LOGIC COLLOQUIUM 78
M. Boffa, D. van Dalen, K. McAloon (eds.)

CONSTRUCTIVE THEORIES OF FUNCTIONS AND CLASSES[1)]

SOLOMON FEFERMAN

Dedicated to the memory of my friend and colleague, Karel de Leeuw

Introduction and contents. These lectures were designed to acquaint a general logical audience with basic features of Bishop's approach to constructive mathematics (BCM) and with work on a certain formal system T_0 in which that can be represented. Several competing and rather different systems have been proposed for the same purpose. Thus, in addition to the intrinsic interest of the subject BCM provides an excellent case study for the process of formalization.

The contents are divided into five parts, only the last of which assumes some prior background; in outline they are as follows.

I. Background and aims. Part I gives an informal introduction to BCM which contrasts it both with everyday non-constructive mathematics as well as with the schools of constructivity previously established by Brouwer and Markov. Towards the end of this part we discuss general criteria of formalization, involving questions of adequacy and accord with the informal body of mathematics being represented.

II. The theory T_0. In part II we present the language and axioms of T_0 and some natural subsystems and extensions. The adequacy of T_0 to BCM is sketched and the question of its accord is discussed. Alternative formal systems proposed by Martin-Löf and Myhill are briefly compared in this connection.

III. Models. A variety of models (in the classical sense of the word) are presented for T_0 and related theories. One main purpose which these serve is to show how developments in BCM, when formalized in T_0, generalize corresponding parts of classical mathematics and certain recursion-theoretic analogues. They are also used to obtain consistency and independence results for some statements of mathematical interest.

IV. Realizability interpretations. In contrast to models, the method of realizability (originating with Kleene) is distinctively associated with interpretations

1) Text of lectures presented at Logic Colloquium 78 (University of Mons, Belgium, August 24-September 1, 1978). I am indebted to the organizers of this conference for proposing the lectures and for their helpful assistance in many ways. Research for and preparation of the text were supported in part by National Science Foundation grant MCS76-07163-A01.

of constructive theories. It is here adapted to the formalism of T_0 so as to obtain more delicate consistency (and conservation) results, in particular as concern axioms of choice and continuity principles.

V. Relations with subsystems of analysis. In this part one combines both proof-theoretical and model-theoretical methods to obtain equivalence (in strength) of various subsystems of the classical system $S = (\Sigma^1_2\text{-AC}) + (BI)$ with subsystems of T_0. For the full system T_0 one has an interpretation in S, but it is an open question whether S is equivalent to T_0.

We concentrate throughout on explanation and statement of results. Proofs are not given but some proof-ideas are indicated. The basic source is Feferman 1975; this has been enriched considerably by the work of Beeson 1977. The latter gives both models and realizability interpretations which are used particularly for continuity principles; his work is described within Parts III and IV. Important contributions to Part V have been made by Aczel, Buchholz, Friedman, Pohlers, and Sieg; detailed references are given in the text. Otherwise we draw principally on the unpublished notes Feferman 1976a, 1976b, and 1976c, which are now largely incorporated in the following.

For the reader seeking a general introduction to the subject of constructivity and its formalizations (especially stemming from the schools of Brouwer and Markov) I would suggest the excellent survey article Troelstra 1977a; this contains an extensive bibliography.

I. Background and aims

1. Ad hoc (local) vs. systematic (global) constructive mathematics. At the local level one deals with particular questions of construction without regard to general principles or methods. Frequently one knows an existential result guaranteeing the existence of a solution to a specific mathematical problem without knowing how it may be calculated, represented, or constructed. One then seeks to produce an explicit solution to the problem. For examples familiar to logicians we have: (i) decidability of p-adic fields (first existence by Ax and Kochen, followed by a concrete decision procedure by Cohen) and (ii) representability of positive definite real polynomials as sums of squares of rational functions (existence by Artin, followed by recursive representations by Robinson and primitive recursive representations by Kreisel).

At the global, systematic level one reconstructs whole portions of mathematics using entirely constructive notions and methods. One of the main reasons advanced for doing this is philosophical; it is based on a conception of mathematics which is opposed to the current underlying platonistic conception and has its source in human thought and constructions. Such systematic redevelopment according to constructive principles was initiated by Brouwer and carried on by Heyting and his students. Subsequently another school of constructivity was developed in Russia by Markov and Shanin (cf. Troelstra 1977a for references on these two schools). Finally the approach (here labeled BCM) was initiated in Bishop 1967 and continued by him and his students. The main features of the first two schools will be described briefly below and those of BCM will be described at length.

2. Constructivity in principle and constructivity in practice (feasibility). No matter how a constructive result is obtained (locally or globally) there is a question of its actual computation or execution. In this respect, even constructive existence results have a non-concrete character. A classical example is provided by Gauss' characterization of the regular polygons which are constructible by ruler and compass; the general theory had to be refined in order to give a feasible construction even of the 17-sided regular polygon. For a (negative) example familiar to logicians, we may mention Tarski's primitive recursive decision procedure for the theory of reals. It has been shown by Fischer and Rabin that any decision procedure for the reals requires exponential time and so is unfeasible by present computational methods.

3. General features of the platonistic conception. We describe these for a point-by-point comparison with the constructivist conception in §4.

3.1. Mathematical entities. These are conceived to be external to us and independent of our thoughts and constructions. In its modern form, the most general mathematical entities are sets and functions (which are interchangeable, cf. 3.3 below). Thus the platonist conception is also called the Cantorian set-theoretical conception of mathematics.

3.2. Mathematical statements are true or false. Hence the logic employed is the classical predicate calculus based on 2-valued semantics. The law of excluded middle $\phi \vee \neg \phi$ leads us to conclude $\exists x\, \psi(x) \vee \forall x \neg \psi(x)$. Thus to prove $\exists x\, \psi(x)$ it is sufficient to prove $\neg \forall x \neg \psi(x)$. This is the basis of the use of the indirect method to obtain existential results: assume $\forall x \neg \psi(x)$ and draw a contradiction. Evidently there is no explicit solution provided by such arguments.

3.3. Interchangeability of sets and functions. Of course the former are reduced to the latter via characteristic functions. Conversely, functions are regarded as many-one relations, which in turn are certain sets of ordered pairs. But the latter are definable as sets, so functions are reduced to sets.

3.4. Extensionality. The principle $\forall x\ (x \in A \leftrightarrow x \in B) \rightarrow A = B$ for sets A,B is justified by the consideration that sets exist independently of us and of any means of definition. Sets, then, may only be distinguished by their members.

3.5. Power set. Since arbitrary subsets of a set are supposed to exist independently and permanently, we may speak of their totality $\mathcal{P}(A)$. This operation may be iterated, leading to the finite-type hierarchy. For transfinite iteration the operation $A \mapsto A \cup \mathcal{P}(A)$ gives a more convenient theory (the cumulative hierarchy).

3.6. Subset formation. Any property $\phi(x)$ of elements of A determines a subset $B = \{x \in A \mid \phi(x)\}$ (separation or comprehension principle). ϕ may contain quantified variables ranging over other sets, in particular over $\mathcal{P}(A)$. Such comprehension principles are impredicative: B is defined in terms of the totality $\mathcal{P}(A)$, which contains B as an element.

3.7. The axiom of choice is usually agreed to be correct on the Cantorian view, since there is no question as to how the choices are to be effected. Then one has the well-ordering theorem and the theory of finite and transfinite cardinals. In consequence, such statements as the continuum hypothesis are taken to have a definite truth-value, though undecided by all set-theoretical principles so far recognized to be correct (or even having some plausibility).

4. General features of the constructivist conception.[2)]

4.1. Mathematical entities are only those which are understood directly by humans or obtained from such by successive human constructions (e.g., by combination into pairs or sequences). The natural numbers 0,1,2,... (denoted as a whole by $\mathbb{N}$) form basic entities which are generated by repeated adjunction of a single unit. Both the processes of construction of mathematical entities and of recognition of their properties are mental activities. Such recognition is the result either of direct intuition or of proofs based on principles inherent in the specific nature of the constructions used. For example, the principle of induction for $\mathbb{N}$ directly follows the manner of its generation.

4.2. Mathematical statements do not communicate questions of truth or falsity; they can only be assertions which communicate results of completed proofs. The use of the logical particles is explained in terms of constructions and proofs, roughly as follows:
(i) a proof of $(\phi \wedge \psi)$ is given by a pair (p,q) of proofs of ϕ and ψ , resp.; (ii) a proof of $(\phi \vee \psi)$ consists of a proof of ϕ or a proof of ψ (together with the information as to which of these is proved); (iii) a proof of

2) For further information and references on §4-§8, cf. Troelstra 1977a .

of $(\phi \to \psi)$ is a constructive operation for which we recognize that it will convert any proof q of ϕ into a proof (pq) of ψ. (iv) a proof of $\exists x\, \psi(x)$ consists of a pair (p,c) where p is a proof of $\psi(c)$; (v) a proof p of $\forall x\, \psi(x)$ is a constructive operation for which we recognize that it will convert any object c (in the intended range of the variable 'x') into a proof (pc) of $\psi(c)$.

Taking $\perp$ to be an identically false statement (e.g. $0 = 1$) which has no proof, negation is defined by $(\neg \phi) = (\phi \to \perp)$; thus proof of a negation of a statement (or of its absurdity) amounts to constructive recognition of the impossibility of proof of that statement. A proof of $\phi \vee \neg \phi$ is only given when one has a proof of ϕ or a proof of its absurdity.

There is a system of intuitionistic logic which is recognized to be correct for this interpretation of the logical operations, but which does not yield such (apparently) unacceptable principles as the law of excluded middle (LEM) or its consequence $\neg \forall x \neg \phi(x) \to \exists x\, \phi(x)$. Heyting has formulated this logic in such a way that classical logic is obtainable from it simply by adjunction of LEM. No further general logical principles have been recognized as constructively evident. (However, there is no generally recognized completeness result for intuitionistic logic.)

4.3. Functions are supposed to be constructive operations, the idea of which was already contained in 4.2 (iii),(v). These are supposed to be given by algorithmic rules of construction which can be effected by finite mechanical steps of computation. For relations with the recursion-theoretic concept of computable function cf. 4.8 below.

4.4. Sets are only given by defining properties , for which we are supposed to know and understand their condition for membership. For example, the condition for $x \in \mathbb{N}$ is that x is generated from 0 by a finite number of applications of the successor operation. If A,B are sets then $A \times B$, which consists of all ordered pairs (x,y) with $x \in A \wedge y \in B$, is a set. So also is B^A , where $x \in B^A$ iff $x : A \to B$, which means that x is a constructive operation such that for each $y \in A$, $x(y)$ (is defined and) belongs to B. Finally, if A is any set and $\phi(x)$ is a well-understood property of members of A then $B = \{x \in A \mid \phi(x)\}$ is a set, with $x \in B \leftrightarrow x \in A \wedge \phi(x)$.

4.5. Non-extensionality. Two rules may have the same values at all arguments (even provably so), but they are not identified unless the rules are recognized to be the same, as rules. (This allows for minor syntactic variations in the presentation of rules.) Two sets may have the same members, but they are not identified unless they are seen to be given by the same properties. (For the notion of intensional identity implicit here, cf. 4.11 below.)

4.6. Non-interchangeability of sets and functions. If B is a subset of A and $f: A \to \{0,1\}$ is such that $\forall x \in A[x \in B \leftrightarrow f(x) = 0]$ then we say that f is a characteristic function of B (rel. to A). Not every (sub)set (of a given set) has a characteristic function. Those which do are called decidable, otherwise undecidable. E.g. the set of exponents n for which Fermat's last theorem is true is (presently) undecidable. If every constructive function on $\mathbb{N}$ is recursive then every subset of $\mathbb{N}$ which is recursively undecidable is undecidable in the constructive sense. In any case, sets are not reducible to functions.

If $f: A \to B$ then the graph of f is a set, namely $R = \{(x,y) \in A \times B | f(x)=y\}$. R has the property $\forall x \in A\, \exists! y\, (x,y) \in R$. The question whether conversely, any such $R \subseteq A \times B$ determines a function $f: A \to B$ is a special case of the following. The conclusion is that functions are not reducible to sets.

4.7. The axiom of choice is considered here in the schematic form

$$\text{(AC)} \qquad \forall x \in A\, \exists y\, \phi(x,y) \to \exists f\, \forall x \in A\, \phi(x,f(x)).$$

This looks like it ought to be admitted using the interpretation of the connectives in 4.2. However we have to be careful : a proof p of the hypothesis, written out as $\forall x[x \in A \to \exists y\, \phi(x,y)]$ gives for each x and each proof q that x belongs to A (i.e. that x has the property which defines A) a proof p^* of $\exists y\, \phi(x,y)$ which is a pair $p^* = (p_1^*, y)$ where p_1^* proves $\phi(x,y)$. But p^* depends on both x and q, i.e. $p^* = p(q,x)$, so that y also is a function of x and q, not of x only as would be required for (AC). Writing $x \in_q A$ for 'q is a proof that x has the property determining A', this informal argument does justify accepting the following modified principle:

$$\text{(AC)}' \qquad \forall x \in A\, \exists y\, \phi(x,y) \to \exists f\, \forall q\, \forall x \in_q A\, \phi(x,f(x,q)).$$

We can derive (AC) only for those sets A for which we have a canonical choice of q, i.e. a function c such that $\forall x[x \in A \to x \in_{c(x)} A]$. $\mathbb{N}$ is an example of such a set; for $n \in \mathbb{N}$, the build-up of $\mathbb{N}$ gives itself the verification that we have a natural number. It may be noted that the principle

$$\text{(AC!)} \qquad \forall x \in A\, \exists! y\, \phi(x,y) \to \exists f\, \forall x \in A\, \phi(x,f(x))$$

is, for the same reasons, no more assured in general than AC. (These principles are dealt with formally in the framework of T_0 in Part IV below.)

4.8. Church's thesis. Let $e,n,m,\ldots$ range over $\mathbb{N}$, and take the usual notation $\{e\}(n)$ for partial recursive function application. The thesis that every (total) constructive function on $\mathbb{N}$ is recursive is referred to as

Church's Thesis in the literature on intuitionism (though it is open to argument whether Church himself had this in mind). Formally we can express it (in a 2nd order language) by

(CT) $\quad \forall f \in \mathbb{N}^{\mathbb{N}} \exists e \forall n[\{e\}(n)\downarrow \wedge f(n) = \{e\}(n)]$.

Note that the converse to Church's thesis is that

$$\forall n \exists m \{e\}(n) \simeq m \rightarrow \exists f \in \mathbb{N}^{\mathbb{N}} \forall n[f(n) = \{e\}(n)].$$

This follows from $(AC)_{\mathbb{N}}$ - which is acceptable by 4.7. There are some schemes related to (CT) which are expressible in 1st order form and follow from (CT) and $(AC)_{\mathbb{N}}$, in particular:

(CT_0) $\quad \forall n \exists m\, \phi(n,m) \rightarrow \exists e \forall n[\{e\}(n)\downarrow \wedge \phi(n,\{e\}(n))]$.

It is of logical interest that almost every known theory T which is informally constructively acceptable is consistent with (CT_0). However, the acceptability of this or of (CT) itself is a matter of dispute. As an example of the kind of argument which can be made against it, consider the following. Let J be a mathematician who works on deep problems of set theory and whose mental behavior is not duplicable by a machine. Then the function f defined by

$$f(n_0,n_1) = \begin{cases} 1 & \text{if on the } n_0\text{th day from now, J proves the } n_1\text{th theorem of ZF} \\ 0 & \text{otherwise} \end{cases}$$

is constructive but not recursive. Perhaps a more convincing argument against (CT) is that under the constructive interpretation of the logical operations if it held we would have to be able to pass constructively from any (proof of) $f \in \mathbb{N}^{\mathbb{N}}$ to a Turing machine e which calculates f. Thus even if human mental behavior is believed to be mechanical in principle, there is no constructive method of duplicating it by Turing machines. An argument for (CT) on the other hand, goes back to what is meant by constructive operation; at least in the form explained in 4.3, this would seem to be justified by Church's thesis in the usual sense that every finite algorithmic procedure can be carried out by a Turing machine.

4.9. Function sets and power sets. If Church's thesis is accepted, the meaning of $\mathbb{N}^{\mathbb{N}}$ is perfectly clear: it consists simply of the total recursive functions. An argument can be made without CT that we understand B^A for any sets A,B (whose condition for membership was given in 4.4) because our conception of constructive operation is supposed to be basic; this does not mean that "we know the totality of all constructive operations from A to B". The question of whether for each set A we have a set $\mathcal{P}(A)$ of all subsets of A seems to be different: even if we accept CT it is not clear that we have an understanding of what constitutes an arbitrary property of elements of $\mathbb{N}$, let alone of any A. The constructive status of $\mathcal{P}(A)$ is not settled; it is of mathematical and logical interest to investigate the effect of assuming its existence.

4.10. Comprehension principles. If a set A is given (understood and accepted) then quantification over A, i.e. the logical operations $\forall x \in A(\ldots)$ and $\exists x \in A(\ldots)$, are understood. Hence any property built using such operations determines a subset of any given set. If the existence of power sets is assumed then this leads us to impredicative comprehension principles, i.e. existence of $\{x \in A | \phi(x)\}$ where in ϕ we can quantify over $\mathcal{P}(A)$. Again, the constructive character of such principles is not settled, while their role and effect is of interest.

4.11. Literal, intensional and extensional identity. In 4.6 we spoke of functions given by the same rule or sets given by the same property. If we concentrate on syntactic representation of rules, properties, etc., then the most obvious notion of sameness to consider is that of literal identity, i.e. identity of syntactic configurations, symbol by symbol. A less definite but common idea is that rules, properties, etc. are mental objects which may have a variety of syntactic representations. For example, and most trivially, this may be by a renaming of bound variables or other symbols. More generally, we may have representations in different, but intertranslatable languages (so that the structure of the formal configurations may actually change). When two syntactic objects represent the same mental object they are said to be in the relation of intensional identity.

Most frequently in mathematics we are concerned with various kinds of defined relations of "equality" $=_A$ on a set A, which are simply equivalence relations. For example, when defining the integers $\mathbb{Z}$ as $\mathbb{N} \times \mathbb{N}$, we take $(n_1, m_1) =_{\mathbb{Z}} (n_2, m_2) \leftrightarrow n_1 + m_2 = n_2 + m_1$. When defining $\mathbb{Z}_p$ we take $x =_{\mathbb{Z}_p} y \leftrightarrow p | (x-y)$ for $x, y \in \mathbb{Z}$. The set $F = B^A$ has defined on it the relation $f =_F g \leftrightarrow \forall x \in A\ \forall y \in A[x =_A y \rightarrow f(x) =_B f(y)]$. All such equality relations are sometimes lumped together(perhaps misleadingly) under the heading of extensional identity relations. In Cantorian mathematics it is common to pass from $(A, =_A)$ to $(A/ =_A)$ so as to replace all equality relations by literal identity using the axiom of extensionality for sets. This practice is neither possible (without extensionality) nor necessary constructively; one simply makes clear for each set A considered what equality relation is being used in a given context. Note. It is common practice to drop the subscript 'A' from $=_A$ once that is fixed in any given discussion.

5. Constructive theory of real numbers. We sketch here how the preceding principles are used to set up a theory of real numbers. First of all, recursion on $\mathbb{N}$ is justified directly by its manner of generation, so we can define successively $+$, $\cdot$ and all further primitive recursive operations. $\mathbb{Z}$ and $=_{\mathbb{Z}}$ are defined as explained in 4.11, and $+$, $\cdot$, $<$ are extended in the standard way to $\mathbb{Z}$. Then $\mathbb{Q}$ is taken to consist of all (x,y) with $x, y \in \mathbb{Z}$ and $y \neq 0$ and $+$, $\cdot$, $<$ are extended to it. Next $\mathbb{Q}^{\mathbb{N}}$ consists of all sequences $\langle r_n \rangle_n$ of rational numbers. Cauchy sequences of rationals are those for which the Cauchy condition is constructively satisfied, i.e. for which we have a rate-of-convergence function $\mu : \mathbb{N} \to \mathbb{N}$ such that

$$\text{(C)} \qquad \forall k > 0 \ \ \forall n, m \geq \mu(k)[\,|r_n - r_m| < \tfrac{1}{k}\,].$$

By the set $\mathbb{R}$ of real numbers is meant the set of all pairs $x = (\langle r_n \rangle, \mu)$ with $\langle r_n \rangle \in \mathbb{Q}^{\mathbb{N}}$ satisfying (C). Then we put $x =_{\mathbb{R}} y$ for $y = (\langle s_n \rangle, \nu)$ if $(r_n - s_n) \to 0$. Real functions (of k arguments) are of course those operations $f\colon \mathbb{R}^k \to \mathbb{R}$ which preserve $=_{\mathbb{R}}$. In particular $+$ and $\cdot$ may be defined as real functions. For example, we may take $(\langle r_n \rangle, \mu_1) + (\langle s_n \rangle, \mu_2) = (\langle r_n + s_n \rangle, \nu)$ with $\nu(k) = \max(\mu_1(2k), \mu_2(2k))$.

The first essential difference is met with inverse and order. Given $x = (\langle r_n \rangle, \mu)$ we seek $x^{-1} = (\langle r_n^{-1} \rangle, \nu)$ but there is no obvious choice of ν unless we know a bound of x away from 0. Define $x > 0(m, k)$ if $\forall n \geq m\ (r_n \geq \frac{1}{k})$, and $x > 0$ if $\exists m,k(x > 0(m,k))$. Then define $x > y$ (or $y < x$) if $(x-y) > 0$ and finally $x \# y$ if $(x > y) \vee (x < y)$. We cannot establish constructively that $x \# y \vee x = y$. Inverse is defined for all $(x, (m, k))$ such that $|x| > 0\ (m,k)$. This is not strictly speaking a subset of $\mathbb{R}$, but only a subset by imbedding; such sets are dealt with systematically in BCM as will be described in § 14 below.

We could of course define $x \geq 0$ by $x > 0 \vee x = 0$, but it is more useful to take $x \geq 0$ to be $\forall k > 0\ \exists m\ \forall n \geq m\,(r_n \geq -\frac{1}{p})$; these are not constructively equivalent definitions, nor is $x \geq 0 \vee x < 0$ constructively justified. Classically, the expression

$$f(x) = \begin{cases} 1 & \text{if } x \geq 0 \\ 0 & \text{if } x < 0 \end{cases}$$

defines a function on $\mathbb{R}$ which is discontinuous, but this does not make constructive sense as a definition. Indeed there is no evident way to obtain a discontinuous function; theoretical reasons for this will be produced later.

6. Brouwer's intuitionism. [3] Brouwer both explored general constructive concepts (e.g. constructive operations, sets or "species", ordinals, etc.) and carried out particular mathematical developments, especially in analysis. He thought it should be possible to prove constructively that every (total) real function is continuous and that every real function on a closed bounded interval $F:[a,b] \to \mathbb{R}$ is uniformly continuous. For this purpose he introduced a new concept of free choice sequence (f.c.s.) $\langle r_n \rangle$ of which we know only a finite amount of information $(r_0,\ldots,r_k)$ at any given time, though we can proceed as far out as needed to make a calculation. The sequence may be produced randomly, e.g. by rolls of a die or observations of some random physical phenomena, rather than by some mechanical law. [4] It makes sense to operate constructively on such sequences to obtain values in $\mathbb{N}$ or $\mathbb{Q}$ or new f.c.s. themselves. For example, the operations $+$ and $.$ are easily defined for f.c.s. Now if $f(\langle r_n \rangle) = \langle s_n \rangle$ and a value s_m has been established, it can only have used a finite amount of information about $\langle r_n \rangle$; from this principle follows the statement of continuity of real functions where the reals are understood in the extended sense to include all those given by f.c.s. By some further (less immediately evident) principles Brouwer also derived the statement concerning uniform continuity.

Choice sequences need not be completely 'free'. They can be considered with or without restrictions on their values. For example, we can consider sequences $\langle r_n \rangle$ restricted by $|r_n| \leq M_n$ where $\langle M_n \rangle$ is given in advance, or is itself produced by some rule depending on earlier values of $\langle r_n \rangle$. Lawless sequences are those which are given without any restriction whatever. At the opposite end, lawlike sequences are those which are completely determined in advance by rules. The theory of reals sketched in § 5 may be interpreted as applying to the latter kinds of Cauchy sequences; for this reason it is sometimes called lawlike analysis.

Brouwer's analysis based on f.c.s. has been studied in various logical formalisms by Kleene, Vesley, Kreisel, Troelstra, van Dalen and others (cf. Troelstra 1977a 1977b for references). Various parts of this have taken settled and coherent form (and have, incidentally, been shown consistent). But efforts to treat the most general concept of f.c.s. have not yet had a convincing outcome. For mathematicians, Brouwer's theory has remained a curiosity;

---

3) A perusal of Brouwer 1975 (collected works, vol.I) is rewarding here.

4) Following a remark of Troelstra, H. Jervell has traced back the idea of f.c.s. to papers of E. Borel in 1912 which grew out of earlier discussions by the French mathematicians on the axiom of choice. At first Brouwer rejected the idea but later (1917) accepted it and expanded it into a theory.

it has largely been of interest to logicians. Moreover, the concepts are rather special to analysis and topology and seem to have little to do with other parts of mathematics. Historically, the actual development of intuitionistic mathematics got hung up around analysis because of the need to clarify Brouwer's ideas there.

It should be remarked that the intuitionistic theory of f.c.s. is inconsistent with classical mathematics, for we can prove $\neg \forall \langle r_n \rangle [\exists m(r_m = 0) \vee \forall m(r_m \neq 0)]$, as is intuitively evident from the 'finite-information' principle. Relatedly, one can disprove $\forall x \in \mathbb{R} (x \geq 0 \vee x < 0)$, etc. This is in contrast to lawlike analysis, which is a part of classical mathematics (if one does not assume (CT_0)).

