Three conceptual problems that bug* me (7th Scandinavian Logic Symposium, Uppsala lecture, Aug.18-20, 1996; Draft)

Solomon Feferman

*__Bug__ (v.t.): to bother, annoy, pester; to prey on, worry

Introduction

I will talk here about three problems that have bothered me for a number of years, during which time I have experimented with a variety of solutions and encouraged others to work on them. I have raised each of them separately both in full and in passing in various contexts, but thought it would be worthwhile on this occasion to bring them to your attention side by side. In this talk I will explain the problems, together with some things that have been tried in the past and some new ideas for their solution.

Types of conceptual problems.

A conceptual problem is not one which is formulated in precise technical terms and which calls for a definite answer. For this reason, there are no clear-cut criteria for their solution, but one can bring some criteria to bear. These will vary from case to case. There are three general types of conceptual problems in mathematics of which the ones that I will discuss today are examples. These are:

1°. **Finding a suitable framework for the direct expression of seemingly problematic notions.** Examples from logic are:

   - The use of non-standard models to develop analysis with infinitesimals.
   - Formal theories for Brouwerian free choice sequences.

Examples from mathematics (outside of logic) are:

- Projective geometry for points at infinity and the line at infinity.
- Distribution theory for the Dirac δ-function.

My problem of this type: Find a suitable framework for the direct expression of structural concepts which admit self-application (e.g., the category of all categories).

N.B. It is to be expected in solving this sort of conceptual problem that there will be *trade-offs*. For example, in non-standard analysis one has to distinguish between internal and external properties of the model. And in projective geometry, one has to give up metric properties.

2°. **Explicating an informal concept.** (This is perhaps what one thinks of first under conceptual problems, that of conceptual analysis.) Examples from logic are:

- The definition of truth for formal languages.
- The definition of effective computability.

Examples from mathematics are:

- The definition of random variable.
- The notion of natural isomorphism.

My problem of this type: How should one define the notion of natural well-ordering?

3°. **Generalizing some heuristically successful analogies.** Examples from logic:

- Effective descriptive set theory, to generalize classical descriptive set theory and the logical theory of hierarchies.
- Admissible set theory, to generalize ordinary and meta-recursion theory.

Examples from mathematics:

- Hilbert spaces, to generalize finite-dimensional linear algebra and theory of integral equations.

- Ideal theory, to generalize prime factorization in integers and in specific algebraic number fields.

My problem of this type: How should one generalize the concepts of large cardinals as they appear in impredicative set theory, admissible set theory, and proof-theoretical ordinal notation systems?

These are not all the types of conceptual problems that arise in practice. Another one which is closely related to the type 1° is:

4°. **Finding the "right" framework in which to carry out certain developments.** Examples from logic are:

- Axiomatic set theory to develop Cantorian theory of ordinals and cardinals.
- Domain theory for denotational semantics of programs.

Examples from mathematics are:

- Category theory for homological algebra.
- Sheaf theory for algebraic geometry.

One has an overlap with problems of type 1°, e.g. to the extent that Cantorian set theory involved problematic notions or methods, and earlier algebraic geometry involved problematic arguments.

A type of problem that one thinks of more in terms of applied mathematics is:

5°. **Modeling mathematically some specific features of our experience.** Here examples from logic are:

- Intuitionistic logics to model constructive reasoning.
- Ramified systems to model the idea of predicativity.

Examples from mathematics are ubiquitous:

- Differential geometry to model "curved space".
- Differential equations to model dynamical physical systems.

The listener will no doubt think of other examples of all these types and perhaps other distinctive types of conceptual problems. I'll not talk further about 4° and 5° here.

I. The problem of self-applicable structural concepts

I assume some familiarity with category theory in this part. The mathematical notion of category isolates an interesting mathematical structure on the class of structure-preserving maps (or "morphisms") between all structures of a given kind. Then, pursued informally, one is naturally and directly led to the following:

(R1) For each usual kind K of mathematical structure (for which we have some usual notion of structure-preserving morphism), there is the category of all structures of that kind, e.g. the category Grp of all groups (group homomorphisms), the category Top of all topological spaces (continuous maps), the category Cat of all categories (functors).

(R2) For any two categories A and B we have the category $(A \to B)$ [or B^A] of all functors from A to B.

