes use of Goodman-style theorems for Feferman's theories (see below) together with an interpretation of \underline{B} in Feferman's theories.

Proof. In [3] is proved that the axiom of extensionality can be deleted from T without loss of arithmetic theorems (if certain minor modifications are also made to the other axioms of T). We can therefore assume that T is one of the nonextensional set theories discussed in [3]. Then forcing and realizability are both sound for T , as proved in [3]. The same proof shows the soundness of realizability using functions recursive in \underline{a} for Ta . Then the proof goes through exactly as for Theorem 4.1.

Next we extend Goodman's theorem to certain theories in the language of Feferman's systems. Our proof will work for any theory T in this language, and any axiom of choice C , such that (i) T proves $C + T$ is realized, and (ii) forcing is sound for T . In [1] we apply the theorem to an unconventional axiom of choice C , which we can show is realized; this is an essential step in our proof that \underline{B} is conservative over *HA*. From results of [2] discussed earlier in this paper, we know for example that these conditions hold for $C =$ dependent choices and $T = BEM + CA$.

THEOREM 6.3. *Let T be a theory in Feferman's language and C an axiom of choice, such that conditions (i) and (ii) are satisfied. Then $T + C$ is conservative over T for arithmetic theorems.*

Proof. There is an obstacle here, in that realizability for T is not "recursive realizability"; the formula erA has a variable e for

an operation, not a number. Hence it is not clear how to define a “relativized realizability”, relative to a generic function \underline{a} . There are at least two ways around this obstacle, one of which is as follows. Let M be a model for T , constructed in the way that Feferman’s model in [5] or [2] is constructed, by using functions recursive in \underline{a} instead of recursive functions. It is not hard (except in case T is EMN , in which case it is still possible) to formalize this model in Ta , so that $M \Vdash A$ is a formula of Ta , for each formula A of Ta ; all the free variables of $M \Vdash A$ are number variables. (The details of this formalization are actually carried out in [2].) Now $M \Vdash erA$ is a formula with a number variable e ; and we can use this formula in place of erA to carry out the argument for Goodman’s theorem. We have to check that $Ta \vdash A$ implies $Ta \vdash (M \Vdash A)$; this is done for T in place of Ta in [2], as mentioned; the same proof works for Ta . Hence $T + C \vdash A$ implies $Ta \vdash (M \Vdash erA)$. Also, since the integers of M are standard, we have $T \vdash (A \leftrightarrow M \Vdash A)$ for arithmetic A . Now the proof is straightforward: Suppose $T + C \vdash A$, with A arithmetic. Then $T \vdash (M \Vdash \bar{e}rA)$, for some e ; then $T \vdash \exists p(p \Vdash (M \Vdash \bar{e}rA))$, for forcing conditions C chosen to make $T \vdash \forall p \exists q \leq p(q \Vdash (M \Vdash erA \rightarrow M \Vdash A))$. Hence $T \vdash \exists q q \Vdash (M \Vdash A)$. Since A is arithmetic we have $T \vdash ((M \Vdash A) \leftrightarrow A)$. Hence $T \vdash \exists q(q \Vdash A)$. Hence $T \vdash A$, by Lemma 2.1. This completes the proof of the theorem.

Open Problem. Is $\underline{B} + AC$ conservative over \underline{B} for arithmetic sentences? Here one must read AC as defined in this paper, that is the axiom of choice at all finite types. It is known that the full set-theoretic axiom of choice implies the law of the excluded middle in \underline{B} (see [1] for a general discussion of axioms of choice in intuitionistic theories). It is worth explaining why the technique of Theorem 6.2 fails to solve the open problem. It is because we cannot get rid of extensionality, as we can in the case of DC . In the case of DC , we have the work of [3], where a model of \underline{B} is given in \underline{B} minus extensionality, using sets of rank $< \omega + \omega$, with defined relations of extensional equality. If we try to verify AC in this model, we can apply AC to get a choice function, but there is no guarantee that the choice function will be extensional. With DC , this problem does not arise, since the function will have domain the integers.

REFERENCES

1. M. Beeson, *Extensionality and choice in constructive mathematics*, to appear in Pacific J. Math.
2. ———, *Principles of continuous choice and continuity of functions in formal*

- systems for constructive mathematics*, *Annals of Math. Logic*, **12** (1977), 249–322.
3. M. Beeson, *Continuity in intuitionistic set theories*, to appear in *Proceedings of the Logic Colloquium at Mons, August 1978*, North-Holland *Studies in Logic* series.
 4. ———, *The underderivability in intuitionistic formal systems of theorems on the continuity of effective operations*, *J. Symbolic Logic*, **40** (1975), 321–346.
 5. S. Feferman, *A language and axioms for explicit mathematics*, In *Algebra and Logic*, Springer Lecture Notes No. 450, 87–139.
 6. H. Friedman, *Set theoretic foundations for constructive analysis*, *Annals of Math.*, **105** (1977), 1–28.
 7. N. Goodman, *The theory of the Gödel functionals*, *J. Symbolic Logic*, **41** (1976), 574–583.
 8. ———, *Relativized realizability in intuitionistic arithmetic of all finite types*, *J. Symbolic Logic*, **43** (1978), 23–45.
 9. G. E. Minc, *Finite investigation of infinite derivations* (Russian, with English summary), *Zapiski naūchyh seminarov LOMI*, **49** (1975), 67–121.
 10. R. Platek, *Foundations of recursion theory*, Stanford dissertation, 1966.
 11. H. Rogers, *Theory of Recursive Functions and Effective Computability*, McGraw-Hill, 1967.
 12. A. S. Troelstra, *Metamathematical investigation of intuitionistic arithmetic and arithmetic of finite types*, Springer Lecture Notes No. 344.

Received March 30, 1977.

INSTITÜT FÜR MATHEMATIK
 DER UNIVERSITÄT BONN
 WEGELERSTRASSE 6
 53 BONN, GERMANY

Current address: Mathematics Instituut
 Rijksuniversiteit Utrecht
 Budapestlaan 6
 3508 TA Utrecht
 Netherlands