7. The (Russian) school of Markov and Shanin. Here one accepts the scheme (CT_0) and the laws of intuitionistic logic, but also the following non-intuitionistic law, called Markov's principle:

$$\text{(MP)} \qquad \forall n[\phi(n) \vee \neg \phi(n)] \wedge \neg \forall n \neg \phi(n) \rightarrow \exists n \phi(n)$$

(where 'n' ranges over $\mathbb{N}$). The intuitive idea for (MP) is that under the hypothesis we can constructively find a solution n of the conclusion simply by performing a search through $\mathbb{N}$. It may be shown that $(CT_0) + (MP)$ is consistent over number theory though (CT_0) is inconsistent with full classical logic there. (The consistency proof can be given by Kleene's recursive realizability, which will be described in IV.) Various parts of analysis can be carried out under these assumptions, continuing the line sketched in §5. For example, if f is continuous on $[a,b]$ then $\inf_{a \leq x \leq b} f(x)$ and $\sup_{a \leq x \leq b} f(x)$ exist. However, it cannot be proved that f takes on its minimum (resp. maximum) in $[a,b]$. The reason is provided by a well-known example due to Specker of a recursively continuous function on $[0,1]$ which has no recursive point at which f takes on its minimum. Various other basic results of classical analysis may also be contradicted by suitable recursion-theoretic examples, e.g. that if f is continuous on $[a,b]$ and $f(a) < 0$ then $\exists x(a < x < b \wedge f(x) = 0)$. (In the Russian school it is admitted that there are some 'peculiarities' to their approach.)

8. Recursive analogues to classical mathematics. We have in mind here a series of studies concerning analogues to classical notions where one uses recursive functions (or functionals or sets) in place of arbitrary objects of the same type. To be mentioned in particular is the work of Dekker and Myhill for set theory, Crossley for order theory, Malcev and Rabin for algebra, and Specker and Lacombe for analysis and topology (cf. Feferman 1975 for references). These have been

carried out informally, with no restriction on the logic or methods employed. In effect, though, at least (CT) is assumed (though not (CT_0), which is classically inconsistent), and indeed a corresponding stronger principle identifying partial functions on $\mathbb{N}$ to $\mathbb{N}$ with partial recursive functions. For example, in the Dekker-Myhill theory of recursive equivalence types one defines

$$(A \sim B) \leftrightarrow \exists f, g[f,g \text{ partial recursive} \wedge f \upharpoonright A : A \to B \wedge g \upharpoonright B : B \to A \wedge (gf) \upharpoonright A = 1_A \wedge (fg) \upharpoonright B = 1_B] ,$$

where A,B may be arbitrary subsets of $\mathbb{N}$. One positive result which is proved for this is a form of the Cantor-Bernstein Theorem:

$$A \sim (B \dotplus C) \wedge B \sim (A \dotplus D) \to A \sim B.$$

In analysis one considers recursive real numbers (i.e. $x = (\langle r_n \rangle, \mu)$ with both $\langle r_n \rangle, \mu$ recursive) and recursive functions of reals (defined in an appropriate way via recursive functionals on $\mathbb{N}^{\mathbb{N}}$). As with the Russian school, a number of 'peculiarities' are met in this version of analysis.

While these pursuits are not constructive they can be relevant to constructive approaches in the following ways. Where a recursion-theoretic analogue gives a positive result, i.e. where a classical theorem carries over, one can often prove the same theorem constructively. On the other hand, when a negative result is obtained by suitable counter example, it is usually possible to use such to get underivability of the classical theorem in a constructive system. However, neither of these is automatic. For example, the least number principle

$$\exists n\, \phi(n) \to \exists n[\phi(n) \wedge \forall m < n \neg \phi(m)]$$

which is frequently applied in recursion-theoretic arguments is not constructively derivable except for decidable ϕ.

9. Bishop's approach. In 1967 Bishop published his Foundations of constructive analysis in which he carried out an informal development of constructive analysis which looked much more like modern analysis than anything done previously by constructivists and which went substantially further mathematically. Bishop works with general notions of function and set regarded in informal constructive terms. He rejects the notion of f.c.s. as being obscure and unnecessary. Instead of trying to prove that all functions of reals are continuous, his view is: there is not much of interest we can say about arbitrary functions from $\mathbb{R}$ to $\mathbb{R}$ or from $[a,b]$ to $\mathbb{R}$. Define $C([a,b],\mathbb{R})$ to be the uniformly continuous functions on $[a,b]$ to $\mathbb{R}$ and $C(\mathbb{R},\mathbb{R})$ to be the functions which

are uniformly continuous on each compact interval. (These definitions will be explained in more detail below.) These are classes of central mathematical interest. In a sense, Bishop is working in lawlike analysis and the notions and principles he uses are contained either directly or implicitly in Brouwer's intuitionism, but simply without f.c.s.. What is novel about Bishop's work is its spirit and execution, which is much more like everyday modern mathematics than anything previously done in a systematic constructive way. Indeed, a (philosophically unprepared) mathematician could pick up Bishop 1967 and read it as a straight piece of classical Cantorian mathematics. What would be puzzling to him is the more involved choice of certain notions and proofs, unless he also saw in what sense these were dictated by constructive requirements. It is this which is least successfully explained by Bishop. One of the main aims of the logical study of BCM is to elicit its underlying principles and to show how they may be interpreted constructively, as well as classically. One is led to consider constructive theories of functions, sets and classes which relate to BCM as theories like Zermelo - Fraenkel relate to Cantorian mathematics. Such systems could have been developed years ago, before Bishop, but it must be acknowledged that the work itself provided both the stimulus and a test for the adequacy of proposed theories.[5)]

10. Note on personal viewpoints. Bishop is a confirmed constructivist, as was Brouwer. Just as with Brouwer, he places the doing of constructive mathematics ahead of its logical study, regarding the latter as inessential. I am not a constructivist (nor a Platonist - it is harder to say what I am.) My main interests are logical and as a logician I am particularly interested in various forms of explicit mathematics (constructive, recursive, predicative, hyperarithmetic, inductive, Borelian, etc.) Of course this kind of position lends itself to greater objectivity, but there is also the possibility of insensitivity to, or neglect of, what are considered by a given school to be essential points.

11. Criteria of formalization. How well does a formal theory T represent an informal body of mathematics M? We judge this in terms of its adequacy and accordance.

(i) T is an adequate formalization of M if every concept, argument and result of M may be represented by a (basic or defined) concept, proof and theorem, resp. of T.

(ii) T is in accordance with (or faithful to) M if every basic concept of T corresponds to a basic concept of M and every axiom and rule of T corresponds to or is implicit in the assumptions and reasoning followed in M (i.e. T does not go beyond M conceptually or in principle).

5) Actually an informal constructive theory of functions and sets was outlined about the same time as Bishop's work in Tait 1968.

Remark. Formalisms always go syntactically beyond what is of ordinary interest, e.g. in practice we never assert $\phi \wedge \phi \wedge \phi$ or $\phi \to \psi$ where ϕ, ψ are unrelated.

We may refine (i),(ii) by considering whether the representation is direct or indirect. The idea of being (i)' directly adequate, resp. (ii)' directly in accordance with M seems clear. We would say that

(i)" T is indirectly adequate to M if there is a theory T^* directly adequate to M which can be translated into T (or otherwise reduced to T in an elementary way).

(ii)" T is indirectly in accordance with M if T can be translated or reduced to a theory T^+ which is directly in accordance with M.

A good formalization of M is one which is both directly adequate to and in accordance with M.

12. Illustrations of these criteria.

12.1. M = elementary number theory (non-analytic and non-algebraic).

$Z = Z^1$ = Peano's arithmetic with all primitive recursive function symbols.

$Z^1(+,\cdot)$ = Peano's arithmetic with just $+,\cdot$.

Z^2 = 2nd order arithmetic with full comprehension.

Z^1 is directly adequate to and directly in accordance with M.

$Z^1(+,\cdot)$ is directly in accordance with M but only indirectly adequate to it (by translation of Z^1).

Z^2 is directly adequate to M but not in accordance with M since the concept of $\mathcal{P}(\mathbb{N})$ as a completed totality is implicitly assumed in the comprehension scheme of Z^2.

12.2. M = classical analysis.

Z^ω = arithmetic in all finite types.

Z^ω is directly adequate to M.

Z^2 is indirectly adequate to M by reduction of the concepts that actually occur in practice to second order terms.

Z^ω is not in accordance with M; for in classical analysis we assume the totality $\mathbb{R}$, but not $\mathbb{R}^{\mathbb{R}}$ as a totality. [$C(\mathbb{R}, \mathbb{R})$ is assumed as a totality in the calculus of variations. Functionals in $C(\mathbb{R}, \mathbb{R}) \to \mathbb{R}$ are treated in modern analysis, but not higher type objects in any essential way.]

12.3. M = Cantorian set theory.

ZF = Zermelo-Fraenkel set theory.

ZF is directly adequate to M; is it in accordance with M? The question is raised since the idea of the cumulative hierarchy does not seem essential to M.

It is evident from these examples that the application of the criteria of formalization are reasonably objective, though there are cases of uncertainty.

13. Formal systems which have been proposed for BCM.

13.1. Bishop 1970, Goodman-Myhill 1972, both considered formalization in $HA^{\omega} + AC$ where HA^{ω} is intuitionistic arithmetic extended to finite types (HA = Heyting's intuitionistic arithmetic). HA^{ω} is directly in accordance with BCM. The question of Bishop's views on, and use of, AC is more delicate and will be taken up below. $HA^{\omega} + AC$ is inadequate to Bishop's theory of sets.

13.2. Martin-Löf 1975 (transfinite type theory). This is directly in accordance with BCM and adequate to everything but Bishop's theory of inductively defined classes (ordinals, Borel sets, etc.); it may also be naturally supplemented for the latter. It thus constitutes a good formalization of BCM. However, it is syntactically complicated, and not as flexible to work with as other theories to be discussed. This will be explained in more detail later. It should be added that Scott 1970 anticipated Martin-Löf 1975 in various ways.

13.3. Myhill 1975 CST (Constructive set theory), Friedman 1977. CST is a subtheory of IZFC/ZFC, intuitionistic ZFC, which like ZFC assumes extensionality and identifies functions with many one relations. Thus it is not directly in accord with constructive views, let alone BCM. It is indirectly adequate to BCM, as will be explained later. Friedman 1977 has considered a number of such theories and characterized their strength; he has also sketched interpretation into constructively justified theories, thus indirectly in accord with BCM.

13.4. Feferman 1975(T_0). This will be described in detail in Part II below. It is directly adequate to all of BCM. Accordance however is a matter of dispute; I shall argue that it is in accord, at least indirectly. T_0 is a type-free theory which is very amenable to metamathematical study and applications.

14. Some general features of BCM. As already said these incorporate the general features of constructivist mathematics outlined in §4, §5: the logic is intuitionistic, functions are given by rules, sets by defining properties; these are not interchangeable, and extensionality is not assumed. We detail in the following slight variants from §4-§5 above; novel points come with the treatment of subsets in 14.6.

14.1. Mathematical entities. The only objects which appear to be considered by Bishop are natural numbers, operations and sets, and such as are generated from these by pairing. Each such is considered to be presented by a finite symbolic expression.

14.2. Identity and equality. Two symbolic expressions are identical if they are presented in the same way - as in 4.11. We may take this to be literal identity or intensional identity. Each set considered has attached to it one or more relations of 'equality'. Notation: Bishop writes $\equiv$ for literal or intensional identity, $=$ for an equality relation on a set. We shall write instead $=$ for the first and $=_A$ for the second (but when there is no ambiguity we drop the subscript).

14.3. Operations and functions: Given sets $(A, =_A)$, $(B, =_B)$, f is an operation from A to B if it is a rule which assigns to each a in A an element $f(a)$ in B. f is a function from A to B if $a_1 =_A a_2 \rightarrow f(a_1) =_B f(a_2)$.

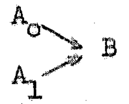
14.4. Function sets. Bishop says that for each A, B there is the set of all functions $F(A,B)$. Of course, if we consider A, B as sets endowed with the literal identity relation this implies that there is the set $O(A,B)$ of all operations from A to B. Since we can form these sets we can consider operations applied to operations, etc. Iterating F (or O) starting with N allows us to obtain the finite type hierarchy over N. We write B^A for $O(A,B)$.

14.5. Integers, Rationals, Reals. This follows §5 for $\mathbb{N}$, $\mathbb{Z}$ and $\mathbb{Q}$. However, Bishop defines $\mathbb{R}$ to consist of the following special class of Cauchy sequences: those $\langle x_n \rangle_{n \geq 1}$ such that $|x_n - x_m| \leq \frac{1}{n} + \frac{1}{m}$ for all n, m, (so the limit x satisfies $|x_n - x| \leq \frac{1}{n}$). $\mathbb{R}^+$ consists of pairs (x,k) where $x = \langle x_n \rangle$ is a real and $x_k > \frac{1}{k}$; for $\varepsilon = (x_k - \frac{1}{k})$ this means $x \geq \varepsilon > 0$. $(x,k) =_{\mathbb{R}^+} (y,\ell)$ is defined to mean $x =_{\mathbb{R}} y$.

14.6. Subsets. $\mathbb{R}^+$ is not a subset of $\mathbb{R}$ in the usual sense, but is one in the following sense of Bishop. By a subset (A,i) of B is meant a set A and operation $i : A \rightarrow B$ such that $a_1 =_A a_2 \leftrightarrow i(a_1) =_B i(a_2)$ (i.e. i is an injection of A into B). For x in B we say $x \in A$ if $x = i(y)$ with $y \in A$. Thus if we take $i(x,k) = x$ we have $(\mathbb{R}^+, i)$ a subset of $\mathbb{R}$. Similarly, intervals (a,b), $[a,b]$ etc. in $\mathbb{R}$ are given as subsets in this generalized sense. But we can also consider subsets in the usual sense where $i =$ identity operation on A.

14.7. Separation. Bishop recognizes that each property P applicable to elements of a set S determines the subset $A = \{x \mid x \in S \wedge P(x)\}$. (Implicitly, this is a subset in the usual sense; equality is defined to be $=_P$ restricted to A.)

14.8. Operations on subsets of a set. Using the more general concept of subset from 14.6, the operations of union and intersection take on more general (categorical) forms. Suppose given subsets (A_o, i_o) and (A_1, i_1) of B:



$(A_0 \cup A_1, i)$ is defined by Bishop to consist of all pairs (k,a) where $k=0$ and $a \in A_0$ or $k=1$ and $a \in A_1$; further, $i(k,a) = i_k(a)$ for $k=0,1$. Note this is essentially a form of disjoint union. $(A_0 \cap A_1, j)$ is further defined to consist of all (a_0, a_1) with $a_0 \in A_0$ and $a_1 \in A_1$ and $i_0(a_0) = i_1(a_1)$; then one takes $(j(a_0, a_1) = i_0(a_0)$. Note that $A_0 \cap A_1$ in this sense is contained in $A_0 \times A_1$.

14.9. Families of subsets of a set. By a family of subsets of a set B, indexed by T, is meant an operation f which associates with each $t \in T$ a subset $f(t) = (A_t, i_t)$ of B, in such a way that equal sets are associated with equal indices. The family is indicated by $\langle (A_t, i_t) \rangle_{t \in T}$ or $\langle A_t \rangle_{t \in T}$ or just $\langle A_t \rangle$. For simplicity, we shall only consider in the following the cases that the index set T is simply supplied with literal identity as its equality relation.

14.10. Operations on families. As an extension of the operations in 14.8 one defines $(\bigcup_{t \in T} A_t, i)$ and $(\bigcap_{t \in T} A_t, j)$ as subsets of B by:

$$\bigcup_{t \in T} A_t = \{(x,t) \mid x \in A_t\} \quad \text{and} \quad i(x,t) = i_t(x);$$

$$\bigcap_{t \in T} A_t = \{g \mid \forall t \in T(g(t) \in A_t) \wedge \forall t_1, t_2 \in T(i_{t_1}(g(t_1)) = i_{t_2}(g(t_2)))\} \quad \text{and}$$

$$j(g) = i_{t_0}(g(t_0)) \quad \text{for} \quad t_0 \in T.$$

The union is again a form of disjoint sum that we call the join of $\langle A_t \rangle$; it will be denoted below by $\sum_{t \in T} A_t$. In effect, the intersection is formed by separation from the cartesian product $\prod_{t \in T} A_t = \{g \mid \forall t \in T(g(t) \in A_t)\}$, on which equality $g_1 = g_2$ is defined by $\forall t \in T(g_1(t) =_{A_t} g_2(t))$, i.e. $\forall t \in T(i_t(g_1(t)) =_B i_t(g_2(t)))$.

14.11. Pre-joined families. An alternative definition of family which Bishop says could be considered is a subset A of $B \times T$. Certainly, given any such A we can define $f(t) = A_t = \{x \mid (x,t) \in A\}$. Then $A = \bigcup_{t \in T} A_t$ (extensionally). Thus we call such a family pre-joined, i.e. its prescription already guarantees existence of its join. In general though we need a join axiom which tells us that if f is a family $\langle A_t \rangle_{t \in T}$ then the join $J(T,f)$ exists.

14.12. Borel sets. These are inductively generated in a topological space from certain basic sets by the operations of union and intersection applied to countable families.

Abstractly this has the form:

(i) $B_0 \subseteq B$

(ii) $(f: \mathbb{N} \to B)$ implies $(J(\mathbb{N}, f) \in B \wedge I(\mathbb{N}, f) \in B)$.

(iii) if a property holds of all elements of B_0 and holds of $J(\mathbb{N}, f)$ and $I(\mathbb{N}, f)$ whenever it holds of $f(n)$ for each n, then it holds of all elements of B.

14.13 Principles in the general theory of integration. The Borel sets in a measure space are used in the development of integration theory in Bishop 1967, Ch.7. The theory of measure and integration was redeveloped by Bishop-Cheng 1972 without the use of Borel sets. This is an abstract theory, i.e. one starts with an arbitrary 'integration space' X and associates with it a certain completion L(X). It was pointed out by Friedman that the basic definitions in the latter approach make prima-facie use of the power-set operation which, as we have seen, is constructively problematic. However, this is only necessary if one wants L(X) again to be a set. The notion of being a member of L(X) does not require the power-set axiom and in that sense one can carry out abstract integration theory without this principle or the generalized inductive principles behind Borel sets. For more detailed examination of the issue here cf. Feferman 1978 §4. In any case, the potential (albeit marginal) mathematical utility of both generalized inductive and power-set principles in BCM makes them of interest for logical study. The former are incorporated directly in T_0, since they have constructive character.

15. General features of BCM, contd: Existential definitions and witnessing information. Notions which are defined classically using existential information are frequently replaced in BCM (as well as in other schools of constructive mathematics) by corresponding notions in which witnessing information is explicitly shown. This is required to carry out constructive operations on the objects satisfying the given notion.

15.1. Examples.

(i) $\mathbb{R}^+$. We have already explained its definition in §5 (as needed to make the operation of inverse constructive).

(ii) Limits of sequences of reals. By a convergent sequence of reals is meant a triple $(\langle x_n \rangle, x_0, m)$ where x_0 and each $x_n (n \geq 1)$ belongs to $\mathbb{R}$ and m is a function of positive integers such that (for all $k \geq 1$) $|x_n - x_0| \leq \frac{1}{k}$ for all $n \geq m(k)$; m is called a modulus-of-convergence function for the sequence.

(iii) Continuous functions. By a (uniformly) continuous function f on a compact interval $I = [a,b]$ is meant a pair (f,ω) with $f \in F(I,\mathbb{R})$ (i.e. $f: I \to \mathbb{R}$ preserving $=_{\mathbb{R}}$) and $\omega: \mathbb{R}^+ \to \mathbb{R}^+$ such that $|f(x)-f(y)| \leq \epsilon$ whenever $|x-y| \leq \omega(\epsilon)$; ω is called a modulus-of-uniform-continuity function for f. The set of all continuous functions from I to $\mathbb{R}$ is denoted by $C(I, \mathbb{R})$; it is a subset of $F(I,\mathbb{R})$ by $i\,(f,\omega) = f$. (This function ω is needed for example when performing the operation of integration of f over I.)

15.2. The logical pattern. In each case we are spelling out a property $P(x)$ involving existential quantifiers. In the above examples (i)-(iii) $P(x)$ is, respectively:

(i) $\exists k \geq 1 (x_k > \frac{1}{k})$

(ii) $\exists x_o \forall k \geq 1 \exists m \forall n \geq m (|x_n - x_o| \leq \frac{1}{k})$

(iii) $\forall \epsilon > 0 \exists \delta > 0 \forall x \forall y (|x-y| < \delta \rightarrow |f(x)-f(y)| < \epsilon)$.

With each such property is associated another $P^*(x,w)$ where w is some witnessing information that <u>realizes</u> or <u>verifies</u> $P(x)$. These properties are related by

(1) $\exists w P^*(x,w) \rightarrow P(x)$ and

(2) if (AC) is assumed, $P(x) \rightarrow \exists w P^*(x,w)$.

We shall also call P^* a <u>spelled-out form of</u> P. For example, in case (ii), w consists of x_o together with m as a function of k.

15.3. <u>Having your cake and eating it too</u>. Often Bishop defines a set A in what appears to be the form $\{x|P(x)\}$ (the <u>unofficial</u> definition) but then says that he is <u>really</u> defining the set $A^* = \{(x,w)|P(x,w)\}$ (the <u>official</u> definition).[6] He speaks of x being a member of A^* when it is really only x coupled with some side information w which can be considered to be a member. Then one treats operations on A^* as if they were operations on A. For example, the set A of (uniformly) continuous functions on I to $\mathbb{R}$ is officially defined as the set A^* of all (f,ω) satisfying the condition of 15.1(iii). But ω is not explicitly revealed in the operation $\int : C(I,\mathbb{R}) \rightarrow \mathbb{R}$ when written in the form $\int_a^b f(x)dx$.

There is a certain casualness in Bishop 1967 about mentioning the witnessing information as one goes along. Constructivity in theory requires that it be mentioned, but one is looser in practice in order to keep that from getting too heavy. Practice then looks very much like everyday analysis and it is hard to see what the difference is unless one takes the official definitions seriously.

15.4. <u>A concrete illustration</u>. The preceding is illustrated by a relatively simple example of a proof from Bishop 1967, but in which several spelling-out steps are already implicitly involved. This is for the theorem that <u>every continuous function on a compact interval has a l.u.b.</u> The reader should compare the following with the original as indicated by the page references.

<u>Definition</u> (p.34). c is called a l.u.b. of A (for $A \subseteq \mathbb{R}$ and $c \in \mathbb{R}$) if (i) $x \leq c$ for all x in A and (ii) for each $\epsilon > 0$ there exists x in A with $c - x < \epsilon$.

<u>Spelled out</u>, (ii) requires that we provide a function g such that $\forall \epsilon > 0$ ($g(\epsilon)$ in A and $c - g(\epsilon) < \epsilon$).

6) With reference to 15.1(i)-(iii) the reader should compare Bishop 1967 pp. 18, 19, pp.26, 27 and p.34, resp.

Theorem (p.34). Suppose A is a subset of $\mathbb{R}$ such that for each $\varepsilon > 0$ there exist $y_1, \ldots, y_n$ in A such that for each x in A at least one of the numbers $|x-y_1|, \ldots, |x-y_n|$ is $\leq \varepsilon$. (Such a set is called totally bounded). Then l.u.b. A exists.

Spelled out, the definition of being totally bounded (contained in the statement of this theorem) requires that we have two functions h and j such that for each $\varepsilon > 0$, $h(\varepsilon)$ is a finite sequence $(n, \langle y_1, \ldots, y_n \rangle)$ with each y_i in A (so $n = \ell h(h(\varepsilon))$) and for each x in A and $\varepsilon > 0$,

$$j(x, \varepsilon) \leq \ell h(h(\varepsilon)) \quad \text{and} \quad |x - y_{j(x,\varepsilon)}| \leq \varepsilon .$$

With this understanding, the proof of the above theorem proceeds as follows. Given any $k \geq 1$, choose $y_1, \ldots, y_n$ in A such that for each x in A, some $|x-y_j| \leq \varepsilon$. The choice of $\langle y_1, \ldots, y_n \rangle$ is given by $h(1/k)$ and of j by $j(x, 1/k)$. Next it is shown that

(1) for some m with $1 \leq m \leq n$ we have $y_m \geq \max\{y_1, \ldots, y_n\} - k^{-1}$.

For, each y_i is given as a Cauchy sequence of rationals $y_i = \langle y^i_p \rangle_{p \geq 1}$, i.e. each y^i_p is in Q and $|y^i_p - y^i_q| \leq \frac{1}{p} + \frac{1}{q}$ (from p.15). Take $p = 4k$ and find m such that $y^m_p \geq y^i_p$ for $i = 1, \ldots, n$. Then for $q \geq p$

$$(y^m_q - y^i_q) + \frac{1}{k} = (y^m_q - y^m_p) + (y^m_p - y^i_p) + (y^i_p - y^i_q) + \frac{1}{k} \geq (y^m_q - y^m_p) + (y^i_p - y^i_q) + \frac{1}{k} \geq 0$$

since $|y^i_q - y^i_p| \leq \frac{2}{p} = \frac{1}{2k}$. It follows that $y^m_q \geq (y^i_q - \frac{1}{k})$ for $q \geq p$ and then that $y_m \geq y_i - \frac{1}{k}$ as required for (1). Note that m in(1) is found as a function of k, say $m = m(k)$. Let $x_k = y_{m(k)}$. It is easily shown that $\langle x_k \rangle$ is a Cauchy sequence of reals. Using a result from p.27, it is shown that $\langle x_k \rangle$ converges and we can find its limit x_o. We claim that

(2) x_o is a l.u.b. for A.