(R1) and (R2) may be considered to be partial *requirements* (or *criteria*) on a framework that it must meet in order to permit direct expression of self-applicable structural concepts. So far, one has only provided frameworks for versions of (R1) and (R2) which are essentially restricted in one way or another. Thinking of structures as objects $(A, \dots)$ with one or more *domains* A, which are *collections* of objects, on which are defined some *relations, operations*, etc., it is natural to think of what has to be accomplished as being part of a broader framework in which we have the following familiar closure conditions of a set-theoretical nature on objects, collections and operations:

(R3) (i) The set $\mathbb{N}$ of natural numbers is among our collections.

(ii) For any objects a, b we have the ordered pair (a, b).

(iii) For any collections A, B, we have the collections $A \cap B$, $A \cup B$, $A - B$, $A \times B$, $A \to B$, $\wp(A)$, etc.

(iv) For any collection A and sequence of collections $\langle B_x \rangle_{x \in A}$, we have the collections $\bigcup B_x[x \in A]$, $\bigcap B_x[x \in A]$, $\sum B_x[x \in A]$, $\prod B_x[x \in A]$.

This can be considered a further requirement on the desired framework for informal category theory and other general theories of structures. It is natural, to begin with, to seek a set-theoretical foundation for these requirements. (R3) is of course met in systems like ZF or ZFC if we take *all* objects, collections and operations to be sets in the usual way. But we can't then meet

(R1), since there is no set of all groups, etc. Here are two familiar solutions which meet (R1)-(R3) but only in rather restricted forms for one or another of these:

1. **Grothendieck's method of universes.** This is a modification of the preceding attempt at a set-theoretical interpretation. Add to ZFC the hypothesis: "there exist infinitely many inaccessible cardinals". *Universes* are defined to be sets like the V_α for α inaccessible, which satisfy suitably strong closure conditions. For any universe U, the category K_U of all structures of a given kind K which are members of U belongs to a larger universe U'. So this is a modification of (R1). The requirements (R2) and (R3) hold within any universe U.

2. **MacLane's small and large categories** (MacLane 1961). Work in the BG theory of sets and classes. A collection is called *small* if it is a set, i.e. a member of V, and *large* if it is a class which is not a set. Again, (R1) is met in a modified form: By Grp_V we mean the (large) category of all small groups, Top_V is the (large) category of all small topological spaces, Cat_V is the (large) category of all small categories, etc. (R2) is met only for A small, and there are similar restrictions on cartesian power and products in (R3).

Now here are some solutions that I have tried over the years:

3. **Set theory with reflection principle** (Feferman 1969). This is a modification of the Grothendieck approach to avoid using infinitely many inaccessibles, but which also gives some internal sense to the category of all categories, for example. Add a set constant S to ZF (or ZFC) with axioms that express that $(S, \in)$ is an elementary substructure of $(V, \in)$. We think of S as a set which acts as a *surrogate* for V. The category of all structures of a given kind K which belong to S, K_S, is then a surrogate for K_V, and belongs to V. Any set-theoretical properties established of the former transfer to the latter. (Some categorical arguments require S to be inaccessible; this of course can be added, but at the cost of then allowing arbitrarily many inaccessibles again.)

4. **NFU with pairing** (Feferman 1974, unpub.). Jensen (1969) proved that NFU is consistent relative to ZF. Add a pairing operation symbol to its language, and modify the definition of stratification, so that when a, b are assigned a type level n (the same for both) then the type level assigned to (a, b) is also n. (This is not met by usual definitions

of pairing in NF or NFU.) Call the system with pairing and comprehension for stratified formulas in this modified sense, NFUP; Jensen's consistency proof extends to NFUP. (One can also establish consistency in ω-logic, in order to make $\mathbb{N}$ standard, relative to ZFC plus a Ramsey-ish cardinal.) The point of working in NFUP is that the collection of all structures of a usual kind K can be defined by a stratified formula in this modified sense. For simplicity take, for example, the collection K of all structures (A, R) where $R \subseteq A^2$ is a partial ordering on A; assign the elements of A type level 0; then the elements of R, which are pairs, are also of type level 0, so A and R both get assigned level 1 and so K exists. Thus (R1) is met. In particular, in NFUP we can prove $(Cat, Funct, \ldots) \in Cat$. Also (R2) is met if we regard functions in the usual set-theoretic way as many-one relations. Most of (R3) is met, but not, e.g. Cartesian product $\prod B_x[x \in A]$. For, in order for this to be a collection of functions from A to the union of the B_x, we have to have the elements x of A and the elements of each B_x at the same level, say 0. But then each B_x is assigned level 1, and there is no way to represent the *sequence* $\langle B_x \rangle_{x \in A}$ as a function from A into the universe. Another thing that is not met is the formation of $(A, \ldots)/E$ for E a congruence relation on the structure $(A, \ldots)$. Again the problem is that this requires a function whose arguments and values are of different type level. I was not successful in developing a system for stratification of pairs with mixed type level. In any case, I'm not happy with NFU as a framework, without its own intuitive support. But stratification by itself is natural in certain contexts, as we'll see.