For, given any x in A, choose $y_1, \ldots, y_n$ as above; then $x - x_k = (x - y_j) + (y_j - y_{m(k)}) < \frac{2}{k}$ when $j = j(x, 1/k)$. Hence $(x - x_o) = \lim_{k\to\infty} (x - x_k) \leq \lim_{k\to\infty} (2/k) = 0$; thus $x \leq x_o$ for all x in A. The function g required by the definition of l.u.b. is provided by $g(\varepsilon) = x_k$ where k is chosen so that $\frac{2}{k} < \varepsilon$, such k being calculable from the information which presents ε as a member of $\mathbb{R}$.

Corollary (p.35). If $f : [a,b] \to \mathbb{R}$ is continuous then $\{f(x) | x \in [a,b]\}$ has a l.u.b.

Proof. f is implicitly provided with a modulus-of-continuity function ω. Given any $\varepsilon > 0$ choose $a = a_o < a_1 < \ldots < a_n = b$ such that $(a_{i+1} - a_i) < \omega(\varepsilon)$ for each

$i=0,\ldots,n-1$. Let $a \le x \le b$. It is claimed that we can find j such that $|x-a_j| < \omega(\varepsilon)$. To do this we consider $x = \langle x_p \rangle$ and $a_i = \langle a_p^i \rangle$ each presented as a Cauchy sequence of rationals. Every x_p can be compared with each of $a_p^o,\ldots,a_p^i,\ldots,a_p^n$. For p sufficiently large (determined by $\omega(\varepsilon)$), the required j is found from this comparison. This shows that $A = \{f(x) \mid x \in [a,b]\}$ is totally bounded (as required by the official definition), and we can apply the preceding theorem.

The reader may wish to reconstitute one or two other proofs from Bishop 1967 in the same manner. (Another instructive example is provided by the proof of the Baire category theorem, p.87).

15.5 <u>A theoretical setting for</u> 15.3. In Part IV we shall present a theory T_o^* extending T_o in which the basic membership relation is refined to a 3-placed relation $x \varepsilon_w A$ and in which $x \varepsilon A$ is <u>defined</u> as $\exists w(x \varepsilon_w A)$. With reference to the logical pattern of 15.2, 15.3 one has $A^* = \{(x,w) \mid x \varepsilon_w A\}$, so that the 'unofficially' defined A actually determines the 'officially' defined A^*. T_o^* can be reduced to T_o so that in this theory we can have our (constructive) cake and eat it too.

15.6 <u>Remark on witnessing data in classical mathematics</u>. The practice of suppressing official parts of the defining data is also frequent in classical mathematics, e.g. algebraic or topological structures are simply referred to by their underlying sets. However the practice is more wholesale in BCM.

II. <u>The theory</u> T_o.

As presented here this theory is a minor modification of that introduced in Feferman 1975; the differences are explained below. For the reader's convenience a good deal of the material from secs. 2-3 loc. cit. is incorporated in the following. There are also some novel points.

1. <u>The language of</u> T_o ; <u>syntactical notions</u>.

1.1 <u>Variables and constants</u>. The language $\mathcal{L}(T_o)$ is two-sorted.

Individual (or general) variables	$a,b,c,\ldots,x,y,z$
Class variables	$A,B,C,\ldots,X,Y,Z$
Individual constants	$k,s,p,p_1,p_2,d,0,s_{\mathbb{N}},p_{\mathbb{N}},c_n,i,j$
Class constants	$\mathbb{N}$

We use $t,t_1,t_2,\ldots$ to range over variables or constants of either sort.

1.2 <u>Atomic formulas</u> are all those of the form $t_1 = t_2$, $App(t_1,t_2,t_3)$ and $t_1 \varepsilon t_2$. In addition there is an atomic formula $\bot$ with no free variables, interpreted as <u>falsity</u>.

1.3 <u>Logical operations</u>: $\wedge, \vee, \rightarrow, \forall, \exists$

1.4 Formulas are generated from atomic formulas by applying $\wedge, \vee, \rightarrow$, $\exists x, \forall x, \exists X, \forall X$.

Notation: ϕ, ψ, θ range over arbitrary formulas. $(\neg \phi) =_{def} (\phi \rightarrow \bot)$. $(\phi \leftrightarrow \psi)$, $\exists ! x \phi$, $(\forall x \in A)\phi$, $(\exists x \in A)\phi$ are defined as usual. $\underline{x}, \underline{X}, \underline{t}$ represent sequences of variables or terms, $\phi(\underline{x}, \underline{X})$ is written for a formula all of whose free variables are among $\underline{x}, \underline{X}$. In such expressions as $\exists y \psi(\underline{x}, y, \underline{X})$ it is assumed that y is not in the list $\underline{x}$ and similarly for $\exists Y \psi(\underline{x}, \underline{X}, Y)$.

2. Informal interpretation of the language. The individual variables range over the full universe of discourse of T_0, hence are also called general variables. These are to be thought of as mental objects (including rules and properties) or as symbolic representations of such objects. Then $=$ is interpreted as intensional identity or, in the latter view as literal identity of syntactic objects. The relation $App(t_1, t_2, t_3)$ is understood to hold when t_1 is a (rule or) operation which has value t_3 at the argument t_2. Since it is not assumed that every operation is total we shall write $t_1 t_2 \simeq t_3$ for $App(t_1, t_2, t_3)$. We write $t_1 t_2 \downarrow$ for $\exists z\, App(t_1, t_2, z)$. (This notation is expanded below.)

The class variables range over one-placed properties. $t \in X$ is interpreted as: t has the property (given by) X. (We may also think of 'class' as short for 'classification' ; classes are not conceived extensionally.) Class variables only range over a part of the universe of discourse. To express that an individual t happens to be a class we simply use the formula

$$C\ell(t) =_{def} \exists X(t = X).$$

The constants k, s are certain basic combinatory operations which permit one to form the constant operations and carry out the process of substitution, resp.; p, p_1, p_2 are operations of pairing and projection. d is an operation which gives definition-by-cases on $\mathbb{N}$, where $\mathbb{N}$ denotes the class of natural numbers, with least element 0 and operators of successor $s_{\mathbb{N}}$ and predecessor $p_{\mathbb{N}}$. The constants c_n, i, j represent certain class-formation operations, corresponding respectively to comprehension, inductive generation and join.

3. Application terms. The language $\mathcal{L}(T_0)$ is formally extended by a binary operator $\alpha(-,-)$ which is interpreted as the operation of application. We use $\tau, \tau_1, \tau_2, \ldots$ to range over the terms of this extended language, which are called application terms (a.t.'s). We write $\tau_1 \tau_2$ for $\alpha(\tau_1 \tau_2)$. Thus the a.t.'s are generated from the variables and constants of $\mathcal{L}(T_0)$ by closure under α. Since a.t.'s may not have defined values, relations between these are explained as formulas in $\mathcal{L}(T_0)$ in the following way:

$$(\tau \simeq x) =_{\text{def}} \begin{cases} (t = x) & \text{when } \tau \text{ is a term } t \text{ of } \mathcal{L}(T_o) \\ \exists y_1 \exists y_2 (\tau_1 \simeq y_1 \wedge \tau_2 \simeq y_2 \wedge y_1 y_2 \simeq x) & \text{when } \tau \text{ is } \tau_1\tau_2 \end{cases}$$

$(\tau_1 \simeq \tau_2) =_{\text{def}} \forall x [\tau_1 \simeq x \leftrightarrow \tau_2 \simeq x]$

$(\tau \downarrow) \quad =_{\text{def}} \exists x (\tau \simeq x)$.

We write $(\tau_1 = \tau_2)$ for $(\tau_1 \simeq \tau_2)$ when $(\tau_1\downarrow) \wedge (\tau_2\downarrow)$ is known or assumed.

$\phi(\tau) =_{\text{def}} \exists x(\tau \simeq x \wedge \phi(x))$.

In particular, $(\tau \in X)$ is $\exists x(\tau \simeq x \wedge x \in X)$.

Parentheses in $\tau_1\tau_2 \ldots \tau_n$ are supposed to be associated to the left as $(\ldots(\tau_1\tau_2)\ldots)\tau_n$. We write (τ) for τ, (τ_1,τ_2) for $p\tau_1\tau_2$ and $(\tau_1,\tau_2,\ldots,\tau_n)$ for $(\tau_1,(\tau_2,\ldots,\tau_n))$. This explains the notation $z(\tau_1,\ldots,\tau_n)$ or $z(\underline{\tau})$ for any $n \geq 1$. Finally, $\tau' =_{\text{df}} s_{\mathbb{N}}\tau$.

4. Further syntactical notions .

4.1 Stratified formulas are those (in $\mathcal{L}(T_o)$) which contain class variables or constants only on the right-hand side of $\in$ atomic formulas, i.e. only in contexts of the form $t \in X$ or $t \in \mathbb{N}$, where t is an individual variable or constant. Thus all the other atomic formulas of a stratified formula are relations of equality and applications between individual terms. Formally, stratified formulas may be thought of as 2nd order formulas with the sort of individuals specifying the 1st order level.

4.2 Elementary formulas are those stratified formulas which contain no bound class variables. These are also sometimes called predicative formulas, the others being impredicative. In an elementary formula $\phi(\underline{x}, \underline{X})$ classes are not referred to in any general way; we only require understanding membership in the given X_i . Elementary formulas may also be considered as the 1st order (stratified) formulas.

4.3 Comprehension notation. Let n be the Gödel number $\ulcorner\phi(x, \underline{y}, \underline{Z})\urcorner$ of ϕ with a specified inclusive list of variables $x, \underline{y}, \underline{Z}$. We put

$$\{x|\phi(x,\underline{y},\underline{Z})\} =_{\text{def}} c_n(\underline{y},\underline{Z}).$$

This shows the process of class formation by comprehension as a uniform function of the parameters $\underline{y}, \underline{Z}$ of the defining formula ϕ . The instances of comprehension are not all equally evident; the most evident are those corresponding to elementary ϕ, and only those are immediately accepted in T_o .

5. Axioms and logic of T_0.

5.1 The logic of T_0 is that of the intuitionistic two-sorted predicate calculus with equality. There is in addition a basic ontological axiom relating the sorts, namely $\forall X \exists x(X = x)$. Note that this justifies the formalism in 4.3 above where one applies operations to classes.

5.2 Non-logical axioms. These fall into several groups.

A P P(Applicative axioms)

(o) (unicity) $x\,y \simeq z \wedge x\,y \simeq w \rightarrow z = w$

(i) (constants) $kxy\downarrow \wedge\ kxy = x$

(ii) (substitution) $sxy\downarrow \wedge\ sxyz \simeq xz(yz)$

(iii) (Pairing, projections) $pxy\downarrow \wedge\ p_1z\downarrow \wedge p_2z\downarrow \wedge p_1(pxy) = x \wedge\ p_2(pxy) = y$

(iv) (definition by cases on $\mathbb{N}$) $x,y \in \mathbb{N} \rightarrow (x=y \rightarrow dxyab = a) \wedge (x \neq y \rightarrow dxyab = b)$.

(v) (zero, successor, predecessor) $x,y \in \mathbb{N} \rightarrow x'\downarrow \wedge p_{\mathbb{N}}y\downarrow \wedge p_{\mathbb{N}}(x') = x \wedge x' \neq 0 \wedge (x' = y' \rightarrow x=y)$.

The remaining axioms are class existence axioms. Note from I(iii) that $c_n(y_1,\ldots,y_m, Z_1,\ldots,Z_p)\downarrow$ for all $\underline{y}, \underline{Z}$.

E C A (Elementary comprehension). For each elementary $\phi(x,\underline{y},\underline{z})$ we take:

$$\exists X\{\{x|\phi(x,\underline{y},\underline{Z})\} = X \wedge \forall x[x \in X \leftrightarrow \phi(x,\underline{y},\underline{Z})]\}$$

$\mathbb{N}$ (Natural numbers)

(i) (closure) $0 \in \mathbb{N} \wedge \forall x(x \in \mathbb{N} \rightarrow x' \in \mathbb{N})$

(ii) (induction) $\phi(0) \wedge \forall x(\phi(x) \rightarrow \phi(x')) \rightarrow \forall x \in N\,\phi(x)$

where $\phi(x) = \phi(x,\ldots)$ is an arbitrary formula of $\mathcal{L}(T_0)$.

J (Join)

$\forall x \in A \exists Y(fx \simeq Y) \rightarrow \exists X\{j(A,f) = X \wedge \forall z[z \in X \leftrightarrow \exists x \exists y(z = (x,y) \wedge x \in A \wedge y \in fx)]\}$.

IG (Inductive generation)

(i) (closure) $\exists I\{i(A,R) = I \wedge \forall y((y,x) \in R \rightarrow y \in I) \rightarrow x \in I\,]\}$

(ii) (induction) $\forall x \in A\{\forall y(y,x) \in R \rightarrow \phi(y)) \rightarrow \phi(x)) \rightarrow (\forall x \in i(A,R))\phi(x)$

where $\phi(x) = \phi(x,\ldots)$ is an arbitrary formula of T_0. This completes the list of axioms: $T_0 = APP + ECA + \mathbb{N} + J + IG$.

6. Relation of T_0 as given here with that of Feferman 1975.

(a) Previously we took a one-sorted language with an additional predicate $C\ell(x)$; the variables $A,B,C,\ldots,X,Y,Z$ were introduced by convention to range over $C\ell$.

(b) We defined $x' = (x,0)$ and $p_{\mathbb{N}} = p_1$ and took the axiom $(x,y) \neq 0$; then the axiom APP(v) as given here was derived.

(c) The constant $\mathbb{N}$ was defined as $i(A,R)$ for

$$A = \{x|x = 0 \vee x \simeq (p_{\mathbb{N}}x)'\} \text{ and}$$

$$R = \{z|\exists x,y\,(z = (y,x) \wedge x = y')\} \qquad \text{(predecessor relation).}$$

The reason for listing $\mathbb{N}$ as a special axiom here is so as to consider the strength of T_o - IG while still including $\mathbb{N}$.

(d) Previously we used definition by cases on the universe instead of just on $\mathbb{N}$, as will be explained next.

7. Some variants of the axioms which will be considered.

D_V (Definition by cases on the universe). This is the following in place of APP(v):

$(x = y \vee x \neq y) \wedge dxyab\downarrow \wedge (x = y \rightarrow dxyab = a) \wedge (x \neq y \rightarrow dxyab = b)$.

Then $T_o + D_V$ is equivalent to the system T_o of Feferman 1975.

Note. The clause $(x,y \in \mathbb{N} \rightarrow x = y \vee x \neq y)$ was not needed in APP(v) since it is derivable from the other axioms.

CA_1 (First order comprehension). This is another denotation of ECA.

CA_2 (Second order comprehension). The same scheme as for CA_1 except that ϕ may now be any stratified formula in $\{x|\phi(x,\underline{y},\underline{Z})\}$.

SEP_1 (First order separation). This is CA_1 restricted to formulas of the form $x \in A \wedge \psi(x,\underline{y},\underline{Z})$ (with parameters $\underline{y},\underline{Z},A$). $\{x \in A|\psi(x,\underline{y},\underline{Z})\}$ is written for $\{x|x \in A \wedge \psi(x,\underline{y},\underline{Z})\}$.

SEP_2 (Second order separation). The same as SEP_1 with any stratified ψ.

$\mathbb{N}\upharpoonright$ (Restricted induction on N). Here one replaces the induction schema $\mathbb{N}$(ii) by the specific instance

$$0 \in X \wedge \forall x(x \in X \rightarrow x' \in X) \rightarrow \mathbb{N} \subseteq X$$

(where $(X \subseteq Y) =_{def} \forall x(x \in X \rightarrow x \in Y)$). Note that $\mathbb{N}\upharpoonright$ can be used to derive any instance

$$\phi(0) \wedge \forall x(\phi(x) \rightarrow \phi(x')) \rightarrow \forall x \in \mathbb{N}\, \phi(x)$$

for which $\{x|\phi(x,\dots)\}$ is known to exist as a class.

IG$\upharpoonright$ (Restricted induction for IG). Analogously replaces IG(ii) by

$$\forall x \in A\,\{\forall y((y,x) \in R \rightarrow y \in X) \rightarrow x \in X\} \rightarrow i(A,R) \subseteq X\,.$$

Remark. We could formulate a generalized inductive definition axiom GID expressing for any elementary $\psi(x,X)(=\psi(x,X,\dots))$ in which X occurs only positively, the existence of a class I which is the least predicate satisfying $\forall x[\psi(x,X) \rightarrow x \in X]$. This stronger axiom is not evidently constructive (GID$\upharpoonright$ can be derived if one accepts the impredicative comprehension principle CA_2).

Note. IG itself becomes still more evident if one writes $y <_R x$ for $(y,x) \in R$. Then $i(A,R)$ is the class of $<_R$ - accessible elements of A, which may be pictured as the elements of A sitting atop well-founded trees branching by the $<_R$ relation.

MIG$\upharpoonright$ (Restricted monotone inductive definition)

$$\forall X \exists Y[fX \simeq Y] \wedge \forall X_1,X_2[X_1 \subseteq X_2 \rightarrow fX_1 \subseteq fX_2] \rightarrow \exists I[fI \subseteq I \wedge \forall X[fX \subseteq X \rightarrow I \subseteq X]]\,.$$

By adjunction of a suitable constant, we could also express I uniformly as a function of f. Here f represents a monotone operation on classes to classes. This is stronger than $GID\restriction$ (in the presence of ECA), since positivity is weaker than monotonicity.

8. Product and power class axioms.

P (Product axiom) $\forall x \in A\, \exists Y(fx \simeq Y) \rightarrow \exists X\, \forall z[z \in X \leftrightarrow \forall x \in A(zx \in fx)]$.

We shall prove (11.2) that P follows from T_0; however, it does not if CA_1 is replaced by SEP_1.

POW^+ (Strong power class axiom). $\forall A\, \exists B\, \forall x[x \in B \leftrightarrow \exists X(X \subseteq A \wedge x = X)]$.

POW (Weak power class axiom).

$$\forall A\, \exists B\, \forall x[x \in B \rightarrow \exists X(X \subseteq A \wedge x = X) \wedge \forall X(X \subseteq A \rightarrow \exists Y(X \equiv Y \wedge Y \in B))]$$

where $X \equiv Y$ is the relation of extensional equality between classes,

$$(X \equiv Y) =_{def} (X \subseteq Y) \wedge (Y \subseteq X).$$

As remarked previously, the constructive status of POW^+ (or even POW) is unclear.

Note. Each of these can be expressed uniformly by adjunction of suitable constants.

9. Theories related to T_0 which are to be considered here.

9.1 $EM_0 = APP + ECA + \mathbb{N}$

$EM_0\restriction = APP + ECA + \mathbb{N}\restriction$.

Remark. The notation 'EM' comes from 'Explicit mathematics'. $EM_0\restriction$ is a minimal constructive theory in the present framework. We shall consider the effect of adding, variously, D_V, J, $IG\restriction$ and IG to these theories.

9.2 A theory of sets S_0.

$$S_0 = APP + SEP_1 + \mathbb{N} + J + P + IG.$$

The class variables here can be interpreted as sets ("small" classes). (The product axiom P is added here according to the remark made in § 8.)

9.3 A theory of sets and classes $T_0(S)$. This may be obtained from T_0 by adjoining a constant S for the class of all sets, for which suitable closure conditions are expressed by additional axioms. A theory of this character was presented in Feferman 1978, but of greater generality, since sets there are taken to be pairs (A,E) where E is an equality relation on A.

10. Consequences of $EM_0\restriction$. Throughout the following all statements are to be considered semi-formally. They are all provable in $EM_0\restriction$.

Notation. We shall often use letters like 'f', 'g', 'h' for individual variables being treated in operator situations. The letters 'k', 'n', 'm', 'p' are now reserved to range over $\mathbb{N}$.

10.1 Abstraction. For each application term τ and variable x we can find a term τ^* with variables $\subseteq$ var $(\tau) - \{x\}$ such that

$$\tau^* \downarrow \wedge \forall x[\tau^* x \simeq \tau].$$

τ^* is denoted $\lambda x(\tau)$ or $\lambda x(\tau[x])$. (Intuitively, $\tau^*\downarrow$ because it always denotes a rule, whether or not $\tau[x]\downarrow$ at any x.) The proof is carried out by induction on τ (just as in total combinatory theories):

(i) If τ is x, we take skk for $\lambda x(x)$ since $skkx \simeq kx(kx) = x$.
(ii) If τ is an individual term $t \neq x$, we take kt for $\lambda x(t)$ since $ktx = t$.
(iii) If $\tau = \tau_1\tau_2$ we want $\tau^* x \simeq \tau_1[x]\tau_2[x]$. But $\tau_1[x]\tau_2[x] \simeq (\tau_1^* x)(\tau_2^* x) \simeq s\tau_1^*\tau_2^* x$, so we can take $\tau^* = s\tau_1^*\tau_2^*$ in this case. (Only APP(i),(ii) are applied here).

10.2 <u>The recursion theorem.</u> We can find a fixed r such that for all f:

$$rf\downarrow \text{ and for } g = rf,\ \forall x(gx \simeq fgx).$$

For the proof take $h = \lambda y\,\lambda x f(yy)x$, so $hy\downarrow$ for all y. In particular $hh\downarrow$ and $hh = \lambda x(f(hh)x)$. Thus $g = hh$ serves as rf. We can take

$$r = \lambda f((\lambda y \lambda x f(yy)x)(\lambda y \lambda x f(yy)x)).$$

10.3 <u>Recursion on $\mathbb{N}$</u>. For any a, f we can find g satisfying:

$$gx \simeq \begin{cases} a & \text{if } x = 0 \\ f(x, g(p_{\mathbb{N}} x)) & \text{if } x \in \mathbb{N} \text{ and } x \neq 0 \end{cases} .$$

Namely, g is found by the recursion theorem so as to make

$$gx \simeq dx\,0\,a(f(x, g(p_{\mathbb{N}} x))).$$

It follows that

(i) $g0 = a$

(ii) $gn' \simeq f(n, gn)$ for any n.

Then for any a, A, f we have:

(iii) $a \in A \wedge \forall n \forall y[y \in A \rightarrow f(n,y) \in A] \rightarrow \forall n(gn \in A)$,

with the conclusion being proved from the hypothesis by restricted induction, since $\{x | gx \in A\}$ is a class.

10.4 <u>Arithmetic in $EM_0\restriction$</u>. It follows from the preceding that all primitive recursive functions can be defined in $EM_0\restriction$. Furthermore, every arithmetical formula is equivalent in this theory to an elementary formula, hence defines a class. Thus the scheme of induction for arithmetical formulas holds in $EM_0\restriction$. Hence the intuionistic system HA of (Heyting's) arithmetic is contained in $EM_0\restriction$.

10.5 <u>Bounded and unbounded minima</u>. Using recursion on $\mathbb{N}$ we can define $\prod_{m \leq n} fm$ so that $\forall m \leq n(fm \in N) \leftrightarrow \prod_{m \leq n} fm \in N$ and $\prod_{m \leq n} fm = 0 \leftrightarrow \exists m \leq n(fm \leq 0)$.

Then further we can obtain $(\mu m \leq n)(fm \simeq 0)$ which is defined under the same conditions. By the recursion theorem one finds g such that

$$g(f,n) \simeq \begin{cases} (\mu m \le n)\ (fm \simeq 0) & \text{if } \exists m \le n(fm \simeq 0) \\ g(f,n') & \text{otherwise.} \end{cases}$$

Let $\mu f = g(f,0)$. It is seen that μf is defined and equal to $\mu n(fn \simeq 0)$ just in case $\exists n(fn = 0 \wedge \forall m < n(fm \in \mathbb{N}))$.

10.6 Partial recursive functions; forms of Church's Thesis. The enumeration of partial recursive functions $\{k\}$ for $k \in \mathbb{N}$ can now be defined as usual in $EM_0\restriction$ with $\{k\}(n) \simeq U(\mu n T_1(k,n,m))$. We now introduce the function-mapping notation:

$$(f : A \to B) =_{def} \forall x \in A\,(fx \in B).$$

The three forms of Church's thesis described in I.4.8 can be formulated in $EM_0\restriction$ as follows.

CT_0 is the scheme $\forall n \exists m \phi(n,m) \to \exists k\, \forall n[\{k\}(n)\downarrow \wedge\, \phi(n, \{k\}(n))]$

CT_1 is $\forall f[(f:\mathbb{N} \to \mathbb{N}) \to \exists k\, \forall n(fn = \{k\}(n))]$

CT_2 is $\forall f[\forall n(fn\downarrow\ \to fn \in \mathbb{N}) \to \exists k\, \forall n\,(fn \simeq \{k\}(n))]$.

(CT_2 was suggested by Beeson; it expresses that every partial function on $\mathbb{N}$ coincides with a partial recursive function there.) Obviously $CT_2 \to CT_1$, and as we remarked in 4.8, $CT_1 + AC_{\mathbb{N}} \to CT_0$. The consistency of these with T_0 will be taken up in Parts III and IV.

10.7 Elementary operations on classes. We now turn to uses of ECA. The following give operations having class values as functions of the individual and class parameters shown.

$V = \{x | x = x\}$

$\Lambda = \{x | \bot\}$

$\{a,b\} = \{x | x = a \vee x = b\}$

$-A = \{x | x \notin A\}$

$A \cup B = \{x | x \in A \vee x \in B\}$, $A \cap B = \{x | x \in A \wedge x \in B\}$

$A \times B = \{z | \exists x \in A\, \exists y \in B\ \ z = (x,y)\}$

$B^A = \{f | f:A \to B\}$ (also denoted $A \to B$)

$Df = \{x | fx\downarrow\}$

$f[A] = \{y | \exists x \in A(fx \simeq y)\}$.

Evidently all these have the form $\{x | \phi(x,\ldots)\}$ with ϕ elementary.

10.8 The finite type hierarchy and HA^ω. The finite type symbols (f.t.s.) are generated by the following elementary inductive definition: $\dot{0}$ is a f.t.s. and if ρ,σ are f.t.s. then so also are $\rho \dot{\times} \sigma$ and $\rho \dot{\to} \sigma$, where $\dot{0} = (0,0)$, $u \dot{\times} v = (1,u,v)$ and $(u \dot{\to} v) = (2,u,v)$. The f.t.s. are enumerated by a function on

$\mathbb{N}$ whose range thus forms a class that we denote by FTS. Using recursion on $\mathbb{N}$ we can define a function g satisfying:

$$g\dot{o} = \mathbb{N}\,,\quad g(u\dot{\times}v) = gu \times gv \quad\text{and}\quad g(u \dot{\rightarrow} v) = (gu \rightarrow gv)$$

where $\times, \rightarrow$ are the operations defined in 10.7. For each $\sigma \in$ FTS we denote by N_σ the value $g\sigma$. Thus

$$N_{\dot{o}} = \mathbb{N}\,,\quad N_{\rho\dot{\times}\sigma} = N_\rho \times N_\sigma \quad\text{and}\quad N_{\rho\dot{\rightarrow}\sigma} = (N_\rho \rightarrow N_\sigma) = N_\sigma^{N_\rho}\,.$$

For each <u>particular</u> σ in FTS we can prove in $EM_o\upharpoonright$ that

$$C\ell(N_\sigma).$$

However, the statement

$$\forall\sigma \in \text{FTS}\,[C\ell\,(N_\sigma)]$$

requires a proof by induction using the impredicative property $\exists X(g\sigma \simeq X)$. This can be carried out in EM_o but not $EM_o\upharpoonright$.

Gödel's notion of <u>primitive recursive functional of finite type</u> (Gödel 1958) can be interpreted in $EM_o\upharpoonright$ simply by using recursion on $\mathbb{N}$. The basic scheme is to pass from $f_o \in N_{\rho \dot{\rightarrow} \sigma}$ and $f_1 \in N_{(\rho\dot{\times}0\dot{\times}\sigma \dot{\rightarrow} \sigma)}$ to a g satisfying

$$g(x,0) \simeq f_o(x) \quad\text{and}\quad g(x,n') \simeq f_1(x,n,g(x,n)),$$

which is obtained using 10.3 uniformly in x. Now we can prove by restricted induction on $\mathbb{N}$ that $x \in N_\rho \rightarrow g(x,n) \in N_\sigma$, hence $g \in N_{\rho\dot{\times}0 \dot{\rightarrow} \sigma}$ as required.

The theory HA^ω of intuitionistic arithmetic in all finite types has variables of each type $\sigma \in$ FTS and constants for the primitive recursive functionals.[7)] The axioms are those of HA together with the defining schemata for all these functionals and, finally, the induction scheme for all formulae of the language. Interpreting the variables of type σ to range over N_σ, each formula of HA^ω is equivalent to an elementary formula with finitely many constants $N_{\sigma_1},\ldots,N_{\sigma_m}$. Hence it defines a class under ECA. It follows that

$$HA^\omega \subseteq EM_o\upharpoonright\,.$$

<u>Remark</u>. An intensional form of HA^ω, denoted $I\text{-}HA^\omega$ is obtained by adjoining a functional at each type level which decides equality between objects of that type. It is easily seen that

$$I\text{-}HA^\omega \subseteq EM_o\upharpoonright + D_V\,.$$

10.9 <u>The extensional finite type hierarchy</u>. The system $E\text{-}HA^\omega$ obtained from HA^ω by adjoining extensionality axioms in all finite types can also be interpreted in $EM_o\upharpoonright$. However, here we interpret the variables of type σ to

7) Cf. Troelstra 1973, Part I §6 for a precise description of HA^ω and the systems $I\text{-}HA^\omega$, $E\text{-}HA^\omega$ below.

range over M_σ, defined together with an equality relation $=_\sigma$ by induction as follows:

$$M_0 = \mathbb{N} \text{ and } n =_0 m \leftrightarrow n = m$$

$$M_{\rho \dot{\times} \sigma} = M_\rho \times M_\sigma \text{ and } x =_{\rho \dot{\times} \sigma} y \leftrightarrow p_1 x =_\rho p_1 y \wedge p_2 x =_\sigma p_2 y$$

$$M_{\rho \dot{\to} \sigma} = \{f \mid f \in M_\rho \to M_\sigma \wedge \forall x, y \in M_\rho (x =_\rho y \to fx =_\sigma fy)\} \text{ and}$$

$$f =_{(\rho \dot{\to} \sigma)} g \leftrightarrow \forall x \in M_\rho (fx =_\sigma gx).$$

Equality between objects of type σ is interpreted as $=_\sigma$ so as to obtain $\text{E-HA}^\omega \subseteq EM_0 \upharpoonright$.

10.10 Classes with equality relations. These are simply pairs (A, E) where $E \subseteq A \times A$ and E is an equivalence relation on A. Then we write $x =_A y$ for $(x,y) \in E$ (though this notation is ambiguous, since E is not uniquely associated with A). While we can operate on classes-with-equality in T_0 (or its subtheories), we proceed more generally than in Bishop 1967 and work with classes per se.

10.11 Integers, rationals and reals in $EM_0 \upharpoonright$. Our definitions here follow I, 14.5 (i.e. essentially Bishop 1967 Ch.2)

$\mathbb{Z} = \mathbb{N} \times \mathbb{N}$; $(n,m) =_{\mathbb{Z}} (p,q) \leftrightarrow n + q = m + p$.

$+$ is defined on $\mathbb{Z}$ by $(n,m) +_{\mathbb{Z}} (p,q) = (n+p, m+q)$, and so on for $\cdot, <$ on $\mathbb{Z}$. $\mathbb{N}$ is embedded in $\mathbb{Z}$, and subscripts are dropped.

$\mathbb{Q} = \{(x,y) \mid x \in \mathbb{Z} \wedge y \in \mathbb{Z} \wedge y \neq 0\}$; $(x,y) =_{\mathbb{Q}} (u,v) \leftrightarrow x \cdot v = y \cdot u$

$(x,y) +_{\mathbb{Q}} (u,v) = (xv + yu, uv)$, and so on for $\cdot, <$ on $\mathbb{Q}$.

$\mathbb{Z}$ is embedded in $\mathbb{Q}$, and subscripts are dropped.

$\mathbb{Z}^+ = \{n \mid n \in \mathbb{Z} \wedge n > 0\}$. We write x_n for xn when $x \in \mathbb{Z}^+ \to A$

$\mathbb{R} = \{x \mid x \in \mathbb{Z}^+ \to \mathbb{Q} \wedge \forall n, m \in \mathbb{Z}^+ (|x_n - x_m| \leq \frac{1}{n} + \frac{1}{m})\}$

$x =_{\mathbb{R}} y \leftrightarrow \forall k \in \mathbb{Z}^+ \exists m \in \mathbb{Z}^+ \forall n \geq m (|x_n - y_n| \leq \frac{1}{k})$.

$x +_{\mathbb{R}} y = \lambda u (x_{2u} + y_{2u})$,

$x \cdot_{\mathbb{R}} y = \lambda u (x_{2ku} \cdot y_{2ku})$ where $k = \max(k_x, k_y)$ and for each x, k_x is the least integer greater than $|x_1| + 2$.

$\mathbb{R}^+ = \{(x,n) \mid x \in \mathbb{R} \wedge x_n > \frac{1}{n}\}$ and $\mathbb{R}^{0+} = \{x \mid x \in \mathbb{R} \wedge \forall n \in \mathbb{Z}^+ (x_n \geq -\frac{1}{n})\}$.

$\mathbb{Q}$ is embedded in $\mathbb{R}$, and the subscripts are dropped.

The elementary properties of these number systems can be developed in $EM_0 \upharpoonright$ directly following Bishop 1967. Then the complex numbers $\mathbb{C}$ can be introduced as usual and their properties derived in the same way.

10.12 Continuous functions and classical analysis. Given any two classes A, B with equality relations $=_A$, $=_B$ respectively, the function class F(A,B) is defined by

$$F(A,B) = \{f \mid f : A \to B \wedge \forall x,y \in A(x =_A y \to f(x) =_B f(y))\}.$$

This is a subclass of $(A \to B)$. In particular in analysis one is interested in $F(\mathbb{R}, \mathbb{R})$ and $F([a,b], \mathbb{R})$ where $a,b \in \mathbb{R}$ and $[a,b] = \{x \in \mathbb{R} \mid a \leq x \leq b\}$. Next for the (uniformly) continuous functions on $[a,b]$, following I.15.1, one takes

$$C([a,b],\mathbb{R}) = \{(f,w) \mid f \in F([a,b], \mathbb{R}) \wedge w \in F(\mathbb{R}^+, \mathbb{R}^+) \wedge \forall e \in \mathbb{R}^+ \forall x,y \in [a,b](|x-y| \leq w(e) \to |f(x) - f(y)| < e)\},$$

and one takes

$$C(\mathbb{R}, \mathbb{R}) = \{(f,u) \mid f \in F(\mathbb{R},\mathbb{R}) \wedge \forall a, b \in \mathbb{R}(a<b \to (f,u(a,b)) \in C([a,b], \mathbb{R}))\}$$

i.e. the functions continuous on each compact interval $[a,b]$. Starting with this as basis, classical real analysis is pursued in $EM_0\restriction$ just as in Bishop 1967, Ch. 2. The only point which requires careful checking is that only restricted induction on $\mathbb{N}$ is applied throughout. We shall return to this observation in §14 below.

11. Consequences of the join axiom. Here, unless otherwise specified, we work within $EM_0\restriction + J$.

11.1 Families and joins. By a family of classes $\langle B_x \rangle_{x \in A}$ is meant an operation f such that $\forall x \in A\, \exists Y(fx \simeq Y)$ and where we write B_x for fx. The join axiom guarantees the existence of a class $\sum_{x \in A} B_x$ with the defining property

$$z \in \sum_{x \in A} B_x \leftrightarrow \exists x \in A\, \exists y[z = (x,y) \wedge y \in B_x].$$

From this we can define the union operation on classes-with-equality as explained by Bishop (I.14.10) above. By a pre-joined family on A is meant a class $B \subseteq A \times V$. Associated with such is a family in the preceding sense by $fx(= B_x) = \{y \mid (x,y) \in B\}$; extensionally this makes $B \equiv \sum_{x \in A} B_x$.

11.2 Products. Suppose given a family $f = \langle B_x \rangle_{x \in A}$. Let $J = \sum_{x \in A} B_x$. Then we can define

$$\prod_{x \in A} B_x = \{g \mid \forall x \in A((x,gx) \in J)\}$$

since $(x,gx) \in J \leftrightarrow gx \in B_x$ for $x \in A$. From this we can treat intersection of a family of classes-with-equality (I.14.10).

Remark. There is no evident way to derive the product axiom P from the remaining axioms in the theory S_0 of sets in 9.2.

11.3 The passage to transfinite types. Using the operations of 10.7 we know that

$$C\ell(a) \wedge C\ell(b) \to C\ell(a \times b) \wedge C\ell(a \to b) .$$

Then by enumerability of FTS in 10.8 we can carry out an induction to prove

$$\forall \sigma \in \mathrm{FTS}\,[C\ell(N_\sigma)] .$$

However this requires unrestricted induction on $\mathbb{N}$ in the theory EM_0. Using J we can then form $\sum_{\sigma \in \mathrm{FTS}} N_\sigma$ and $\prod_{\sigma \in \mathrm{FTS}} N_\sigma$, which is the first move to transfinite types, at type level ω. Then by successively applying the operations $\times$ and $\to$ again we can move up to level $\omega\cdot 2$, then $\omega\cdot 3, \ldots, \omega^2$, etc. A general pursuit of this would be based on a theory of ordinals, which are treated in terms of well-founded trees ("tree ordinals") in constructive mathematics and based on IG in the framework of T_0. This is taken up next. In any case we see that the passage to lower transfinite types can be effected in $EM_0 + J$. Precise limits for this are provided by the proof theory of $EM_0 + J$ (Part V).

12. Consequences of the inductive generation axiom. Here we move to full T_0.

12.1 Tree ordinals. Define $\underline{0} = (0,0)$, $x^+ = (1,x)$, $\sup_a f = (2,a,f)$. These are distinct. Note $\sup_{\mathbb{N}} f = (2,\mathbb{N},f)$. $\mathcal{O}_1$ is inductively generated as the least class such that

$$\underline{0} \in \mathcal{O}_1,\ x \in \mathcal{O}_1 \to x^+ \in \mathcal{O}_1, \text{ and } (f\colon \mathbb{N} \to \mathcal{O}_1) \to \sup_{\mathbb{N}} f \in \mathcal{O}_1 .$$

Then $\mathcal{O}_1 = i(A,R)$ for suitable A, R. Classically the members of $\mathcal{O}_1$ represent countable ordinals with

$$|\underline{0}| = 0,\ |x^+| = |x| + 1,\ |\sup_{\mathbb{N}} f| = \sup_{n \in \mathbb{N}} |f(n)| + 1 .$$

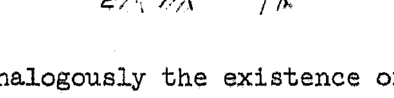
(Note that x^+ can be dropped in favor of $\sup_{\mathbb{N}} \lambda n(x)$.) We can picture members of $\mathcal{O}_1$ as well-founded trees:

Definition by recursion on $\mathcal{O}_1$ is a consequence of the recursion theorem. $\mathcal{O}_2$ is inductively generated as the least class such that

$$\underline{0} \in \mathcal{O}_2,\ x \in \mathcal{O}_2 \to x^+ \in \mathcal{O}_2,\quad (f\colon \mathbb{N} \to \mathcal{O}_2) \to \sup_{\mathbb{N}} f \in \mathcal{O}_2$$

$$\text{and}\quad (f\colon \mathcal{O}_1 \to \mathcal{O}_2) \to \sup_{\mathcal{O}_1} f \in \mathcal{O}_2 .$$

The last is pictured by



Then we can obtain analogously the existence of a class $\mathcal{O}_n$ for each $n \in \mathbb{N}$. Using join we can carry this on to define transfinite tree classes: $\mathcal{O}_a$ for $a \in \mathcal{O}_1$, and more generally for $a \in \mathcal{O}_b$ with any given b. (For more details on this material cf. Feferman 1975 pp.99-100, also Feferman 1978 §5.)

12.2 IG and Borel sets. For simplicity we indicate the treatment in Baire space $\mathbb{N}^{\mathbb{N}}$. Let $s_n (n \in \mathbb{N})$ be an enumeration of all finite sequences in $\mathbb{N}$ and $G_n = \{g \in \mathbb{N}^{\mathbb{N}} \mid \bar{g}(\ell h(s_n)) = s_n\}$; these are the basic clopen sets. We take $\mathcal{B}$ to be the smallest class such that each $G_n \in \mathcal{B}$ and if $f : \mathbb{N} \to \mathcal{B}$ then $\bigcap_{n\in\mathbb{N}} fn \in \mathcal{B}$ and $\bigcup_{n\in\mathbb{N}} fn \in \mathcal{B}$. In more general spaces we can follow Bishop's treatment via complemented sets. In any case IG suffices for this.

It would be appropriate at this point to take up the questions of adequacy and accord of T_0 and its subtheories with BCM. However we first complete our discussion of the consequences of various axioms with a look at D_V and POW (which are only marginally related to actual BCM).

13. Non-extensionality as a consequence of D_V.

13.1 Non-extensionality of operations. The following is proved in Feferman 1975 §3.4 as a consequence of $APP + D_V$:

$$\neg \forall f,g[\forall x(fx \simeq gx) \to f = g].$$

The idea of the proof is first (using D_V) to associate with each f an f^* such that $Df = Df^*$ and $f^* x \simeq 0$ whenever $x \in Df$. Then, if extensionality is assumed we have f total (i.e. $\forall x[fx\downarrow]$) iff $f^* = \lambda x(0) = 0^*$. Again using full definition by cases D_V we obtain a total operation e such that $Tot(f) \leftrightarrow ef = 0^*$. Diagonalization produces a contradiction.

13.2 Consistency of extensionality of operations in T_0. Denote by EXT_{Op} the statement $\forall f,g[\forall x(fx \simeq gx) \to f = g]$. Using extensional term models for APP (due to Barendregt) it will be shown in Part III that $T_0 + EXT_{Op}$ is consistent. Hence the use of D_V in 13.1 is essential. (It will also be shown that $T_0 + D_V$ is consistent.)

13.3 Non-extensionality of classes. Denote by $EXT_{C\ell}$ the statement $\forall A,B [\forall x(x \in A \to x \in B) \to A = B]$. It is also proved in Feferman 1975 §3.4 that $\neg EXT_{C\ell}$ holds under the assumptions $APP + ECA + D_V$. The idea is to associate with each f the class $cf = \{x \mid fx\downarrow\}$; c itself is total. Then if $EXT_{C\ell}$ held we would have $Tot(f) \leftrightarrow cf = V$, from which one can proceed as in 13.1. With reference to 13.2 we have the following.

Question: Is $T_0 + EXT_{C\ell}$ consistent?
We can of course ask similar questions for the addition of $EXT_{C\ell}$ to sub-theories like EM_0, S_0 etc., for all of which the answer is not known.

13.4 Discussion. D_V is a perfectly reasonable axiom if we regard the entities of our universe as being syntactic objects and = as literal identity. It is less evident if the entities are viewed as mental objects and = is interpreted as intensional identity; however, it appears from writings of Kreisel and of Troelstra (cf. Troelstra 1975) that here also D_V is to be accepted. Then, far from being disturbing, the results of 13.1 and 13.3 add support to the basic non-extensional viewpoint of constructive mathematics as presented in I.4.5, 4.11.

14. Status of the power-class axioms.

14.1 Inconsistency of POW with Join. To be more precise it is shown that APP + ECA + J proves ¬POW, the weak power-class axiom of §8 above. Indeed, suppose that there is a weak power-class C of V, so $\forall x[x \in C \to C\ell(x)] \wedge \forall X \exists Y(Y \in C \wedge X \equiv Y)$. Let $B = \sum_{x \in C} x$ and $A = \{x | x \in C \wedge (x,x) \notin B\}$. Then $A \equiv a$ for some a in C, and $a \in a \leftrightarrow a \in A \leftrightarrow a \in C \wedge (a,a) \notin B \leftrightarrow a \in C \wedge (a \notin a) \leftrightarrow a \notin a$, which gives a contradiction.

14.2 Consistency of POW with EM_0. This will be proved in Part III. It may also be shown that S_0 + POW is consistent where S_0 is the theory of sets described in 9.2. Even though S_0 contains J we cannot derive a contradiction as in 14.1, since we don't have a universal class V in S_0.

The consistency of further axioms introduced in § 7 above (such as 2nd order comprehension) will also be taken up in Part III.

15. Adequacy of (subtheories of) T_0 to BCM.

15.1 Adequacy of T_0. The development outlined in §§10-12 provides a basis for the formalization of BCM in T_0. Moreover, this is accomplished by following the informal mathematics as explained in I.14 and I.15. The official intended definitions come to the forefront in the process of formalization and must always be kept in mind. When informal concepts and proofs are spelled out accordingly, one is in a position from which formalization in T_0 can proceed directly. (This was illustrated in I.15.4). One may thus conclude that T_0 is directly adequate to BCM (as exemplified in Bishop 1967). It is of logical interest to see next how much of BCM can be carried out in theories weaker than T_0.

15.2 The role of IG. Obviously IG is used only for the theory of Borel sets in Bishop 1967, which in turn figured in the theory of measure and integration. As was explained in I.14.13, this was superseded by a treatment without Borel sets in Bishop-Cheng 1972. The latter makes prima facie use of the axiom POW, but just to form (a complete integration space) L(X) as a class from any integration space X; however integration theory only requires the notion of f being a member of L(X), which is definable without the assumption of POW.[8] The conclusion is that IG is unnecessary for the development of abstract integration theory in this sense.

8) The role of axioms like POW in abstract constructive integration theory is studied in Feferman 1978 §4.3. A modified form of POW for this purpose can be derived in the theory of sets and classes $T_0(S)$ (cf. 9.3 above); this yields the class of subsets of any given class.

15.3 Dispensability of the join axiom. We have seen in 11.3 that the axiom J is needed to effect the passage to transfinite types (e.g. to $\sum_{\sigma\in FTS} N_\sigma$ and $\prod_{\sigma\in FTS} N_\sigma$). However in actual analysis one never considers families of varying type but only families of subsets of a given set. In these cases one can try to eliminate J by replacing the notion of family by that of pre-joined family (11.1). It may be verified that, except for the theory of Borel sets, this replacement does indeed leave the treatment of analysis in BCM unaffected.

15.4 Adequacy of restricted induction on $\mathbb{N}$. An example where unrestricted induction on $\mathbb{N}$ was used in an essential way was given in 11.3, namely to prove that for all $\sigma \in FTS$, N_σ is a class. Similarly, the principle of unrestricted induction in IG is used only to show that the objects in the Borel hierarchy actually are classes. But for BCM without transfinite types and without Borel theory it appears that only restricted induction on $\mathbb{N}$ is needed. This has been verified in detail by Friedman (unpublished, but cf. 18.2 below).

15.5 Adequacy of $EM_0\restriction$. Putting 15.1 - 15.4 together we conclude that $EM_0\restriction$ is adequate to all of Bishop 1967 except for that part involving the theory of Borel sets, and to all of Bishop-Cheng 1972 except for treating L(-) as an operation from classes to classes. This is of logical (and epistemological) interest because, as will be shown in Part V, $EM_0\restriction$ is conservative over HA.

16. The question of accordance of T_0 (or its subtheories) with BCM.

16.1 Sets vs. classes. Bishop does not speak of classes and it is questionable whether he would countenance a universal class V. In this respect, T_0 is not explicitly in accordance with BCM. The theory of sets S_0 (9.2) is here in greater direct accord. Incidentally, S_0 is adequate to the same part of practice as T_0.

16.2 The question of operations with unbounded domains. There is no explicit discussion by Bishop of operations with unbounded domains like k, s and the resulting $d = \lambda x(x)$, $e = \lambda x \lambda y(xy)$, etc. However, the idea of such does seem to be implicit in his view of operations simply as rules; it is further implicit in his use of operations such as Cartesian product and power on sets, since no class of all sets is assumed as an object. It is my conclusion from these arguments that the use of operations with unbounded domains is implicitly in accordance with BCM. However, this is clearly subject to debate, especially since it leads us to talk about combinations like (xx) which appear foreign to practice.

Remark. There is a simple formal device which permits us to replace unbounded combinatory operations by corresponding bounded ones and still achieve much the same mathematical effects. Namely, one introduces formal "external" operation symbols on (variable) classes (A,B,...) e.g. $k_{A,B}$, id_A, $e_{A,B}$ etc. (writing the arguments as subscripts) with axioms like:

$k_{A,B} \in (A \to (B \to A))$, $\forall x \in A \forall y \in B\ (k_{A,B}xy = x)$,

$id_A \in (A \to A)$, $\forall x \in A(id_A(x) = x)$,

$e_{A,B} \in (A \to ((A \to B) \to B))$, etc.

What is lost here is the possibility of reducing recursion (on $\mathbb{N}$, or any $i(A,R)$) to the combinators, since those reductions make essential use of the possibility of self-application (xx). Thus in such a step one must supplement the $\mathbb{N}$, resp. IG axioms, by suitable axioms for recursion operators.

16.3 The other principles. These are comprehension, natural numbers, join and inductive generation. If we are to judge the axioms for these separately from the issues in 16.1, 16.2, we must naturally consider them in weaker forms that apply as well to sets and are given by external rather than internal operations. In particular CA_1 is to be replaced by SEP_1. With such modifications in mind, it should be clear from I.14, 15 and 10-12 above that these principles are called for in BCM.

Remark. Beeson has also raised a question (in conversation) about the constructivity of the join axoim, as formulated uniformly using j. His point is that the result should depend not only on A and f but also on a proof of $\forall x \in A[C\ell(fx)]$.

16.4 Conclusion. The issues in dispute are those in 16.1 and 16.2. I believe a case can be made - based on Bishop's views of operations given by rules and sets by properties - that the use of both operations with unbounded domains and of classes (as well as sets) makes T_0 implicitly in accordance with BCM. However, there is little support for explicit, direct accordance.

Remark. Since $EM_0\restriction$ is conservative over HA and the latter is certainly in direct accordance with BCM, the former is consequently in indirect accordance with it.

17. Comparison with Martin-Löf 1975.

17.1 Character of Martin-Löf's system. This is a kind of logic-free transfinite type theory which is denoted TT.[9] There are terms for objects a,b,c,..., and terms for types A,B,C,... . The informal idea is that each object is of a unique type. The basic propositions are of the form

$$a \in A \quad \text{and} \quad a = b\ ,$$

where $a \in A$ is read: a is of type A. TT is based on a natural deduction system (cf. Prawitz 1971) for deriving such propositions from hypotheses of the form $x_i \in A_i$. For example, suppose one has inferred $b[x] \in B[x]$ from $x \in A$.[10]

9) As stated by Martin-Löf, a significant earlier attempt to formulate such a theory was made in Scott 1970.

10) We simplify here the form of assumptions actually given by Martin-Löf for TT.