4. **Non-extensional type-free theory of operations and classes** ("Explicit Mathematics", Feferman 1975). The point of departure for this work was to give an axiomatic account of Errett Bishop's approach to constructive mathematics. But the framework is much broader and has a variety of applications in constructive, predicative and classical mathematics. The first main difference from set theory is that it separates the role of operations and classes, giving them independent status. The second is that it is non-extensional. The informal model for the system T_0 of explicit mathematics is that all individuals are given by some sort of syntactic presentation, and in particular operations (which are in general partial) are given by defining rules, and the classes (or classifications of individuals) are given by defining properties. However, within T_0, these notions are dealt with abstractly.

Thus the universe V of individuals in T_0 is closed under pairing, and comprises among its individuals partial operations $f, g, \ldots$ and classes $A, B, \ldots$. That is, since operations and classes are thought of as given by syntactic presentations, they may also be treated as individuals. In particular, classes are Janus-faced; when considering what members they have, we are thinking of them as given by defining properties. But when treating them as individuals, we are just thinking of them in the manner that they are specified. When forming new classes $\{x : \phi(x, y, \ldots, A, \ldots)\}$ from given class parameters $A, \ldots$, we must treat each parameter A in its class or property guise; thus ϕ must be stratified in the sense that class variables and constants in ϕ appear only to the right of $\in$ subformulas. But otherwise, self-membership is both reasonable and possible, e.g. $V \in V$. Operations may have operations or classes as arguments and/or values. In particular, $\langle B_x \rangle_{x \in A}$ is treated as an operation f which is defined for each $x \in A$ and has value B_x, and we can form the Cartesian sum $\sum B_x[x \in A]$, and product $\prod B_x[x \in A]$ of these classes over A. (This overcomes the obstacle to such in NFUP.) In T_0 we meet the requirements (R2) and (R3) except for $\wp(A)$—for, there is no class $C = \wp(V)$. The reason is instructive. Otherwise the identity operation f takes each individual x in C to itself as a class, and we could then form the class $B = \sum x[x \in C] = \{(x, y) : x \in C \wedge y \in x\}$, and finally the Russell class $R = \{x : x \in C \wedge (x, x) \notin B\}$. But this blocks meeting (R1), since similarly there is no class B of all structures (A, a) with $a \in A$, and no class of all groups $(G, \ldots)$, etc.

Note: Some set-theoretically oriented folk think that sacrificing extensionality at the outset makes systems of explicit mathematics a nonstarter. But the evidence of the massive amount of work in Bishop-style mathematics demonstrates that extensionality is an easily dispensable principle. One simply deals with structures $(A, \ldots)$ where each domain carries an "equality" relation E on it, which is a suitable congruence relation, instead of dealing with A/E. Constructivity is not an issue here; the method works equally well in classical mathematics. For example, the real numbers may be identified with the collection of all Cauchy sequences $\langle s_n \rangle$, where two such sequences are declared to be "equal" if they have the same limit. I have made the case for non-extensionality at length in (Feferman 1975, 1979). Of course this is one of the trade-offs of trying to work in the explicit mathematics framework, but I do not consider it a major trade-off.

5. **A theory of partial operations and partial classes** (Feferman 1977). In trying to see whether we can also meet (R1) in an explicit mathematics style framework, I tried to see what happens if we give up stratified comprehension in T_0. But then in order to avoid inconsistency, one has to restrict comprehension in other ways. One way to do this is by a notion of *partial classes* A, for which we have two (disjoint) relations of *membership* $\in$ and *non-membership* $\tilde{\in}$, but where we need not have $\forall x[x \in A \vee x \tilde{\in} A]$; a class that satisfies the latter is called a *total* class. One can formulate a theory $\widetilde{T}$ which is similar to T_0 in its principles for operations, but much more liberal in its existence principles for classes, provided that we allow them to be partial. (This grafts ideas of Fitch and Gilmore onto the application structure.) But now, for (R1), the best one can do is form the partial category $\mathrm{Grp}_{\mathrm{Tot}}$ of all total groups, and similarly the partial category of all total categories, etc. So this is still not a happy solution.