Then we have terms for application, abstraction and Cartesian product related by the rules

$$\lambda x\, b[x] \in (\Pi x \in A)\, B[x], \quad \frac{a \in A}{(\lambda x\, b[x])(a) = b[a]}, \quad \text{and} \quad \frac{a \in A,\; c \in (\Pi x \in A) B[x]}{c(a) \in B[a]}.$$

Similarly there are rules for pairing, projection and join $(\Sigma x \in A)\, B[x]$. Special cases of product and join are $(A \to B)$ and $A \times B$. There are rules for the natural numbers N using 0, s (successor) and recursion on N. (Finite initial segments N_k of N are provided for too.) Further, with each A is associated the identity relation I on A as a function of $(x,y) \in A \times A$. Finally, there is a V_o which is supposed to be the type of all small types, and is closed under the introduction rules for types; moreover, there is for each V_n a corresponding V_{n+1}. To prove that A is a type in the system one proves $A \in V_n$ for some n. The syntax of the predicate calculus is represented in TT via the (Curry-Howard) correspondence between formulas and types. When a type A is thought of as a proposition then $(a \in A)$ is thought of as 'a is a proof of the proposition A'. From this, the intuitionistic predicate calculus is derived using the (Brouwer-Heyting) explanation of the logical operators in terms of proofs (I.4.2 above)

Remark. Logic is assumed informally in the explanation of the rules.

17.2 Comparison of the system with T_o. TT does not provide for inductively generated types in general, but rules for them can be adjoined along the same lines, following Martin-Löf 1971. With or without such rules, the system can be interpreted in T_o (each V_n is inductively generated by certain closure conditions involving $V_o, \ldots, V_{n-1}$). Furthermore, the system with no universes contains $EM_o\!\restriction$.[11)]

17.3 Adequacy of TT to BCM. By the preceding, TT (as given by Martin-Löf) is adequate to the same portion of BCM as $EM_o\!\restriction$; when supplemented by inductively generated types as suggested in 17.2 it is also adequate to the same portion of BCM as all of T_o actually serves to formalize.

17.4 Accordance with BCM. The types of TT can be interpreted as sets in Bishop's sense. Following I.14 - I.15 above it should be granted that TT is in direct accordance with BCM, at least insofar as concerns basic concepts and principles. The one reservation has to do with its heavily syntactic formulation for the conditions to introduce and use the various kinds of terms. This is in turn necessitated by the requirement that each object is assigned a type. Thus we cannot have an 'internal' function f of which it is proved $\forall x \in A[fx$ is a type$]$ (as done in T_o by $\forall x \in A[C\ell(fx)]$) but must use 'external' objects $B[x]$ of which $B[x] \in V_n$ is proved (for some n) under the hypothesis $x \in A$. There are no indications in Bishop's writings that would lead one necessarily to take such a formal approach. In this respect, the looseness which T_o enjoys owing to its type-free character seems more in accord with BCM.

11) The exact relationships are not known to me.

Remark. The syntax of TT is evidently somewhat more complicated than that of T_0. Some simplification could presumably be made by assuming all of intuitionistic logic at the outset. In any case, it is much easier to form a variety of models and interpretations of T_0, as we shall see in Parts III, IV.

18. Comparison with Myhill's and Friedman's extensional systems.

18.1 The character of these systems. The system CST introduced in Myhill 1975 is a subsystem of IZF(N) + DC by which is meant Zermelo-Fraenkel set theory over the natural numbers (as urelements) with the logic restricted to be intuitionistic and with the axiom scheme of dependent choices added. The notions of pair and of function are both defined here just as usual in ZF : in other words functions are identified with graphs of many-one relations. One takes (over the usual axioms of N) the axioms of extensionality, unordered pair, union, Δ_0-separation, domain and ranges of functions, the set $(A \to B)$ of all functions $f:A \to B$ for given sets A,B (which is taken in place of the power set axiom) and the replacement scheme

$$(\forall x \in A)\exists! y\, \phi(x,y) \to \exists z[\mathrm{Fun}(z) \wedge \mathrm{Dom}(z) = A \wedge \forall x \in A\, \phi(x, z(x))].$$

Finally, the principle DC

$$(\forall x \in A\, \exists y \in A\, \phi(x,y) \to \forall x \in A\, \exists z[z:N \to A \wedge z(0) = x \wedge \forall n \in N\, \phi(z(n), z(n'))]$$

is taken, but not AC, since that is shown to contradict Church's thesis in the system.[12)]

Friedman 1977 considers a number of subsystems and extensions of CST (all contained in IZF/N). The weakest of these is denoted $\underset{\sim}{B}$. In $\underset{\sim}{B}$, induction on N is restricted, replacement is taken only to form $\{\{x \in A \mid \phi(x,y)\} \mid y \in A\}$ for Δ_0 formulas ϕ, and DC is also taken only for such ϕ. The other systems considered are denoted T_1, T_2, T_3 and T_4. We shall not describe them here;[13)] however CST is equivalent to T_2.

Remarks. (i) Axioms of inductive generation are not taken in CST. They are derivable in T_4.

(ii) Friedman 1977 shows that $\underset{\sim}{B}$ itself is reducible to HA and is conservative for Π^0_2 sentences. (Beeson 1979 shows that it is conservative for all sentences.) By contrast, $\underset{\sim}{B}$ + classical logic is equivalent to Zermelo set theory.

18.2 Adequacy of these systems to BCM. The system $\underset{\sim}{B}$ is adequate to the same portion of BCM as the system $EM_0\restriction$ (cf. the discussion in Friedman 1977 p.7). Though formally stronger, the system CST does not seem to have any further power for the actual mathematics involved. The adequacy in both cases is indirect. The definitions of concepts do not follow Bishop's official spelled-out definitions, but rather the corresponding classical ones which use extensionality. For example,

12) The reason why AC but not DC is problematic in the framework of T_0 will be explained in Part IV below.

13) The reader may find it useful to read my review of Friedman 1977 which appeared in Math. Reviews 55(1978) No.7748.

the real numbers $\mathbb{R}$ are taken to be equivalence classes of Cauchy sequences. The positive real numbers $\mathbb{R}^+$ are those $x \in \mathbb{R}$ such that $(\exists n \in \mathbb{N}^+)(x > \frac{1}{n})$. Thus all the distinctions and use of witnessing data required by Bishop in order to carry out constructive operations are essentially ignored.

18.3 The question of accordance. These systems differ in two essential respects from the constructive point of view which is basic to BCM, at least as described in I.4. Namely, extensionality is accepted, in violation of I.4.5, and functions are defined in terms of sets (of ordered pairs), in violation of I.4.6. It is plain then that any set theory which contains the extensionality axiom and defines the notion of function in this way - and in particular CST and $\underset{\sim}{B}$ - is not in direct accordance with BCM.

By the reduction of $\underset{\sim}{B}$ to HA due to Friedman (and Beeson) referred to above, $\underset{\sim}{B}$ is certainly indirectly in accordance with BCM.[14)] As to CST, Myhill 1975 gives a constructive reduction via a realizability interpretation. More sharply, Friedman 1977 obtains reduction of the equivalent T_2 to intuitionistic ramified analysis in levels $< \epsilon_0$ (which in turn is interpretable in our T_0 minus IG, cf. Part V below). The system T_3 is also reduced loc. cit. to an intuitionistic theory of one inductively defined set, which is certainly justified by BCM and is contained in our T_0. Finally, the theory T_4 is reducible to the full 2nd order theory of species which is contained in $EM_0 + CA_2$; but the accordance of the latter with BCM is open to dispute.

It should be mentioned that in Beeson's contribution to this volume he shows for a number of intuitionistic extensional theories of sets how to interpret them in their subtheories without extensionality. This is followed in Beeson 1979 by certain realizability interpretations to reduce the latter theories to sub-theories of T_0, in particular of $\underset{\sim}{B}$ to HA (conservatively).

III. Models

Throughout this part models will be understood in the usual set-theoretical sense and thus will satisfy classical logic. This does not hold for the interpretations to be dealt with in Part IV.

1. A model of T_0 over any model of APP (presented in Feferman 1975 sec. 4.1.) Let

$$\mathfrak{A} = \langle V, App, k, s, d, p, p_1, p_2, 0, s_N, p_N \rangle$$

be any model of the axioms APP of T_0 (in II.5). Here $x \in \mathbb{N}$ is interpreted as $x \in \omega$ where ω is the least subset of V containing 0 and closed under $x \to x' = s_N x$. (The identification of $\mathbb{N}$ as a member of V will be explained in a moment.) Abbreviations for application terms, pairing, comprehension are

14) Friedman also gives an informal argument for the constructive justification of $\underset{\sim}{B}$ (and stronger theories) by interpretation in a theory of species of finite type.

taken just as in II.3,4. Now we take codes in V for the class constants and operations, e.g. as follows:

$$\mathbb{N} = (0,0),\ c_n z = (1,n,z),\ j(a,f) = (2,a,f) \text{ and } i(a,r) = (3,a,r).$$

Next $C\ell_\alpha$ and ϵ_α are defined by transfinite recursion on α; at stage α one has a structure $(\mathfrak{A}, C\ell_\alpha, \epsilon_\alpha)$ in which the formulas of $\mathcal{L}(T_o)$ are interpreted by taking ϵ_α for 'ϵ' and letting the class variables range over $C\ell_\alpha$. In this definition we shall also use 'ϵ' in its ordinary set-theoretic extensional sense; the context serves to avoid ambiguity.[15)]

(1) $C\ell_o = \{\mathbb{N}\}$ and $x \epsilon_o \mathbb{N} \leftrightarrow x \epsilon \omega$.

(2) (i) $C\ell_\alpha \subseteq C\ell_{\alpha+1}$ and $x \epsilon_{\alpha+1} a \leftrightarrow x \epsilon_\alpha a$, for $a \epsilon C\ell_\alpha$;

(ii) for each elementary $\phi(x,y_1,\ldots,y_m, Z_1,\ldots,Z_p)$ and $n = \ulcorner\phi(x,\underline{y},\underline{Z})\urcorner$, and for any $y_1,\ldots,y_m \epsilon V$ and $a_1,\ldots,a_p \epsilon C\ell_\alpha$ we have $c = c_n(y_1,\ldots,y_m, a_1,\ldots,a_p) \epsilon C\ell_{\alpha+1}$ and

$$x \epsilon_{\alpha+1} c \leftrightarrow (\mathfrak{A}, C\ell_\alpha, \epsilon_\alpha) \models \phi(x,y_1,\ldots,y_m, a_1,\ldots,a_p);$$

(iii) for each $a \epsilon C\ell_\alpha$ and $f \epsilon V$ such that $\forall x(x \epsilon_\alpha a \rightarrow fx \epsilon C\ell_\alpha)$, we have $c = j(a,f) \epsilon C\ell_{\alpha+1}$ and $z \epsilon_{\alpha+1} c \leftrightarrow \exists x,y(x \epsilon_\alpha a \wedge y \epsilon_\alpha fx)$;

(iv) for each $a \epsilon C\ell_\alpha$ and $r \epsilon C\ell_\alpha$ we have $c = i(a,r) \epsilon C\ell_{\alpha+1}$ and $x \epsilon_{\alpha+1} c \leftrightarrow \forall I \subseteq V\{\forall u[u \epsilon_\alpha a \wedge \forall w((w,u)\epsilon_\alpha r \rightarrow w \epsilon I) \rightarrow u \epsilon I] \rightarrow x \epsilon I\}$;

(v) $C\ell_{\alpha+1}$ has only those elements obtained by (i)-(iv).

(3) For limit λ, $C\ell_\lambda = \bigcup_{\alpha<\lambda} C\ell_\alpha$ and $\epsilon_\lambda = \bigcup_{\alpha<\lambda} \epsilon_\alpha$.

For the final model of T_o we take

(4) $C\ell = \bigcup_\alpha C\ell_\alpha$ and $\epsilon = \bigcup_\alpha \epsilon_\alpha$, so that for $a \epsilon C\ell_\alpha$ and any x, $x \epsilon a \leftrightarrow x \epsilon_\alpha a$.

The axioms for T_o are verified to hold in $\mathfrak{B} = (\mathfrak{A}, C\ell, \epsilon)$ in a straightforward way. In particular, by (1) we have full induction on $\mathbb{N}$ (with respect to any properties) and by (2)(iv) we have full induction on $i(A,R)$ for any A,R. It is further to be noted that in checking elementary comprehension for $c = \{x|\phi(x,y_1,\ldots,y_m, a_1,\ldots,a_p)\}$ we need only know the meaning of $x \epsilon a_i (1 \leq i \leq p)$. This is where the predicative character of ϕ is used in an essential way. A new idea is needed if one wishes to satisfy stronger comprehension schemes; that is explained in the next section.

Remarks. (i) V, defined by $\{x|x=x\}$, is interpreted as a certain code c_n in the domain V of $\mathfrak{A}$; then $x \epsilon V$ has the same meaning in the model as extensionally.

15) In Feferman 1975 we used the symbol 'η' in place of 'ϵ' in $\mathcal{L}(T_o)$ in order to distinguish the two uses.

(ii) If the constant d of $\mathfrak{A}$ satisfies $\forall x,y,a,b[dxyab\downarrow]$ and $(x=y \rightarrow dxyab = a) \wedge (x \neq y \rightarrow dxyab = b)$ then the full definition-by-cases axiom D_V is of course satisfied so we have a model of $T_0 + D_V$ in this case.

(iii) There is an obvious modification of the construction above to get a model of T_0 starting with any $C\ell_0$ and ϵ_0 such that $\mathbb{N} \in C\ell_0$ and $x \epsilon_0 \mathbb{N} \leftrightarrow x \in \omega$, as long as card $(C\ell_0) \leq$ card (V) (so that there is room for all the codes).

2. Modification to obtain a model of $T_0 + CA_2$. (Feferman 1975, Addendum). The argument here will be much less constructive. Starting with any model $\mathfrak{A}$ of APP as in §1, let $\mathfrak{M} = (\mathfrak{A}, \mathcal{P}(V), \epsilon)$ where $\mathcal{P}(V)$ is the set of all subsets of V and ϵ is the standard membership relation. To each stratified formula $\psi(x,\underline{y},\underline{Z},X)$ is assigned a Skolem function $F_\psi(x,\underline{y},\underline{Z}) = X$ which has the property:

$$\mathfrak{M} \models \{\exists X\, \psi(x,\underline{y},\underline{Z},X) \rightarrow \psi(x,\underline{y},\underline{Z},F_\psi(x,\underline{y},\underline{Z}))\}\ .$$

Given any stratified $\phi(x,\underline{y},\underline{Z})$ take $\psi(x,\underline{y},\underline{Z},X) = \forall x[x \in X \leftrightarrow \phi(x,\underline{y},\underline{Z})]$, and $G_\phi = F_\psi$, so $G_\phi(\underline{y},\underline{Z}) = \{x \in V \mid \mathfrak{M} \models \phi(x,\underline{y},\underline{Z})\}$. We choose codes f_n for the F_ψ for each stratified ψ and in place of (2)(ii) in §1 take, for $n = \ulcorner\psi(x,\underline{y},\underline{Z},X)\urcorner$:

(2)(ii)' for each $x, y_1,\dots,y_m \in V$ and $a_1,\dots,a_p \in C\ell_\alpha$ we have
$c = f_n(x,y_1,\dots,y_m, a_1,\dots,a_p) \in C\ell_{\alpha+1}$ and
$x \epsilon_{\alpha+1} c \leftrightarrow x \in F_\psi(x,y_1,\dots,y_m, A_1,\dots,A_p)$, where $A_i = \{x \mid x \epsilon_\alpha a_i\}$ for each $i=1,\dots,p$.

With the resulting $(\mathfrak{A},C\ell,\epsilon)$ then defined as in §1(4), let a^* be the extension $\{x \mid x \epsilon a\} \subseteq V$ for each a in $C\ell$, and let $C\ell^* = \{a^* \mid a \in C\ell\}$. Then $C\ell^*$ is closed under the F_ψ and so $\mathfrak{M}^* = (\mathfrak{A}, C\ell^*, \epsilon)$ is an elementary substructure of $\mathfrak{M}$. It follows that for each stratified ϕ and any sets $A_1,\dots,A_p$ we have $\{x \mid \mathfrak{M}^* \models \phi(x,y_1,\dots,y_m, A_1,\dots,A_p\}$ $(= G_\phi(\underline{y}, A))$ in $C\ell^*$. Finally it is proved by induction on stratified ϕ that

$$(\mathfrak{A}, C\ell, \epsilon) \models \phi(x,\underline{y},\underline{a}) \leftrightarrow \mathfrak{M}^* \models \phi(x,\underline{y},\underline{a}^*).$$

From this it follows that $(\mathfrak{A}, C\ell, \epsilon)$ is a model of $T_0 + CA_2$.

Remarks. The axioms for $\mathbb{N}$ and IG could be subsumed under CA_2 since their 2nd order definitions in $\mathfrak{M}^*$ are absolute.

3. The recursion-theoretic model. Take $V = \omega$ and $App(x,y,z) \leftrightarrow \{x\}(y) \simeq z$ in the sense of ordinary recursion theory. Taking 0 and $x \mapsto x'$ to be standard, we can easily choose constants k,s,d,p,p_1,p_2,S_N,P_N so as to obtain $\mathfrak{A}$ satisfying the axioms APP. Note that D_V is automatically satisfied. In addition then to T_0 (or $T_0 + CA_2$ if we follow §2) we also have Church's thesis for partial functions

(CT_2) $\qquad \forall f\, \exists e\, \forall n(fn \simeq \{e\}(n))$

true in $(\mathfrak{A}, C\ell, \epsilon)$ simply by taking $e = f$. $\mathbb{N}^{\mathbb{N}}$ is just the class of recursive functions. Hence CT_1 is also true. On the other hand CT_0, being classically false, is not satisfied.

With reference to II 10.8-10.9, it may be seen that $\langle N_\sigma \rangle_{\sigma \in FTS}$ is interpreted in this model as the hierarchy HRO of <u>hereditarily recursive operations</u> and $\langle M_\sigma \rangle_{\sigma \in FTS}$ as the hierarchy HEO of <u>hereditarily effective operations</u> (cf. Troelstra 1973, 124-127). Going on to 10.11 - 10.12 one sees that the reals $\mathbb{R}$ are interpreted as in <u>recursive analysis</u>, and so on for $C([a,b], \mathbb{R})$, etc. Finally, with reference to 12.1, it is seen that $\mathfrak{O}_1, \mathfrak{O}_2, \ldots, \mathfrak{O}_a, \ldots$ are interpreted as forms of the Church-Kleene <u>constructive ordinal notation classes</u>.

<u>Remark</u>. Any "enumerative" generalization of recursion theory gives rise to a model of APP which, when extended to a model of T_0 as in § 1 yields other interesting interpretations of its concepts. For further examples of such and §§5-6 below cf. Feferman 1978, 3.2 - 3.4.

4. <u>Independence results from BCM</u>. Since all of BCM can be safely formalized in T_0, any model $\mathfrak{B} = (\mathfrak{A}, C\ell, \epsilon)$ of T_0 automatically provides independence results for ϕ which are classically true but for which $\mathfrak{B} \not\models \phi$. For example, if we take the recursion-theoretic model $\mathfrak{A}$ of § 3 to begin with then the example due to Specker of a recursively (uniformly) continuous function on [0,1] which does not take on a recursive minimum shows that <u>the theorem of the minimum is underivable in BCM</u>. Indeed, to be more precise and even stronger, by §2 <u>it is not derivable in $T_0 + D_V + CA_2$ with classical logic</u>. Similarly for the other examples giving 'peculiarities' of recursive analysis and of the Russian school of constructive analysis (cf. I.7-8).

<u>Remark</u>. The obverse of the point here is that if ϕ is a mathematical statement for which (classical) $T_0 + D_V + CA_2 \vdash \phi$ then ϕ has a recursion-theoretic interpretation or 'analogue'.

5. <u>Generating models of</u> $APP + D_V$. Given any infinite set V, we can generate a model $\mathfrak{A}$ of APP from the following information: (i) a pairing operation $P : V^2 \xrightarrow{1-1} V$ and projections $P_i : V \to V$ for which $P_i(P(x_1, x_2)) = x_i$, (ii) an embedding of $(\omega, 0, ')$ in V, and (iii) any collection $\mathfrak{F}$ of partial $F : V \rightharpoonup V$ for which card $(\mathfrak{F}) \leq$ card (V). Then we can define constants for $\mathfrak{A}$ so that $APP + D_V$ is satisfied and $pxy = P(x,y)$, $p_i x = P_i(x)$, $s_N n = n'$, $(p_N n') = n$ and such that for each $F \in \mathfrak{F}$ there exists $f \in V$ which represents F, i.e. $fx \simeq F(x)$ for all x. (By non-extensionality, each F will have many representations.) To obtain $\mathfrak{A}$ we simply use pairing to build codes for the constants $k, s, \ldots$ of $\mathfrak{A}$ as well as for each $F \in \mathfrak{F}$. Then we regard the axioms of APP as inductive closure conditions on the relation $App(x,y,z)$. In particular (seeing to it that sx and sxy are always defined in a simple way) one wants

$$xz \simeq u \wedge yz \simeq w \wedge uw \simeq v \to (sxy)z \simeq v.$$

6. Full set-theoretic models of $APP + D_V$. In particular, let $V = R_\lambda$ (the set of sets in the cumulative hierarchy of rank $< \lambda$) for some limit $\lambda > \omega$. Define $0, '$ and predecessor on ω and pairing and projections as usual. Let $\mathfrak{F}$ be the class of all functions which (as sets) are members of R_λ. By § 5 we obtain a model $\mathfrak{A}$ of $APP + D_V$ in which every set-theoretic function is represented. Proceed to build a model $\mathfrak{B} = (\mathfrak{A}, C\ell, \varepsilon)$ of $T_0 + D_V + CA_2$ over $\mathfrak{A}$. In $\mathfrak{B}$, $\mathbb{N}$ has the same elements as ω and $(\mathbb{N} \to \mathbb{N})$ consists of representatives of all functions from ω to ω. For the type symbols $1 = (0 \dot{\to} 0)$, $2 = (1 \dot{\to} 0)$, etc. we have $(M_1 / =_1) \cong (\omega \to \omega)$ and $(M_2 / =_2) \cong ((\omega \to \omega) \to \omega)$, etc. Further $(\mathbb{R} / =_{\mathbb{R}})$ is isomorphic to the reals in the set-theoretical sense, and the class of all functions from $\mathbb{R}$ to $\mathbb{R}$ which preserve $=_{\mathbb{R}}$ is isomorphic (modulo the defined equality between such functions) with the set of all real functions in the set-theoretic sense. Now $C([a,b], \mathbb{R})$ consists of representatives of all uniformly continuous functions.

Suppose λ is inaccessible. Each element a of $\mathfrak{O}_1$ has a naturally associated ordinal $|a| < \omega_1$ and $\omega_1 = \{|a| : a \in \mathfrak{O}_1\}$. More generally for any $\mathfrak{O}_a$, we have $\omega_{|a|} = \{|b| : b \in \mathfrak{O}_a\}$. The Borel hierarchy in $\mathbb{N}^{\mathbb{N}}$ as explained in II.12.2 consists of representatives of the full Borel hierarchy in Baire space in the set-theoretic sense.

7. Generalizing classical, recursive and constructive mathematics.

7.1 It follows from § 3 and § 6 that any mathematical theorem ϕ of $T_0 + D_V + CA_2$ with classical logic automatically generalizes a theorem of recursive mathematics and of classical set-theoretic mathematics.

7.2 It also follows that for any sub-theory T of $T_0 + D_V + CA_2$ which is recognized as being constructively valid (so, the logic may be restricted) any mathematical theorem ϕ of T generalizes one from classical, recursive and constructive mathematics. In particular, this applies to $T = T_0$ (if II.15.4 is accepted).

Remarks. (i) In a certain sense Martin-Löf's TT can also be considered to have both set-theoretic and recursion-theoretic models, so 7.2 would also apply to it. (ii) Myhill's CST (and related theories) has immediate set-theoretic models, but no direct recursion-theoretic model and, as we have seen in II.17, its constructive interpretation is in dispute.

8. Term models. In the framework of T_0 these have been given by Beeson 1977 (1.3) which is followed here; however, the ideas are familiar from combinatory calculi (cf. Barendregt 1971, 1977). As will be explained in 8.2, the method works to give a model $\mathfrak{A}$ of APP but not of $APP + D_V$.

8.1 Reduction of terms. Let $\tau, \tau_1, \tau_2, \ldots$ range over application terms as explained in II.1.3. A reduction relation $\tau_1 \geq \tau_2$ is defined inductively by the following clauses, where we write $\bar{n}$ for $0\underbrace{''\cdots'}_{n}$:

(i) $\tau \geq \tau$

(ii) $\tau_1 \geq \tau_2 \wedge \tau_2 \geq \tau_3 \rightarrow \tau_1 \geq \tau_3$

(iii) $\tau_1 \geq \tau_1^* \wedge \tau_2 \geq \tau_2^* \rightarrow \tau_1\tau_2 \geq \tau_1^*\tau_2^*$

(iv) $k\tau_1\tau_2 \geq \tau_1$

(v) $s\tau_1\tau_2\tau_3 \geq \tau_1\tau_3(\tau_2\tau_3)$

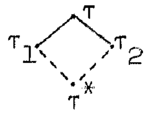
(vi) $p_1(p\tau_1\tau_2) \geq \tau_1$, $p_2(p\tau_1\tau_2) \geq \tau_3$

(vii) $p_{\mathbb{N}}(s_{\mathbb{N}}\tau) \geq \tau$

(viii) $d\bar{n}\bar{n}\tau_1\tau_2 \geq \tau_1$, $n \neq m \rightarrow d\bar{n}\bar{m}\tau_1\tau_2 \geq \tau_2$

(ix) $\tau_1 \geq \tau_2$ only as required by (i) - (viii).

We shall use $\tau_1 \geq \tau_2$ for literal identity of terms. τ_1 is said to be in <u>normal form</u> (<u>or irreducible</u>) if whenever $\tau_1 \geq \tau_2$ we have $\tau_1 = \tau_2$. The set of terms in normal form is denoted by NF. Note every term reduces to a term in NF. The <u>Church-Rosser theorem</u> (or $\Diamond$ property) for $\geq$ is proved by standard methods:

CR. If $\tau \geq \tau_1$ and $\tau \geq \tau_2$ then for some τ^*, $\tau_1 \geq \tau^*$ and $\tau_2 \geq \tau^*$.  .

As a corollary one has <u>unicity of normal form</u> in the sense that

$$\tau_1, \tau_2 \in NF \wedge \tau \geq \tau_1 \wedge \tau \geq \tau_2 \rightarrow \tau_1 = \tau_2 \ .$$

8.2 <u>The model of normal terms</u>. The domain V of the model $\mathfrak{A}$ taken by Beeson 1977 is NF (which is also used here to denote $\mathfrak{A}$). The application relation is:

$$A_{pp}(\tau_1, \tau_2, \tau_3) \leftrightarrow \tau_1, \tau_2, \tau_3 \in NF \wedge (\tau_1\tau_2 \geq \tau_3) \ .$$

The constants, all of which are in NF, denote themselves in this model. NF is a model of APP because all $\bar{n}$ are in NF and $\bar{n} \neq \bar{m}$ for $n \neq m$. The following easy lemmas are proved by Beeson, where $APP_{\mathbb{N}}$ denotes APP plus the $\mathbb{N}$-closure axiom $0 \in \mathbb{N} \wedge \forall x (x \in \mathbb{N} \rightarrow x' \in \mathbb{N})$.

(a) If $\tau \in NF$ then $APP_{\mathbb{N}} \vdash (\tau\downarrow)$.

(b) If τ_1, τ_2 are closed terms and $\tau_1 \geq \tau_2$ then $APP_{\mathbb{N}} \vdash (\tau_1\downarrow \rightarrow \tau_2\downarrow \wedge \tau_1 = \tau_2)$.

(c) If τ is closed and $APP_{\mathbb{N}} \vdash (\tau\downarrow)$ then $\exists \tau^* \in NF(\tau \geq \tau^*)$.

(d) If T is any sub-theory of $T_0 + CA_2$ with $APP_{\mathbb{N}} \subseteq T$ and if $T \vdash (\tau \in \mathbb{N})$ then for some $n \in \omega$, $T \vdash (\tau = \bar{n})$.

When coupled with realizability methods in Part IV the lemma (d) allows one to obtain the numerical instantiation property and the disjunction property for such T.

<u>Remark</u>. An essential difference of D_V from the other applicative axioms appears here. In the proof of Lemma (a) we use that if $n \neq m$ then $APP_{\mathbb{N}} \vdash \bar{n} \neq \bar{m}$. To

try to obtain a corresponding result for D_V we would want that if $\tau_1, \tau_2 \in NF$ and $\tau_1 \neq \tau_2$ then $APP_{\mathbb{N}} + D_V \vdash (\tau_1 \neq \tau_2)$. But that isn't so - for example, $k, 0 \in NF$ and $k \neq 0$ but $(k \neq 0)$ is not provable (since we can construct an applicative model $\mathfrak{A}$ in which $k = 0$). For the same reason we can't prove the disjunction property for $T_0 + D_V$.[16)]

8.3 An extensional term model (Barendregt 1971). Instead of taking V as in 8.2, one takes V to be the set of all equivalence classes $[\tau]$ of terms for the least equivalence relation $\equiv$ such that $\tau_1 \geq \tau_2 \rightarrow \tau_1 \equiv \tau_2$. Then we take $App([\tau_1],[\tau_2],[\tau_3]) \leftrightarrow \tau_1\tau_2 \equiv \tau_3$. By the CR property if an equivalence class contains some $\tau \in NF$, that τ is unique. Then one sees that $[\bar{n}] \neq [\bar{m}]$ whenever $n \neq m$ so that we obtain a model of APP. Since every application term τ denotes $[\tau]$ in the model, it satisfies $(\tau\downarrow)$, i.e. every operation here is total. In addition the model may be shown to satisfy the axiom of extensionality for terms. This shows D_V to be essentially required for the non-extensionality result of II.13.1.

9. Continuous function models. There are again models due to Beeson 1977 (1.2), once more without D_V.

9.1 Continuous partial function application. The idea here is to form a model of APP which is a kind of untyped version of the class of countable functionals of finite type (which are hereditarily continuous in a certain sense) due to Kleene 1959 (and Kreisel, same volume). One takes V to be the class of all partial functions f from ω to ω. A relation $App(f,g,h)$ (or $fg \simeq h$) is defined for members of V as follows. For each n, the value of h at n is supposed to depend on only a finite amount of information about g. Let $(f)_n = \lambda x.f(x,n)$ with (x,y) a primitive recursive pairing function. More precisely, $h(n)$ is obtained, when defined, by $(f)_n$ acting continuously on $(g)_n$ so that if $(f)_n((\overline{g_n})(m)) = 0$ no information is given by the initial segment $\overline{(g)}_n(m)$, and if $(f)_n((\overline{g_n})(m)) = k+1$ then k is unique and we put $h(n) = k$. It is shown by Beeson that the natural numbers can be embedded in V and the constants interpreted in such a way as to form a model $\mathfrak{A}$ of APP. (We can't do the same for $APP + D_V$ because definition by cases on V is not continuous.) Now form a model $\mathfrak{B}$ of (classical) $T_0 + CA_2$ from APP by § 2.

9.2 Consistency of continuity properties. It can be shown that the model $\mathfrak{B}$ satisfies the following statements of interest:

(i) Any operation $f : \mathbb{N} \rightarrow \mathbb{N}$ is continuous (in the product topology).

(ii) Any function from a complete separable metric space X to a separable metric space Y is continuous.

16) Another explanation of the difficulty is due to Klop 1977, who has shown that there is no Church-Rosser theorem for the calculus with the $\geq$ relation augmented as follows to correspond to the D_V axiom: $d\tau\tau\tau_3\tau_4 \geq \tau_3$ and $\tau_1, \tau_2 \in NF \wedge \tau_1 \neq \tau_2 \rightarrow d\tau_1\tau_2\tau_3\tau_4 \geq \tau_4$.

It follows that these continuity properties are consistent with $T_o + CA_2$, even allowing classical logic. Moreover, by modifying the model so as to take V to be the class of all partial recursive functions we can also satisfy Church's thesis CT_2 for partial functions. Hence, to the extent that constructive mathematics is contained in $T_o + CA_2 + CT_2$, we cannót prove constructively the existence of discontinuous functions on the spaces of interest to us in ordinary analysis. The main results of Beeson 1977 are in certain respects stronger positive results for a variety of intuitionistic theories T, to the effect that if a term can be proved in T to define a function (between suitable spaces) then it can proved to be continuous. This will be explained more precisely in Part IV (cf. also Beeson's corresponding results for intuitionistic theories of sets in this volume).

10. Topological models. (The material of this section and its application in 11.4 was developed in collaboration with my student Jan Stone.) Let S be a topological space and $\mathfrak{F}$ a family of partial continuous functions from S to S with card $(\mathfrak{F}) \leq$ card (S). ω is assumed disjoint from S and is considered with its discrete topology. We use $+, \Sigma$ for the operations of disjoint sum of topological spaces. Let $J = \{0\}^*$ be the closure of $\{0\}$ under pairing. Then define S_a for $a \in J$ by

$$S_o = \omega + S \quad \text{and} \quad S_{(a,b)} = S_a \times S_b \,.$$

Finally, take $V = \sum_{a \in J} S_a$. Thus pairing and projection make sense on V and we have $\omega \subseteq V$. A model $\mathfrak{M} = (V, \simeq, k, s, p, p_1, p_2, d, 0, s_N, p_N)$ of APP is generated as indicated in §5, only now defining $d\,x\,y\,u\,v$ just for $x, y \in \omega$. The choice of codes can be arranged in such a way that

for each f, the partial function $\lambda x(fx)$ is continuous on V.

We illustrate the argument for k, s where we take $k=1$, $kx = (1,x)$, $kxy = x$, $s=2$, $sx = (2,x)$, $sxy = (2,x,y)$ and $sxyz \simeq xz(yz)$. It is proved by induction on the generation of the relation App that if $f^* \to f$ and $x^* \to x$ then $f^*x^* \to fx$. For example, if f is $(2,x,y) = (2,(x,y))$ and $f^* \to f$ and $z^* \to z$ then in a suitable neighborhood of f, we have $f^* = (2, x^*, y^*)$ where $x^* \to x$, $y^* \to y$. Hence by induction it follows that $x^*z^*(y^*z^*) \to xz(yz)$. Again the model does not work to give the axiom V because full definition by cases is not continuous.

11. Applications to independence of Cantor-Bernstein statements.

11.1 Cardinality relations. These relations between classes are defined in the language of T_o as follows:

(i) $(X \sim Y) \leftrightarrow \exists f,g[f{:}X \to Y \wedge g{:}Y \to X \wedge \forall x \in X(g(fx) = x) \wedge \forall y \in Y(f(gy) = y)]$

(ii) $(X \preccurlyeq_1 Y) \leftrightarrow \exists f[f{:}X \to Y \wedge \forall x_1, x_2 \in X\,(fx_1 = fx_2 \to x_1 = x_2)]$

(iii) $(X \prec_2 Y) \leftrightarrow \exists g\,[g{:}\,Y \to X \wedge \forall x \in X\ \exists y \in Y(gy = x)]$

(iv) $(X \prec_3 Y) \leftrightarrow \exists z(Y \sim X + Z)$.

(The operation $X_0 + X_1$ is $\sum_{i \in \{0,1\}} X_i$.) The statement of <u>Cantor-Bernstein</u> can be given in one of three forms corresponding to (ii)-(iv):

$$(CB)_i \qquad X \preccurlyeq_i Y \wedge Y \preccurlyeq_i X \rightarrow X \sim Y .$$

The converse in each case is trivial. It will be shown that each of these statements is constructively unprovable, by suitable independence arguments. The first such results were obtained by van Dalen 1968 in the informal framework of Brouwer's theory of free choice sequences where maps between suitable topological spaces are necessarily continuous. We give different arguments here for the framework of T_0 .

11.2 <u>Independence of</u> CB_1 <u>from</u> $T_0 + D_V + CA_2$. This is by failure of the recursion-theoretic analogue of CB_1 . Let $\mathfrak{B} = (\mathfrak{A}, C\ell, \varepsilon)$ be a model of $T_0 + D_V + CA_2$ built from the Structure $\mathfrak{A}$ of ordinary recursion theory in §3. Let $X = \mathbb{N}$ and $Y \subseteq \mathbb{N}$ any member of $C\ell$ with $\mathbb{N} \preccurlyeq_1 Y$ but Y not recursively enumerable (e.g. such Y can be chosen co-r.e.). There is no map in the model from $\mathbb{N}$ onto Y, otherwise Y would be r.e.; thus $\mathbb{N} \not\sim Y$.

11.3 <u>Independence of</u> CB_2 <u>from</u> $T_0 + CA_2$. Here we use an example from van Dalen 1968 but apply §10 instead to get the independence result. Let X be $2^{\mathbb{N}}$ considered as a topological space, and $Y = X + E$ where E consists of a single point, thus isolated in Y. There are continuous maps $F : X \xrightarrow[\text{onto}]{} Y$ and $G : Y \xrightarrow[\text{onto}]{} X$. Let $S = X + Y$ and $\mathfrak{F} = \{F, G\}$. By §11 we form a model $\mathfrak{B} = (\mathfrak{A}, C\ell, \varepsilon)$ of $T_0 + CA_2$ with F, G represented in $\mathfrak{A}$, and every $\lambda x(fx)$ in $\mathfrak{A}$ being partial continuous on $V = \Sigma_a S_a$. Thus if $X \sim Y$ in this model we would have X homeomorphic to $X + E$, which is false.

<u>Question</u>: Is CB_2 independent from $T_0 + D_V + CA_2$? That would of course follow if the recursion-theoretic analogue of CB_2 is false.

11.4 <u>Independence of</u> CB_3 <u>from</u> $T_0 + CA_2$. Here we use an example due to Hanf (cf. Halmos 1963) of a pair of topological spaces X, Z with $X \sim X + Z + Z$ but $X \not\sim X + Z$, where $\sim$ is the relation of being homeomorphic. It follows for $Y = X + Z$ that $X \sim Y + Z$ so CB_3 fails for this topological interpretation. Now we form a model $\mathfrak{B}$ of $T_0 + CA_2$ over $S = X + Z + Z$ by §10, in which the maps giving the homeomorphism $X \sim X + Z + Z$ are included and every map is continuous on V. By Hanf's result, $X \not\sim Y$ in the sense of cardinal equivalence in this model. (Another example of van Dalen 1968 can also be adapted to this purpose.)

<u>Remark</u>. The recursion-theoretic analogue of CB_3 <u>is</u> true by Dekker-Myhill 1960. In their argument the fact that the universe V is $\mathbb{N}$ is used in an essential way. This is an example of a positive recursive analogue of a classical set-theoretic result which is not subsumed under a theorem of T_0 or even $T_0 + CA_2$.

<u>Question</u>. Is CB_3 independent from $T_0 + D_V + CA_2$?

12. A model of the weak power-class axiom. In the preceding we mostly chose different models of APP to get various consistency and independence results; the only exception was in §2. Here we modify the construction of $C\ell$ and ε so as to get a model of POW. This will also satisfy $EM_0 + D_V + IG$ but not the join axiom J. (Recall inconsistency of POW with J from II.14.1).

Let $\mathfrak{A}$ be the recursion-theoretic model of $APP + D_V$. We introduce a new code $\mathbb{C}$ for the "class of all classes". Now instead of defining $C\ell_\alpha$, ε_α simultaneously we first define $C\ell$ and then ε. (This procedure would not be possible if closure under join were required.) Take $C\ell_0 = \{\mathbb{N}, \mathbb{C}\}$ and

$$C\ell_{n+1} = C\ell_n \cup \{c_k(\underline{y},\underline{a}) \mid k = \ulcorner\phi(x,\underline{y},\underline{Z})\urcorner \text{ with } \phi \text{ elementary and } a_1,\ldots,a_m \in C\ell_n\}$$

$$\cup\ \{i(a,r) \mid a, r \in C\ell_n\}.$$

Put $C\ell = \bigcup_{n<\omega} C\ell_n$. For $a \in C\ell_n$ we define $x \varepsilon_n a$ as follows:

$$x \varepsilon_0 \mathbb{N} \leftrightarrow x \in \omega\ , \quad x \varepsilon_0 \mathbb{C} \leftrightarrow x \in C\ell\ .$$

For $c \in C\ell_n$, $x \varepsilon_{n+1} c \leftrightarrow x \varepsilon_n c$. For $c = c_k(\underline{y},\underline{a})$, $k = \ulcorner\phi(x,\underline{y},\underline{Z})\urcorner$, with the $a_i \in C\ell_n$, and $c \notin C\ell_n$ we put

$$x \varepsilon_{n+1} c \leftrightarrow (\mathfrak{A}, C\ell, \varepsilon_n) \models \phi(x,\underline{y},\underline{a})\ ,$$

and for $a, r \in C\ell_n$ and $c = i(a,r)$ with $c \notin C\ell_n$ we put

$$x \varepsilon_{n+1} c \leftrightarrow \forall I \subseteq \omega\{[\forall u(u \varepsilon_n a \wedge \forall w((w,u) \varepsilon_n r \rightarrow w \in I) \rightarrow u \in I] \rightarrow x \in I\}.$$

It may be seen that the resulting model satisfies $EM_0 + D_V + IG$ (plus CT_2 as in §3). Furthermore it satisfies

$$\forall x[x \varepsilon \mathbb{C} \leftrightarrow \exists X(x = X)].$$

Given any A we can form a weak power class $P(A)$ of A by taking $P(A) = \{fAb \mid b \varepsilon \mathbb{C}\}$ where $fAB = A \cap B$.

Remark. One can also arrange to satisfy CA_2 by using the method of §2. CA_2 can't be derived from $CA_1 + POW$ without join.

IV Realizability interpretations

1. Background. The distinctive effect of restriction of the logic to be intuitionistic is of course not shown by standard models of the kind considered in the previous part III. The following are some special properties which are typically enjoyed to some extent or other by various intuitionistic theories T: (i) The disjunction property (DP), i.e. if $T \vdash (\phi \vee \psi)$ then $T \vdash \phi$ or $T \vdash \psi$; (ii) the existential definability property (ED), i.e. if $T \vdash \exists x\, \phi(x)$ then for some term τ, we have $T \vdash \phi(\tau)$; and for T containing arithmetic: (iii) the property $(ED)_{\mathbb{N}}$ holds, i.e. if $T \vdash \exists n\, \phi(n)$ then for some (specific) n, $T \vdash \phi(\bar{n})$; (iv) T is consistent with the schematic form of Church's thesis CT_0; (v) T is closed under Church's Rule CR_0, i.e. if $T \vdash \forall n \exists m\, \phi(n,m)$ then for some (specific)

e, $T \vdash \forall n\, \phi(n, \{\bar{e}\}(n))$; and finally for T containing function variables: (vi) T is consistent with various forms of the axiom of choice AC, and is closed under corresponding choice rules. While, as remarked by Kreisel and Troelstra, these properties are neither necessary nor sufficient for T to be constructive, much of the metamathematics of constructive theories revolves around their verification.

The basic methods to obtain such results are by realizability interpretations.[17)] These were introduced by Kleene in 1945 with his notion of recursive realizability. Many extensions and variants have since been applied, due to Kreisel, Troelstra, de Jongh, J.R. Moschovakis, Friedman, Beeson and others. A rather complete survey can be found in Troelstra 1973 Ch.III or Troelstra 1977a §4; it may be helpful for the reader to look at these references in connection with this part.

It is useful to distinguish formal or internal realizability from informal or external realizability interpretations, though very often these are coupled. In the former one associates with each formula ϕ of $\mathcal{L}(T)$ a new formula ϕ_r with one additional free variable f, written $f\, r\, \phi$. In the latter one defines a relation between mathematical objects f of some sort and formulas ϕ. (Kleene's recursive realizability was of this type: he defined a relation between numbers $f \in \omega$ and formulas of arithmetic.) External realizability interpretations can often be regarded as the reading of a formal $f\, r\, \phi$ in a specific model M; that is the approach we shall take here. In any case the idea of $f\, r\, \phi$ is that f packages the constructive information (witnesses, proofs) which verifies ϕ; the definitions are thus closely related to the informal interpretation of the logical connectives in I.4.2.

By a realizability interpretation of $\mathcal{L}(T)$ in $\mathcal{L}(T')$ is meant an association $\phi \mapsto f\, r\, \phi$ with each formula ϕ (of the language of T) of a formula $f\, r\, \phi$ (of the language of T') having at most one additional free variable f. (Thus every sort of variable of $\mathcal{L}(T)$ must also be included among those of $\mathcal{L}(T')$.) This interpretation is said to be sound for T in T' if for each theorem ϕ of T we have a term τ such that $T' \vdash (\tau\, r\, \phi)$.

2. Formal realizability of $\mathcal{L}(T_0)$ in itself. This was introduced in Feferman 1975;[18)] variants from Feferman 1976b and Beeson 1977 will be explained below. When ϕ is written $\phi(\underline{x}, \underline{X})$ we write $f\, r\, \phi(\underline{x}, \underline{X})$ for $f\, r\, \phi$; when concentrating on a distinguished variable as in $\exists x\, \phi$ we may write $fr(\exists x\, \phi(x))$. The interpretation is defined inductively as follows:

17) Another method to obtain some of these properties is due to Kripke; cf. Smorynski's chapter on Kripke models in Troelstra 1973. These models will be applied at one point in Part V below.

18) It was pointed out by Beeson that the clause there for disjunction needed correction, as given in (iii) below.

(i) $[f\, r\, \phi] = \phi$ for ϕ atomic

(ii) $[f\, r(\phi \wedge \psi)] = [(p_1 f) r\, \phi \wedge (p_2 f) r \psi]$

(iii) $[f\, r(\phi \vee \psi)] = [p_1 f \in \mathbb{N} \wedge (p_1 f = 0 \rightarrow (p_2 f) r\, \phi) \wedge (p_1 f \neq 0 \rightarrow (p_2 f) r\, \psi)]$

(iv) $[f\, r(\phi \rightarrow \psi)] = \forall z[z\, r\, \phi \rightarrow (fz) r\, \psi]$

(v) $[f\, r(\forall x\, \phi)] = \forall x[(fx) r\, \phi]$

(vi) $[f\, r(\exists x\, \phi(x))] = [(p_1 f) r\, \phi(p_2 f)]$

(vii) $[f\, r(\forall X\, \phi)] = \forall X[(fX) r\, \phi]$

(viii) $[f\, r\, \exists X\, \phi(X)] = [C\ell(p_2 f) \wedge (p_1 f) r\, \phi\, (p_2 f)]$.[19)]

When it is necessary to distinguish this from other realizability interpretations to be defined later, we shall subscript this r as r_1. Note that $fr(\neg\, \phi)$ is equivalent to $\forall z \neg (zr\, \phi)$.

3. <u>Essentially</u> $(\vee, \exists)$-<u>free formulas</u>. This class of formulas are such as can be realized in a canonical way (if at all) and for that reason play a distinguished role. We call ϕ <u>essentially</u> $(\vee, \exists)$-<u>free</u> if it is built up from formulas of the form $(\tau\!\downarrow)$, $C\ell(\tau)$, $(\tau \in X)$ and $(\tau_1 \simeq \tau_2)$ by $\wedge$, $\rightarrow$ and $\forall$ applied to either sort of variable. Note that the existential information in the first three formulas, written as $\exists x(\tau \simeq x)$, $\exists X(\tau \simeq X)$ and $\exists x\, (\tau \simeq x \wedge x \in X)$ can be represented by the application term τ itself. The following lemmas are easily established for $r = r_1$.

(1) <u>For each</u> ϕ, <u>the formula</u> $(fr\phi)$ <u>is essentially</u> $(\vee, \exists)$-<u>free</u>.

(2) <u>With each essentially</u> $(\vee, \exists)$-<u>free</u> ϕ <u>is associated a term</u> τ_ϕ <u>with free variables contained in those of</u> ϕ <u>such that</u> $APP_{\mathbb{N}} \vdash [\phi \rightarrow (\tau_\phi\, r\, \phi)]$.

(3) <u>If</u> ϕ <u>is essentially</u> $(\vee, \exists)$-<u>free then</u> $APP_{\mathbb{N}} \vdash [(fr\phi) \rightarrow \phi]$.

Here (2) and (3) are proved by a simultaneous induction in order to take care of the case of implication, where we put $\tau_{(\phi \rightarrow \psi)} = k(\tau_\psi)$. Because of (2) we call τ_ϕ the <u>canonical realizer</u> of ϕ for ess. $(\vee, \exists)$-free ϕ.

<u>Remark</u>. Formulas of the kind that we call essentially $(\vee, \exists)$-free are often called <u>almost negative</u> in the intuitionistic literature.

4. <u>The scheme 'To assert is to realize'</u>. This scheme consists of all formulas of the following form:

$$\text{(A-r)} \qquad \phi \leftrightarrow (\exists f)(fr\, \phi)$$

which expresses equivalence of the assertion of ϕ with its realizability. By (2), (3) of §3 each instance of (A-r) in which ϕ is ess. $(\vee, \exists)$-free is derivable in $APP_{\mathbb{N}}$. The scheme as a whole is itself realizable:

(1) <u>for any formula</u> ϕ <u>we can find a</u> τ <u>such that</u>

$$APP_{\mathbb{N}} \vdash \tau r [\, \phi \leftrightarrow \exists f\, (fr\, \phi)\,].$$

19) The clause for $(\phi \rightarrow \psi)$ does not completely mirror the requirements for a constructive proof as expressed in II.4.2, which calls for <u>constructive recognition of</u> $\forall z[zr\phi \rightarrow (fz) r\, \phi]$ when $zr\phi$ is read 'z is a proof of ϕ'; similar remarks apply to the universal generalization cases.

For the proof one defines τ_1, τ_2 which realize each implication. Thus τ_1 is to be chosen so that $\forall z[zr\phi \to (\tau_1 z)r(\exists f(fr\phi))]$. This makes use of the fact from § 3 that $(zr\phi)$ is ess. $(\vee,\exists)$-free and so has a canonical realizer. The converse construction τ_2 is equally easy, and uses the fact (3) from § 3.

Remark. (A-r) is suggested by the basic tenet of constructive reasoning I.4.2, that a statement is to be asserted only if it is proved.

Note. It may be necessary to distinguish (A-r) for different realizability interpretations. For example we write (A-r_1) for that of §2.

5. Axioms of choice. The most general scheme considered here for the axiom of choice takes the form:

(AC) $$\forall x \in X\, \exists y\, \phi(x,y) \to \exists f\, \forall x \in X\, \phi(x, fx)$$

for variable X. Special cases of this can be formulated for each term $\mathbb{A}$ which denotes a class, e.g. $\mathbb{N}$, $\mathbb{N} \to \mathbb{N}$, etc. We write $(AC)_{\mathbb{A}}$ for the restriction of the scheme to $X = \mathbb{A}$. There is a close connection between the schemes (AC) and (A-r); we have:

(1) (A-r_1) implies (AC).

For suppose $\forall x[x \in X \to \exists y\, \phi(x,y)]$. By (A-r) for $r = r_1$ we can find g such that $\forall x,z[zr(x \in X) \to (gxz)r\, \exists y\phi(x,y)]$. But by clause (i) of the definition of r_1, $(x \in X) \to 0r(x \in X)$, so $p_1(gx0)r\,\phi(x, p_2(gx0))$. But then $\phi(x, p_2(gx0))$ holds, so $f = \lambda x(p_2(gx0))$ is a choice function. (Note that the proof just uses the APP axioms.) It may be of interest to the reader to see which instances of (A-r_1) are implied by (AC); they form a wide class.

While full (AC) will thus be realized in the r_1-interpretation, this will not hold for other r_i to be considered. A special consequence of (AC) which will be realized even when (AC) is not, is the axiom scheme of dependent choices:

(DC) $$\forall x \in X\, \exists y \in X\, \phi(x,y) \to \forall x_0 \in X\, \exists f \in X^{\mathbb{N}}[f0 = x_0 \wedge \forall n\, \phi(fn, fn')].$$

Using the axioms APP + $\mathbb{N}$ we can derive (DC) from (AC) in essentially the standard way. Namely, given g such that $\forall x \in X[gx \in X \wedge \phi(x, gx)]$ we define f by primitive recursion to satisfy $f0 \simeq x_0$, $fn' \simeq g(fn)$. Then it is proved by full induction on $\mathbb{N}$ that $\forall n[fn\downarrow \wedge\, \phi(fn, fn')]$.

6. The theory $T_0^{(-)}$. We do not have a soundness theorem for r_1-realizability of T_0 in itself. The problem arises with the elementary comprehension scheme CA_1(i.e. ECA). To realize $\exists X\{c_n(\underline{y},\underline{Z}) = X \wedge \forall x[x \in X \leftrightarrow \phi(x,\underline{y},\underline{Z})]\}$ for $n = \ulcorner\phi(x,\underline{y},\underline{Z})\urcorner$ we have to show how to convert any u with $ur\,\phi(x,\underline{y},\underline{Z})$ into a w such that $wr(x \in X)$ where $x = c_n(y,\underline{Z})$ and conversely. But for $r = r_1$ we have $wr(x \in X) \leftrightarrow x \in X$. Thus we would have to obtain $\phi \leftrightarrow \exists u(ur\,\phi)$, which is only generally true for essentially $(\vee,\exists)$-free formulas. However, this difficulty

suggests an obvious modification of CA_1 to a scheme $CA_1^{(-)}$, where CA_1 is taken only for essentially $(\vee,\exists)$-free ϕ. By $EM_0^{(-)}$ we mean the axiom system $APP + CA_1^{(-)} + \mathbb{N}$, and by $EM_0^{(-)}\upharpoonright$, the same theory with induction on $\mathbb{N}$ restricted. We claim that $EM_0^{(-)}\upharpoonright$ serves to obtain the same mathematical consequences in BCM as $EM_0\upharpoonright$ (II.10), and similarly for $EM_0^{(-)}\upharpoonright + J$, $EM_0^{(-)} + J$ in place of $EM_0\upharpoonright + J$, $EM_0 + J$, resp. (II.11). The reason is very simple: in the formalization of BCM by following Bishop's official definitions, we never make essential use of $\vee$ or $\exists$ in defining properties of sets - since the witnessing information is always required to accompany the presentation of the elements of those sets (recall I.15). Hence $CA_1^{(-)}$ always suffices in place of CA_1. The only difference appears when we enter the theory of ordinals and Borel sets (II.12). Here one must make a slight modification in the IG axiom to achieve the same results. For example, previously we took $\mathcal{O}_1 = i(A,R)$ where $A = \{x \mid x=0 \vee x=(p_2x)^+ \vee x = \sup_{\mathbb{N}}(p_2^2 x)\}$, and $R = \{(y,x) \mid (x=(p_2x)^+ \wedge y = p_2x) \vee x = \sup_{\mathbb{N}}(p_2^2 x) \wedge \exists n(y = p_2^2 xn)\}$. Thus R is defined using $\exists$ in an essential way. We now modify IG to $IG^{(-)}$ by taking $i(A,S) = I$ to satisfy instead

$$\forall x \in A\{\forall y,z[(z,y,x) \in S \rightarrow y \in I] \rightarrow x \in I\}$$

as the closure axiom and then taking a corresponding induction principle. This has the same effect as the previous IG with $(y,x) \in R \leftrightarrow \exists z[(z,y,x) \in S]$. Let $T_0^{(-)} = EM_0^{(-)} + J + IG^{(-)}$. It is thus seen that $T_0^{(-)}$ serves to obtain the same mathematical consequences in BCM as T_0 (II.10 - II.12). (The theory $T_0^{(-)}$ was introduced in Feferman 1975, where the soundness result of the next section was outlined.)

7. Soundness theorem for r_1-realizability of $T_0^{(-)}$ in itself. It is usually a routine matter to verify soundness of the axioms and rules of intuitionistic logic for any reasonable realizability interpretation (cf. Troelstra 1973 Ch.III). For the present r_1 interpretation, soundness of the logical part of $T_0^{(-)}$ is easily verified using the APP axioms to provide the requisite constructions. Going on to the non-logical axioms, it is straightforward in each case to verify soundness of each axiom or scheme on the basis of the corresponding principles themselves. The reason in the case of $CA_1^{(-)}$ has already been explained in §6, by use of the properties of essentially $(\vee,\exists)$-free formulas from §3. In the case of the induction scheme on $\mathbb{N}$ we are required to give a τ which realizes $\phi(0) \wedge \forall x(\phi(x) \rightarrow \phi(x')) \rightarrow \forall x(x \in \mathbb{N} \rightarrow \phi(x))$. This must show how to convert any z_0, z_1 with $z_0 r \phi(0)$ and $\forall x,w[w\, r\, \phi(x) \rightarrow (z_1 xw) r \phi(x')]$ into some $u = \tau z_0 z_1$ with $\forall x[x \in \mathbb{N} \rightarrow (ux) r \phi(x)]$; u is defined by recursion (uniformly from z_0, z_1) to satisfy $u0 = z_0$ and $ux' \simeq z_1 x(ux)$ and the required conclusion is proved by induction on $\mathbb{N}$. One proceeds similarly for the realization of $IG^{(-)}$. In this way we conclude that the following holds for $r = r_1$:

(1) with each theorem ϕ of $T_0^{(-)}$ can be associated a term τ such that $T_0^{(-)} \vdash (\tau r \phi)$.

Corresponding results can be obtained for the various subtheories $EM_0^{(-)}\upharpoonright$, $EM_0^{(-)}$, $EM_0^{(-)}+J$ which have been considered. In addition, if we take $CA_2^{(-)}$ to be the restriction of CA_2 to essentially $(\vee,\exists)$-free stratified formulas, we get:

(2) r_1 - realizability is also sound for $T_0^{(-)}+CA_2^{(-)}$ in itself.

Now as we have seen in §4, the scheme (A-r) is itself r - realized in a trivial way; moreover (A-r) implies (AC) by §5. Finally, it is impossible to realize $\perp$ in a consistent theory, since $(fr_1\perp)\leftrightarrow\perp$. We may thus conclude that:

(3) r_1-realizability is sound for $T_0^{(-)}+(A\text{-}r)$ in $T_0^{(-)}$ and the same holds for $T_0^{(-)}+CA_2^{(-)}+(A\text{-}r)$ in $T_0^{(-)}+CA_2^{(-)}$; hence $T_0^{(-)}+CA_2^{(-)}+(A\text{-}r)$ and (therefore also) $T_0^{(-)}+CA_2^{(-)}+AC$ are consistent.

Coupled with the following we see now the metamathematical power of realizability interpretations applied to theories dealt with in this paper emerging.

8. Consistency of Church's Thesis (CT_0). Let $T=T_0^{(-)}+CA_2^{(-)}+(A\text{-}r)$ for $r=r_1$. Then T is consistent with CT_0, by the following argument. We read r_1-realizability in a model of T whose applicative part is the ordinary recursion-theoretic model of III.3. It is to be shown for any ϕ that '$\forall n\exists m\,\phi(n,m)\rightarrow\exists e\forall n\,\phi(n,\{e\}(n))$' is realized in this model by suitable f. This is to convert any g realizing $\forall n\exists m\,\phi(n,m)$ into (fg) realizing $\exists e\forall n\,\phi(n,\{e\}(n))$. Now from the hypothesis we have $\forall n[(p_1(gn))r\,\phi(n,p_2(gn))]$. We take $\{e\}(n)=p_2(gn)$ for all n, from which description f is obtained very simply.

9. Closure properties of $T_0^{(-)}$ and related theories.

9.1 r_2 - realizability. In order to obtain the disjunction and existence properties for $T_0^{(-)}$ we modify r_1 in a manner due to Kleene, (called q-realizability, cf. Troelstra 1973 p. 189); this is here denoted by r_2. Only the following clauses are varied:

(iii)' $[fr(\phi\vee\psi)]=[p_1f\in\mathbb{N}\wedge(p_1f=0\rightarrow\phi\wedge(p_2f)r\,\phi)]\wedge[p_1f\neq 0\rightarrow\psi\wedge(p_2f)r\,\psi]$

(iv)' $[fr(\phi\rightarrow\psi)]=\forall z[\phi\wedge(zr\,\phi)\rightarrow(fz)r\psi]$

(vi)' $[fr\,\exists x\,\phi(x)]=[\phi(p_2f)\wedge(p_1f)r\,\phi(p_2f)]$

(viii)' $[fr\,\exists X\,\phi(X)']=C\ell(p_2f)\wedge\phi(p_2f)\wedge(p_1f)r\,\phi(p_2f)$.

It is easily checked that the soundness theorem for r_2-realizability holds for each of the sub-theories of $T_0^{(-)}+CA_2^{(-)}$ considered. In addition, (A-r) is r_2-realized just as before.

9.2 The ED property. Suppose T is any theory for which we have soundness of r_2 - realizability e.g. any of the theories just indicated. Then if $T \vdash \exists x\, \phi(x)$ there is a term τ such that $T \vdash \tau r(\exists x\, \phi(x))$. Hence by clause (vi)' for r_2 we have $T \vdash \phi(p_2 \tau)$. Thus T enjoys the existential definability property.

9.3 The DP and $ED_{\mathbb{N}}$ properties. For these we need a special argument due to Beeson 1977.[20] Let T be a subtheory of $T_0^{(-)} + CA_2^{(-)}$ for which we have soundness of r_2 - realizability of T in T. Assume also that $APP_{\mathbb{N}} \subseteq T$. It was stated in III.8.2 that if $T \vdash (\tau \in \mathbb{N})$ then for some n, $T \vdash (\tau = \bar{n})$. (The proof made use of the model NF of normal terms.) Now if $T \vdash \exists n\, \phi(n)$ it follows that for some τ, $T \vdash (p_2 \tau) \in \mathbb{N} \wedge \phi(p_2 \tau)$. Hence T has the $ED_{\mathbb{N}}$ property; it is a corollary that T has the DP property.)

Remark. It is a little more work to obtain closure under Church's rule. One way is to formalize the properties of r_2 - realizability of any finite subtheory of T within T.

10. Inconsistency of T_0 with AC. We next turn to the question of obtaining corresponding properties for T_0. This section shows that r_1-(or r_2-) realizability is not sound for T_0, since T_0 is inconsistent with AC. This will lead us to consider a new realizability interpretation (which does not verify full AC). The proof of contradiction of $T_0 + AC$ is by an argument due to Friedman (originally given for $T_0 + D_V + AC$). Let $X = \{x \mid \exists n(xx \in \mathbb{N} \to xx \neq n)\}$. Recall here that we are using the conventions $\exists n\, \phi(n) \leftrightarrow \exists x(x \in \mathbb{N} \wedge \phi(x))$ and $(xx \in \mathbb{N}) \leftrightarrow \exists y(xx \simeq y \wedge y \in \mathbb{N})$, i.e. $\exists n(xx \simeq n)$. Trivially by definition $\forall x \in X \exists n[xx \in \mathbb{N} \to xx \neq n]$. Hence if AC is assumed we can find an f such that $\forall x \in X[fx \in \mathbb{N} \wedge (xx \in \mathbb{N} \to xx \neq fx)]$. In particular, $f \in X \to ff \in N \wedge ff \neq ff$ so $\neg(f \in X)$. Note that $\forall x[xx \in \mathbb{N} \to x \in X]$ holds since we can always find $n \neq xx$ using definition-by-cases on $\mathbb{N}$. Hence $\neg(x \in X) \to \neg(xx \in \mathbb{N})$. It follows that $\neg(ff \in \mathbb{N})$. But then by logic $ff \in \mathbb{N} \to ff \neq 0$ so $f \in X$ which is a contradiction to our original assumption, namely that AC_X holds. It is of course essential for this argument that X is existentially defined, which is not possible in $T_0^{(-)}$.

11. Realizability for T_0 via a refinement T_0^*.

11.1 The theory T_0^*. The language $\mathcal{L}^*$ of T_0^* is obtained by refining the language $\mathcal{L}$ of T_0 as follows: instead of the two-placed relation $(x \in A)$ of $\mathcal{L}$ we now have a 3-placed relation $(x \in_z A)$. In this language we take

$$x \in A \leftrightarrow_{def} \exists z\,(x \in_z A),$$

which gives a translation of $\mathcal{L}$ into $\mathcal{L}^*$. If ϕ is any formula of $\mathcal{L}$ we denote its translation by ϕ^*. If T is any theory in the language $\mathcal{L}$ of T_0, by T^* we mean the theory whose axioms are exactly the ϕ^* for all axioms ϕ of T.

20) It should be noted for comparison with Beeson 1977 that $T_0^{(-)}$ is there denoted EM and $EM_0^{(-)} + J$ is denoted EMN.

11.2 <u>Translation of $\mathcal{L}^*$ into $\mathcal{L}$</u>. This is accomplished in the following simple way. With each formula ψ of $\mathcal{L}^*$ is associated a formula $\tilde{\psi}$ of $\mathcal{L}$, which is obtained by replacing each atomic formula $(x \,\varepsilon_z\, A)$ by $[(x,z) \in A]$ and which, except for some changes of constants, is otherwise unaffected. Each of the combinatory constants $k, s, d, p, p_1, p_2, 0, s_{\mathbb{N}}, p_{\mathbb{N}}$ is unchanged, but the class formation constants c_k, j, i are replaced by new constants $\tilde{c}_k, \tilde{j}, \tilde{i}$ as will be explained in the next section.

11.3 $\tilde{r}_1$ - <u>realizability</u>. This is an interpretation of $\mathcal{L}^*$ (and thence of $\mathcal{L}$) which will make all of CA realizable. The idea to realize $\exists X[x \in X \leftrightarrow \phi(x)]$ as expressed in $\mathcal{L}^*$ and then translated back into $\mathcal{L}$ i.e. as $\exists X[\exists z((x,z) \in X) \leftrightarrow \tilde{\phi}(x)]$ is to produce an X such that from any z with $(x,z) \in X$ we can find a w with $w r \tilde{\phi}(x)$ and conversely. The simplest way to achieve this is to take $z = w$ and thus to take $X = \{(x,w) \mid w r \tilde{\phi}(x)\}$. It is here where the change of constants enters; if $c_n = \ulcorner\phi(x, \underline{y}, \underline{Z})\urcorner$ we'll have $\tilde{c}_n = \ulcorner(p_2 x) r \tilde{\phi}(p_1 x, \underline{y}, \underline{Z})\urcorner$.

For this purpose $\tilde{r}_1$ - realizability is defined as follows; the translations of c_n, j, i are given by a simultaneous inductive definition. First we write down the clauses defining $f r \phi$ for ϕ in $\mathcal{L}^*$ exactly like those for r_1 in § 2. The only difference appears in the fact that we now have atomic formulas $(x \,\varepsilon_z A)$ in place of the old $(x \in A)$, so we read $[fr(x \,\varepsilon_z A)] = (x \,\varepsilon_z\, A)$. Then we take $\tilde{r}_1$ to be $\tilde{r}$ where

$$(f \,\tilde{r}\, \phi) = (\widetilde{fr\phi}),$$

i.e. we translate the realizability interpretation of $\mathcal{L}^*$ just described. Now $\tilde{c}_n, \tilde{j}, \tilde{i}$ are chosen to satisfy the following:

(i) if $n = \ulcorner\phi(x, \underline{y}, \underline{Z})\urcorner$, $\tilde{c}_n(\underline{y}, \underline{Z}) = \{(x,z) \mid z \,\tilde{r}\, \phi(x, \underline{y}, \underline{Z})\}$

(ii) if $\forall x \in A[C\ell(fx)]$ then $\tilde{j}(A,f) = \{((x,y),(z,w)) \mid (x,z) \in A \wedge (y,w) \in fx\}$, and

(iii) $\tilde{i}(A,R) = i(A_1, R_1)$ where $A_1 = \{(x,(z,f)) \mid (x,z) \in A \wedge \forall y,w[((y,x),w) \in R \rightarrow f(y,w)\downarrow]\}$

and $R_1 = \{((y,fyw),(x,(z,f))) \mid ((y,x),w) \in R\}$.

The choice of the constants is made in such a way that for each of the axioms ϕ from CA_1 (or CA_2), J and IG as expressed in $\mathcal{L}^*$ we can find a τ such that $\tau \,\tilde{r}\, \phi$ is provable from the corresponding axiom (or axiom schema) as expressed in $\mathcal{L}$. It follows that for any theory T over APP which is based on some combination of these axioms and schemata plus the axioms $\mathbb{N}$ or $\mathbb{N}\!\upharpoonright$, <u>we have a soundness theorem for $\tilde{r}_1$ - realizability of T^* in T</u>. Once more the scheme (A-r), as written out in $\mathcal{L}^*$, is trivially $\tilde{r}_1$ - realized. It is a corollary that

(1) $$T_0^* + (CA_2)^* + (A\text{-}r) \ \underline{\text{is consistent}}.$$

Hence any consequence of (A-r) in $\mathcal{L}$ is consistent with $T_0 + (CA_2)$ regarded as translated into $\mathcal{L}^*$. It is simpler to study such consequences in the latter language than to pass through $\tilde{r}$. Note: To get further consistency with Church's thesis CT_0 we simply couple this with the recursion-theoretic model as described in §8.

12. Consequences of "to assert is to realize" in $\mathcal{L}^*$. We assume at least $EM_0^* + (A\text{-}r)$ where r is r_1 for $\mathcal{L}^*$ throughout this section.

12.1 Dependent choices. DC is derived from these assumptions in the following way. Suppose $\forall x \in A\ \exists y \in A\ \psi(x,y)$ holds, i.e. $\forall x,z[x \in_z A \rightarrow \exists y,w(y \in_w A \wedge \psi(x,y))]$. Then there exists g which realizes this statement, so for each x,z with $x \in_z A$, $g(x,z)$ provides us with a triple (y,w,u) such that $y \in_w A \wedge u r \psi(x,y)$. Given $x_0 \in A$ fix some z_0 with $x_0 \in_{z_0} A$. Then using g we define a sequence (x_n, z_n, u_n) by recursion such that $g(x_n,z_n) \simeq (x_{n+1}, z_{n+1}, u_n)$ and $x_n \in_{z_n} A$ and $u_n r \psi(x_n,x_{n+1})$ for each n. Let $f = \lambda n.x_n$. Then passing from the right to the left side of (A-r) we have $f0 = x_0 \wedge \forall n[fn \in A \wedge \psi(fn,fn')]$.

As a corollary of this and 11.3 we have that

(1) $T_0 + CA_2 + DC$ is consistent.

Remark. The argument here brings out the reason why DC can be dealt with constructively even where AC can't in the presence of full comprehension. From $\forall x \in A\ \exists y\ \phi(x,y)$ written in $\mathcal{L}^*$ as $\forall x,z[x \in_z A \rightarrow \exists y\ \phi(x,y)]$ we can merely conclude $\exists f \forall x,z[x \in_z A \rightarrow \phi(x,f(x,z))]$ from (A-r). This is the result which was referred to in I.4.7.

12.2 Canonically realizable classes (choice bases). We write $C(A)$ for the following formula:

$$\exists g[\exists z(x \in_z A) \rightarrow x \in_{gx} A].$$

A is called canonically (or self-) realizable if $C(A)$ holds. $C(A)$ is equivalent to AC_A, i.e. the scheme of choice with base A. For suppose $C(A)$ holds using g and that $\forall x \in A\ \exists y\ \phi(x,y)$. Then as just remarked we find f such that $\forall x,z[x \in_z A \rightarrow \phi(x,f(x,z))]$. It follows that $\forall x \in A\ \phi(x, f(x,gx))$. Conversely if AC_A holds then from $\forall x \in A\ \exists z(x \in_z A)$ (which is trivial by definition in $\mathcal{L}^*$) we conclude $\exists g \forall x \in A(x \in_{gx} A)$, i.e. $C(A)$.

Now $C(\mathbb{N})$ holds because $AC_{\mathbb{N}}$ is a consequence of DC. Also $C(V)$ holds because $\forall x \exists y\ \phi(x,y) \rightarrow \exists f \forall x \phi(x,fx)$ by (A-r). Furthermore, the property C is closed under class constructions which can be defined by essentially $(\vee,\exists)$-free formulas, for which we have found canonical realizers by §3. In particular, if $C(A)$ and $C(B)$ then $C(A \times B)$ and $C(A \rightarrow B)$ hold and if $\phi(x)$ is ess. $(\vee,\exists)$-free then also $C(\{x \in A | \phi(x)\})$. This gives consistency of $T_0 + CA_2$ with an extensive collection of instances of AC. (The consistency of T_0 with $(AC)_{FT}$, i.e. AC_{N_σ} for all $\sigma \in FTS$ was noted by Beeson 1977.)

12.3 The presentation axiom (Aczel). Call (A^*, h) a presentation of A if $h : A^* \to_{onto} A$; this is called a full presentation of A if $\mathbb{C}(A^*)$ holds. Intuitively, in a presentation each element x of A is represented (in possibly more than one way) by $x^* \in A^*$ such that $h(x^*) = x$; x^* contains "additional information" that "verifies" $x \in A$. When we have a full presentation, no further information need be added. The presentation axiom PA is the statement that for every A there exists a full presentation (A^*, h) of A. This was introduced (in a slightly different form) by Aczel in unpublished notes; he observed that it serves to derive the various mathematical consequences of (A-r). In the present framework, PA is a trivial consequence of (A-r); we simply take $A^* = \{(x,z) \mid x \varepsilon_z A\}$ and $h(x,z) = x$.

12.4 Having your cake and eating it too with (A-r) as an implement. In the informal discussion of I.15.3 the attempt to have one's constructive cake and eat it too was taken to be a matter of being casual about showing the witnessing information required by the official definitions. Here we can provide a theoretical framework to justify such practices simply by assuming (A-r) for $r = r_1$ in $\mathcal{L}^*$. In effect, the informal definition of a class A in the form $A = \{x \mid \phi(x)\}$ gives rise to $A^* = \{(x,z) \mid z\, r\, \phi(x)\}$, which corresponds to the official definition. By $\phi(x) \leftrightarrow \exists z (z\, r\, \phi(x))$ we have $x \in A \leftrightarrow \exists z[(x,z) \in A^*]$. A realizable refinement of CA in $\mathcal{L}^*$ allows us to take $\forall x[x \varepsilon_z A \leftrightarrow z\, r\, \phi(x)]$, so that this A^* is exactly the same as the full presentation of A described in 12.3. For example, if we define

$\mathbb{R}^+ = \{x \mid x \in \mathbb{R} \wedge \exists n > 0 (x_n > \frac{1}{n})\}$ we than have $(\mathbb{R}^+)^* = \{(x,n) \mid x \in \mathbb{R} \wedge n > 0 \wedge (x_n > \frac{1}{n})\}$,

just as required by the official definition. Using AC in its weakened form $\forall x \in A \exists y\, \phi(x,y) \to \exists f\, \forall x,z[(x,z) \in A^* \to \phi(x, f(x,z))]$ we can conclude that an inverse function is defined on $(\mathbb{R}^+)^*$ knowing that $\forall x \in \mathbb{R}\ \exists y \in \mathbb{R}\, (x \cdot y = 1)$.

Remark. $T_0^* + (A\text{-}r)$ provides an alternative way of reading Bishop which is in some respects simpler than by T_0, since one can formalize the informal mathematical arguments more directly. (Note: The same ends can be achieved by the presentation axiom instead of (A-r).) It is not meant by this that $T_0^* + (A\text{-}r)$ is in direct accordance with Bishop's views; (that is open to discussion).

13. Closure properties of T_0. In order to obtain the ED and $ED_{\mathbb{N}}$ (and hence DP) properties for T_0, Beeson 1977 introduced a kind of combination of $\tilde{r}$- and q-realizability. His definition (loc.cit.pp.281-282) is complicated by the requirement to have a doubling (X, X^*) of class variables, where the new variable X^* is to correspond to the class of all (x,z) such that $z\, r(x \in X)$. In addition, Beeson also doubled individual variables. This does not seem to be necessary, and the following simpler definition is proposed for the same purpose. With each formula $\phi(x_1, \ldots, x_n, Y_1, \ldots, Y_m)$ of $\mathcal{L} = \mathcal{L}(T_0)$ is associated a formula

fr $\phi(x_1,\ldots,x_n,Y_1,\ldots,Y_m,Y_1^*,\ldots,Y_m^*)$ of $\mathcal{L}$ as follows:

(i) $[\mathrm{fr}(x \in Y)] = [(x \in Y) \wedge (x,f) \in Y^*]$ and $[\mathrm{fr}\,(x = Y)] = [x = (Y, Y^*)]$
$[\mathrm{fr}\,\phi] = \phi$ for the other atomic formulas

(ii) $[\mathrm{fr}(\phi \wedge \psi)] = [(p_1 f)\mathrm{r}\,\phi \wedge (p_2 f)\mathrm{r}\psi]$

(iii) $[\mathrm{fr}(\phi \vee \psi)] = [(p_1 f \in \mathbb{N}) \wedge (p_1 f = 0 \to \phi \wedge (p_2 f)\mathrm{r}\,\phi) \wedge (p_1 f \neq 0 \to \psi \wedge (p_2 f)\mathrm{r}\,\psi)]$

(iv) $[\mathrm{fr}(\phi \to \psi)] = \forall z[\phi \wedge (z\mathrm{r}\,\phi) \to (fz)\mathrm{r}\,\psi]$

(v) $[\mathrm{fr}\,\exists x\,\phi(x)] = [\phi(p_2 f) \wedge (p_1 f)\mathrm{r}\,\phi(p_2 f)]$

(vi) $[\mathrm{fr}\,\exists X\,\phi(X)] = [\exists X, X^*((p_2 f) = (X,X^*) \wedge \phi(X) \wedge (p_1 f)\mathrm{r}\,\phi(X,X^*)]$

(vii) $[\mathrm{fr}\,\forall x\,\phi(x)] = \forall x[(fx)\mathrm{r}\,\phi(x)]$

(viii) $[\mathrm{fr}\,\forall X\,\phi(X)] = \forall X,X^*[f(X,X^*)\mathrm{r}\,\phi(X, X^*)]$.

Remark (added in proof): Beeson has pointed out real difficulties with the proposed realizability of T_0 which are met when looking for suitable reinterpretations of the constants. It is thus not known whether his definition can be simplified in any essential way to serve the same purposes.

14. Applications to continuity properties. Beeson 1977 has used the consequences of realizability such as $AC_{\mathbb{N}}$ and $ED_{\mathbb{N}}$ for the theories T considered in the preceding sections [21] e.g. for $T = T_0^{(-)}$ and $T = T_0$, to prove local continuity rules of the following form, where $\mathbb{A}$, $\mathbb{B}$ are any closed terms for classes:

LCR($\mathbb{A}$, $\mathbb{B}$). If T proves $\mathbb{A}$ is a complete separable metric space and $\mathbb{B}$ is a separable metric space and ϕ is an extensional property and $\forall x \in \mathbb{A}\,\exists y \in \mathbb{B}\,\phi(x,y)$ then T proves that $\forall x \in \mathbb{A}\,\exists y \in \mathbb{B}\,[\phi(x,y) \wedge$ "y is stable for x" $]$.

Here "y is stable for x" stands for $\forall \epsilon > 0\ \exists \delta > 0\ \forall u \in N_\delta^{\mathbb{A}}(x)\ \exists v \in N_\epsilon^{\mathbb{B}}(y)[\phi(u,v)]$, where $N_\epsilon^{\mathbb{A}}(x) = \{u \in \mathbb{A} \mid d_{\mathbb{A}}(x,u) < \epsilon\}$ and $d_{\mathbb{A}}$ is the metric of $\mathbb{A}$. Equality in a metric space is defined by $x_1 =_{\mathbb{A}} x_2 \leftrightarrow d_{\mathbb{A}}(x_1,x_2) = 0$; extensional properties are understood to be those $\phi(x,y)$ for which

$$\phi(x_1,y_1) \wedge x_1 =_{\mathbb{A}} x_2 \wedge y_1 =_{\mathbb{B}} y_2 \to \phi(x_2, y_2).$$

As a corollary of LCR($\mathbb{A}$, $\mathbb{B}$) one has: if T proves that F is a function from $\mathbb{A}$ to $\mathbb{B}$ under the same hypotheses on $\mathbb{A}$, $\mathbb{B}$ then T proves that F is continuous. (Note that the hypothesis means $F: \mathbb{A} \to \mathbb{B}$ and $x_1 =_{\mathbb{A}} x_2 \to F(x_1) =_{\mathbb{B}} F(x_2)$.) Beeson's method of proof of LCR($\mathbb{A}$, $\mathbb{B}$) uses the representation of complete metric spaces with countable dense subset D in the form of the Cauchy sequences from D.

21) The matter is actually more complicated: one must formalize within the T considered the corresponding results for all finite subtheories of T.

This allows one to push the problem back to verification of $LCR(\mathbb{N} \to \mathbb{N}, \mathbb{N})$. Roughly, the idea is that if $\forall x \in (\mathbb{N} \to \mathbb{N}) \exists m\, \phi(x,m)$ is proved, where ϕ is extensional, then for each specific $g \in (\mathbb{N} \to \mathbb{N})$ we can prove $\exists m\, \phi(g,m)$ from $T \cup Diag(g)$. This requires only a finite part $\langle g(0),\ldots,g(n-1)\rangle$ of g. Further by $ED_{\mathbb{N}}$ we can find an m s.t. $\phi(g,\bar{m})$ is proved from the same. By formalizing this argument one gets the desired result. Beeson also has results on local uniform continuity rules for compact spaces and a number of consistency and independence results concerning continuity statements. He has further extended these to other formalisms such as those of Myhill and Friedman, as presented in his contribution to this volume.

Discussion. In a sense, Beeson's results confirm Brouwer's ideas that we should be able to prove that every real function on $\mathbb{R}$ (resp. [0,1]) is continuous (uniformly continuous). But the present results have the advantage that the systems to which they apply also have a set-theoretic interpretation. So one can be sure that if an existence proof $\forall x \in \mathbb{A}\, \exists y \in \mathbb{B}\, \phi(x,y)$ can be formalized in BCM then it yields stability or continuity of solutions which are true in the classical sense. Often such results can be obtained directly by ad hoc arguments. But the continuity results described considered as a part of global (or systematic) constructivity may first point the way to what can be obtained for special problems. In other words, the global results can serve as the stimulus and point of departure for mathematically interesting local results. (Indeed this has been the case with Beeson's studies of stability phenomena in the Plateau problem.)

Question. The property $ED_{\mathbb{N}}$ depends essentially on not having D_V. But one doesn't see why the continuity results should be disturbed by its presence. Do Beeson's results on LCR extend to $T_0 + D_V$ by some other arguments?

V. Relations with subsystems of analysis.

1. Introduction and summary of results. In this part (except for the special §2) we describe results which establish the equivalence of certain subsystems of T_0 with subsystems of classical 2nd order analysis. It is assumed here that the reader is familiar with the designations of various of the latter such as $(\Pi^0_\infty\text{-CA})$, $(\Delta^1_1\text{-CA})$, $(\Sigma^1_1\text{-AC})$, $(\Delta^1_2\text{-CA})$, $(\Sigma^1_2\text{-AC})$, as well as with the principle (BI) of bar induction.[22] When '$\restriction$' is used following designation of a theory we mean that the principle of full induction on $\mathbb{N}$ is replaced by the axiom of induction. We write $T_1 \leq T_2$ to mean that T_1 is proof-theoretically reducible to T_2 (i.e. if $Con(T_2)$ implies $Con(T_1)$ by a finitary argument) and $T_1 \equiv T_2$

22) For descriptions of these and some information about their interrelationships cf. Feferman 1977.

if $T_1 \leq T_2$ and $T_2 \leq T_1$. In connection with the following results one also has much information about which sentences are conserved in one direction or the other; however, for simplicity we do not mention such for the most part. PA denotes classical Peano's arithmetic, HA = Heyting's arithmetic.

(1) $EM_0{\restriction} \equiv HA$, in fact $EM_0{\restriction}$ is a conservative extension of HA.

(2) $EM_0{\restriction} \pm J \equiv (\Sigma^1_1 - AC){\restriction} \equiv PA$

(3) $EM_0 + J \equiv (\Sigma^1_1 - AC)$

(4) $EM_0{\restriction} \pm J + IG{\restriction} \equiv (\Sigma^1_2 - AC){\restriction} \equiv (\Pi^1_1 - CA){\restriction}$

(5) $EM_0 + J + IG{\restriction} \equiv (\Sigma^1_2 - AC)$

(6) $T_0 = EM_0 + J + IG \leq (\Sigma^1_2 - AC) + (BI)$.

In all of these except the conservation result of (1), we can also include classical logic and the axiom D_V on the l.h.s. The exact relationship in (6) is unsettled.

Conjecture. $T_0 \equiv (\Sigma^1_2 - AC) + (BI)$.

Credits. The conservation result in (1) is due to Beeson 1979, by a Kripke-model argument outlined in the next section. The $\equiv$ in (1) comes simply from (2) and the fact that $PA \equiv HA$. Conservation of $(\Sigma^1_1 - AC){\restriction}$ over PA has been established by Barwise-Schlipf 1975 using recursively saturated models; it is also stated by Friedman 1975 where conservation of $(\Sigma^1_2 - AC){\restriction}$ over $(\Pi^1_1 - CA){\restriction}$ (for a certain class of sentences) is announced as well. (The method of recursively saturated models has also been extended to prove the latter in unpublished notes by myself.) The proof-theoretical equivalences $(\Sigma^1_1 - AC){\restriction} \equiv PA$ and $(\Sigma^1_2 - AC){\restriction} \equiv (\Pi^1_1 - CA){\restriction}$ have been established by Sieg. The result (3) is due to Aczel (unpublished); a new method of proof was found by myself (Feferman 1976 c). This method was also used there to establish $\leq$ in (2) and (4)-(6). The relations $\geq$ in (4) and (5) are due to Sieg 1977. Only outlines of the various ideas involved are given in the following. Detailed presentations of the proofs of these and related results will be found in the chapter by Feferman and Sieg in the projected volume "Iterated inductive definitions and subsystems of analysis: recent proof-theoretical studies" (for the Lecture Notes in Mathematics Series) which is to consist of contributions by Buchholz, Feferman, Pohlers and Sieg.

2. $EM_0{\restriction}$ is conservative over HA. We first describe the proof of an easier result from Feferman 1976 a : $EM_0{\restriction}$ in classical logic is conservative over PA. To begin with, the axioms APP (even with D_V) are formally modelled in PA by taking $App(x,y,z) \leftrightarrow \{x\}(y) \simeq z$. Any model $\mathfrak{M}$ of PA thus determines an applicative structure $\mathfrak{A}$. This is used to build a model $(\mathfrak{A}, C\ell, \epsilon)$ of $EM_0{\restriction}$

by the method of III.1, but leaving off the clauses for i and j ; now the process closes off at ω. Namely, $C\ell = \bigcup_{n<\omega} C\ell_n$, $\epsilon = \bigcup_{n<\omega} \epsilon_n$ where $C\ell_o = \{\mathbb{N}\}$ and $x \epsilon_o \mathbb{N} \leftrightarrow x = x$, $C\ell_{n+1} = C\ell_n \cup \{c_k(\underline{y}, \underline{a}) \mid k = \ulcorner\phi(x, \underline{u}, \underline{Z})\urcorner$ and ϕ is elementary and $a_1, \ldots, a_m \in C\ell_n\}$, with $x \epsilon_{n+1} c_k(\underline{y}, \underline{a}) \leftrightarrow (\mathfrak{A}, C\ell_n, \epsilon_n) \models \phi(x, \underline{y}, \underline{a})$. Note that $\mathfrak{M}$ may be non-standard and $\mathbb{N}$ is coextensive with the domain of $\mathfrak{M}$. It may be seen that for each $A \in C\ell$ there exists a formula $\psi(x, \underline{u})$ of arithmetic such that for some choice of parameters $\underline{y}$ in $\mathfrak{M}$, $\forall x[x \in A \leftrightarrow \mathfrak{M} \models \psi(x, \underline{y})]$. Hence the induction axiom $(\mathbb{N}\restriction)$ is verified, and we do indeed have $(\mathfrak{A}, C\ell, \epsilon)$ a model of $EM_o\restriction$. To conclude, conservation holds by the completeness theorem for the classical predicate calculus: if θ is a sentence of arithmetic such that $(EM_o\restriction) \vdash \theta$ but $PA \nvdash \theta$ we can choose $\mathfrak{M} \models \neg\theta$ and get a contradiction.

Now Beeson 1979 has shown $EM_o\restriction$ conservative over HA by an adaptation of this argument to Kripke models, using the completeness theorem for intuitionistic logic in terms of the latter. Given any Kripke model $\mathfrak{M} = \langle(\mathfrak{M}_p)_{p\in P}, \leq\rangle$ of HA one modifies the construction of $(\mathfrak{A}, C\ell_n, \epsilon_n)$ as just described to a construction of $\langle(\mathfrak{A}_p, C\ell_{n,p}, \epsilon_{n,p})_p, \leq\rangle$ for each n and thence of a Kripke model $\langle(\mathfrak{A}_p, C\ell_p, \epsilon_p)_{p\in P}, \leq\rangle$ of $EM_o\restriction$.

Discussion. The significance of this result is given by I.15.5, according to which $EM_o\restriction$ is adequate to essentially all of BCM except for the theory of ordinals and Borel sets. A corresponding result had previously been obtained by Friedman 1977 (conservation of $\underset{\sim}{B}$ over HA for Π^0_2 sentences, strengthened to full conservation by Beeson 1979). Thus this portion of BCM does not really take advantage of the strong constructive principles implicitly accepted by Bishop; on the other hand it is of foundational interest that it is justified by the most elementary of these.

3. $EM_o\restriction + J \leq (\Sigma^1_1\text{-}AC)\restriction$, $EM_o + J \leq \Sigma^1_1\text{-}AC$. The proofs of these results from Feferman 1976c are given by formal models which verify classical logic and D_V. We start again with the recursion-theoretic interpretation $App(x,y,z) \leftrightarrow \{x\}(y) \simeq z$. Now, instead of defining $C\ell$ in transfinite stages, one defines it simply to be the set of Δ^1_1 indices. That is, let $P^1_1(e,x)$ (e=0,1,2,...) be a standard Π^1_1-enumeration of all Π^1_1 sets (predicates of one argument x); then $S^1_1(e,x) \leftrightarrow \neg P^1_1(e,x)$ induces a Σ^1_1-enumeration of all Σ^1_1 sets. We put e in $C\ell$ if the pair of indices $(e)_o$, $(e)_1$ determines a Δ^1_1 set, i.e.

$$C\ell(e) \leftrightarrow_{def} \forall x[P^1_1((e)_o, x) \leftrightarrow S^1_1((e)_1, x)].$$

Put $x \in a \leftrightarrow_{def} P^1_1((a)_o, x)$ for $C\ell(a)$. To prove closure under CA_1 in this model reduces to showing that if $\phi(x,\underline{y},\underline{Z})$ is elementary and we substitute Δ^1_1 definable sets D_i for the Z_i, the result is also Δ^1_1 (with index e uniformly recursive in given indices d_i for the D_i); this is by the Δ^1_1-substitution theorem of Addison-Kleene-Schoenfield. Formalization of the latter makes use of $(\Sigma^1_1\text{-AC})$. So far, the argument serves to give a model of EM in $(\Sigma^1_1\text{-AC})$; next it is seen that only restricted $(\Sigma^1_1\text{-AC})\upharpoonright$ suffices if one starts with $EM_0\upharpoonright$. To complete the proof, J is verified as follows. Suppose $C\ell(a)$ and $\forall x \in a\, C\ell(\{f\}(x))$, i.e. that

$$\forall x\{P^1_1((a)_0, x) \to \forall y[P^1_1((\{f\}(x))_0,y) \leftrightarrow S^1_1((\{f\}(x))_1,y)]\} .$$

Then we easily obtain a Δ^1_1 index for $j(A,f)$ where a is the index of A.

<u>Remark</u>. By $(\Sigma^1_1\text{-AC})\upharpoonright \leq HA$, this shows that J is really of no use without unrestricted induction. That was already noted informally in II.11.3, where transfinite types were shown to exist in $EM_0 + J$ - but not in $EM_0\upharpoonright + J$.

4. <u>$EM_0\upharpoonright + J + IG\upharpoonright \leq (\Sigma^1_2\text{-AC})\upharpoonright$, $EM_0 + J + IG\upharpoonright \leq (\Sigma^1_2\text{-AC})$ and $T_0 \leq (\Sigma^1_2\text{-AC}) + (BI)$.</u>

The proofs (again from Feferman 1976c) all use the same idea, which simply follows that of §3 one level up. Let $P^1_2(x,e)$, $S^1_2(x,e)$ enumerate the Π^1_2, resp. Σ^1_2 sets. Take $C\ell(a) \leftrightarrow \forall x[P^1_2(x,(e)_0) \leftrightarrow S^1_2(x,(e)_1)]$ and $x \in a \leftrightarrow P^1_2(x,(a)_0)$. Now one applies the A-K-S substitution theorem for Δ^1_2 predicates, which is proved using Σ^1_2-AC. This serves to show $EM_0 + J$ modelled in $(\Sigma^1_2\text{-AC})$ and $EM_0\upharpoonright + J$ in $(\Sigma^1_2\text{-AC})\upharpoonright$. To verify $IG\upharpoonright$ in this model we simply apply A-K-S again: if A,R are Δ^1_2 then the set i(A,R) which is Π^1_1 in A,R is also Δ^1_2 (with index e uniformly recursive in the indices a,r of A, R resp.). The induction <u>axiom</u> of $IG\upharpoonright$ follows immediately by definition of i(A,R) as the least set satisfying the given closure conditions. To obtain the full principle of induction for IG one must apply full (BI), which gives the final result: $T_0 = EM_0 + J + IG \leq (\Sigma^1_2\text{-AC}) + (BI)$.

5. <u>$(\Sigma^1_1\text{-AC})\upharpoonright \leq PA$, $(\Sigma^1_2\text{-AC})\upharpoonright \leq (\Pi^1_1\text{-CA})\upharpoonright$; consequent reductions into T_0.</u>

As was remarked in the survey of credits in §1, one has proof-theoretical arguments due to Sieg for these first two reductions, corresponding to earlier conservation results of Barwise-Schlipf and Friedman. Now $PA \leq HA$ by the negative $(\neg\neg)$ translation as is well known, and $HA \subseteq EM_0\upharpoonright$ so this completes the relations in §1(2). Next $(\Pi^1_1\text{-CA})\upharpoonright$ can be interpreted in the corresponding intuitionistic system $(\Pi^1_1\text{-CA})\upharpoonright^{(i)}$ by the negative translation, and the latter is directly contained in $EM_0\upharpoonright + IG\upharpoonright$. This completes the chain in §1(4).

6. $(\Sigma_1^1\text{-AC}) \le EM_0 + J$. To begin with, $(\Sigma_1^1\text{-AC}) \le (\Pi_1^0\text{-CA})_{<\varepsilon_0}$ by Friedman 1970. (That used a model-theoretic argument; a proof-theoretical one is outlined in Feferman 1977.) As is familiar, $(\Pi_1^0\text{-CA})_{<\varepsilon_0} \equiv RA_{<\varepsilon_0}$ (ramified analysis in levels $<\varepsilon_0$), and $RA_{<\varepsilon_0} \le RA^{(i)}_{<\varepsilon_0}$ by the negative translation. Finally, $RA^{(i)}_{<\varepsilon_0}$ is contained in $EM_0 + J$, using Join to transfinitely iterate the ramified hierarchy up to each ordinal $\alpha < \varepsilon_0$ (full induction up to α follows from full induction on $\mathbb{N}$). This completes the $\equiv$ in §1(3).

7. $(\Sigma_2^1\text{-AC}) \le EM_0 + J + IG\upharpoonright$. By Friedman 1970, $(\Sigma_2^1\text{-AC}) \le (\Pi_1^1\text{-CA})_{<\varepsilon_0}$ and by Feferman 1970, $(\Pi_1^1\text{-CA})_{<\varepsilon_0} \le ID_{<\varepsilon_0}$ where the latter is a classical theory of iterated first-order inductive definitions up to any $\alpha < \varepsilon_0$. The main next step is to show $ID_{<\varepsilon_0} \le ID^{(i)}_{<\varepsilon_0}(\mathfrak{O})$, i.e. to the intuitionistic theory of the classes $\mathfrak{O}_\alpha$ for $\alpha < \varepsilon_0$. This has been established by Sieg 1977. Finally, $ID^{(i)}_{<\varepsilon_0}(\mathfrak{O})$ is contained directly in $EM_0 + J + IG\upharpoonright$. In this way the $\equiv$ in §1(5) is completed.

Remark. Results closely related to those of Sieg 1977 have been obtained independently by Pohlers and Buchholz, by more complicated methods, but which also give more detailed information. Presentation and comparison of all this work will be found in the forthcoming joint volume referred to in § 1.

8. Questions and conjectures.

(i) The conjecture $(\Sigma_2^1\text{-AC}) + (BI) \le T_0$ has already been stated in § 1. What is missing up to now is the proof-theory analogous to that indicated in § 7.

(ii) We have shown $EM_0 + IG + POW$ consistent in III.12. What is the strength of this system and various of its subsystems? It appears that POW cannot be used very effectively with these axioms. I conjecture that $EM_0\upharpoonright$ + POW $\equiv$ HA.

(iii) The set-theoretical model of T_0 in III.6 can be modified to give a model of $S_0 + POW$ by essentially using $\{0,1\}^A$ as a representative of a power set of A for any set A. Now presence of J makes the axiom POW much more effective. What is the strength of $S_0 + POW$? Further, is $S_0 + POW + CT_i$ consistent $(i = 0,1)$?

(iv) What are the strengths of the various theories considered when CA_2 is added?

Stanford University

Bibliography

Barendregt, H. P.

1971 Some extensional term models for combinatory logics and λ-calculi, Dissertation, University of Utrecht.

1977 The type free lambda calculus, in *Handbook of Mathematical Logic*, 1091-1132.

Barwise, J. and Schlipf, J.

1975 On recursively saturated models of arithmetic, in *Model Theory and Algebra*, Lecture Notes in Maths., 498, 42-55.

Beeson, M. J.

1977 Principles of continuous choice and continuity of functions in formal systems for constructive mathematics, Ann. Math. Logic, 12, 249-322.

1979 Extensionality and choice in constructive mathematics (to appear). Continuity in intuitionistic set theories (this volume).

Bishop, E.

1967 *Foundations of Constructive Analysis* (McGraw-Hill, New York).

1970 Mathematics as a numerical language, in *Intuitionism and Proof Theory* (North-Holland, Amsterdam), 53-71.

Bishop, E. and Cheng, H.

1972 Constructive measure theory, Mem. Amer. Math. Soc., 116.

Brouwer, L. E. J.

1975 *Collected Works*, I (North-Holland, Amsterdam).

van Dalen, D.

1968 A note on spread-cardinals, Composito Math., 20, 21-28.

Dekker, J. C. E. and Myhill, J.

1960 Recursive equivalence types, Univ. Calif. Pubs. in Math., 3, 67-214.

Feferman, S.

1970 Formal theories for transfinite iterations of generalized inductive definitions and subsystems of analysis, in *Intuitionism and Proof Theory* (North-Holland, Amsterdam), 303-326.

1975 A language and axioms for explicit mathematics, in *Algebra and Logic*, Lecture Notes in Math., 450, 87-139.

1976a Notes on the formalization of Bishop's constructive mathematics (unpublished notes).

1976b Having one's cake and eating it too (in Bishop's constructive mathematics) (unpublished notes).

1976c On the length of some constructive systems of 'Explicit Mathematics' (unpublished notes).

1977 Theories of finite type related to mathematical practice, in *Handbook of Mathematical Logic* (North-Holland, Amsterdam), 913-971.

1978 Recursion theory and set theory: a marriage of convenience, in *Generalized Recursion Theory II* (North-Holland, Amsterdam), 55-98.

Friedman, H.

1970 Iterated inductive definitions and Σ^1_2 - AC, in *Intuitionism and Proof Theory* (North-Holland, Amsterdam), 435-442.

1975 Systems of second-order arithmetic with restricted reduction (unpublished notes).

1977 Set theoretic foundations for constructive analysis, Ann. of Math., 105, 1-28.

Gödel, K.

1958 Uber eine bisher noch nicht benützte Erweiterung des finiten Standpunktes, Dialectica, 12, 280-287.

Goodman, N. and Myhill, J.

1972 The formalization of Bishop's constructive mathematics, in *Toposes, Algebraic Geometry and Logic,* Lecture Notes in Maths., 274, 83-96.

Kleene, S. C.

1959 Countable functionals, in *Constructivity in Mathematics* (North-Holland, Amsterdam), 81-100.

Klop, J. W.

1977 A counterexample to the CR property of λ-calculus + DMM → M (unpublished note).

Martin-Löf, P.

1971 Hauptsatz for iterated inductive definitions, in *Proc. Second Scandinavian Logic Symposium* (North-Holland, Amsterdam), 179-216.

1975 An intuitionistic theory of types, in *Proc. Logic Colloquium* (North-Holland, Amsterdam), 73-118.

Myhill, J.

1975 Constructive set theory, J. Symbolic Logic, 40, 347-382.

Prawitz, D.

1971 Ideas and results in proof theory, in *Proc. Second*

Scandinavian Logic Symposium (North-Holland, Amsterdam), 235-307.

Scott, D.

1970 Constructive validity, in *Symposium on Automatic Demonstration*, Lecture Notes in Maths., 125, 237-275.

Sieg, W.

1977 Trees in metamathematics (theories of inductive definitions and subsystems of analysis), Dissertation, Stanford University.

Tait, W. W.

1968 Constructive reasoning, in *Logic, Methodology and Philosophy of Science III* (North-Holland, Amsterdam), 185-199.

Troelstra, A. S.

1973 *Metamathematical investigation of intuitionistic arithmetic and analysis* (Ed.), Lecture Notes in Maths., 344.

1975 Non-extensional equality, Fund. Math., 82, 307-322.

1977a Aspects of constructive mathematics, in *Handbook of Mathematical Logic* (North-Holland, Amsterdam), 973-1052.

1977b *Choice Sequences.* A Chapter of Intuitionistic Mathematics (Clarendon Press, Oxford).