6. **A new idea? Explicit mathematics with stratified structural notions.** Let's go back to the trouble with $\wp(V)$ in T_0. As far as a theory of structure goes, the essence of the problem lies in the fact that we can't form a class B of all structures of the form (A, a) where $a \in A$. For then with B as a class parameter in stratified comprehension, we could form the class $B_1 = \{x : (x, x) \in B\}$ and then the class $R = \{x : x \notin B_1\}$. R, being a class, satisfies $(R, R) \in B$ if $R \in R$, so $R \in B_1$, and then $R \notin R$. So $R \notin R$. But then $(R, R) \notin B$, and hence $R \notin B_1$, so $R \in R$. So we have an inconsistency. What point in this argument is to be avoided? It seems to me innocuous to have closure under the operation that takes B to B_1 and also under complementation, that takes B_1 to R. The problem is the *slip* from B to B_1 in which *we forgot how the elements of B are structured.* This is also why in systems like NF, it is not obvious that we might not avoid inconsistency; even though each instance of comprehension is stratified, as we make use of sets in an argument, we forget the stratification conditions that introduced them. What we need is something that gives a more flexible combination of the benefits of typed and untyped systems. Typed systems are too rigid for mathematical practice and they don't permit self-application. Untyped systems can run into trouble by forgetting the typing that is naturally associated with certain constructions. So the idea would be that *part of the information* of what makes a structure (A, a) with $a \in A$ is the typing *relation*, say a of type 0 and A of type 1, and this information should be carried along *with the struc-*

ture, so that we don't forget it. This should be considered as a *relative typing*, not an absolute one as in type theory. Then we could hope to form the class B of all such structures, but we would be barred from forming B_1 as above, because stratification conditions would prevent us from passing to pairs (x, x) in B. Now for any particular structure (A, a) in B we could say that $(B, (A, a)) \in B$, since we can shift the relative typing to have (A, a) treated as an object of type 0 and B as an object of type 1. If this can be carried through in some coherent way, we could then meet all the requirements (R1)-(R3). But so far I have not seen how best to do this.

II. The problem of natural well-orderings

We have lots of examples of natural well-orderings. These are given by orderings of expressions in some notation system S for ordinals: with each $s \in S$ is associated the ordinal $|s|$ that it denotes, and then we define

$$s \leq_S t \leftrightarrow |s| \leq |t|. \tag{1}$$

Everyone is familiar with the notation system for ordinals less than the Cantor ordinal ϵ_0, based on a system of expressions generated by closure under addition and exponentiation to the base ω, or more compactly by combinations of the form

$$\omega^s + t, \tag{2}$$

with $|\omega^s+t| = \omega^{|s|}+|t|$. We shall freely use such suggestive notation, keeping letters like $\alpha, \beta, \ldots$ for ordinals and $s, t, \ldots$ for terms, but with operations on ordinals represented by the same symbols for terms. Thus with the binary operation $f(\alpha, \beta) = \omega^\alpha + \beta$, is associated the formal function $f(s, t) = \omega^s + t$, and the system S of notations for ordinals less than ϵ_0 is generated by f starting with the symbol 0 (for 0). As defined in (1) above, $\leq_S$ is a pre-well-ordering; we could if wished choose a canonical representative of each equivalence class and obtain a well-ordering from it. In general below, we shall deal with pre-well-orderings and not make this passage.

It is also familiar that this ordering relation is primitive recursive (numbering terms as usual) and that by transfinite induction along this ordering (applied to prim. rec. predicates) we can prove the consistency of PA. Moreover, this is best possible, in the sense that for any proper initial segment of this ordering, we can prove transfinite induction applied to arbitrary arithmetical formulas up to that segment in PA. However, we can also define very short "unnatural" prim. rec. well-orderings such that the consistency of PA can be proved from the formal assumption that transfinite induction holds for them. For example, take an ordering which looks like 0,1,2,... as long as we don't reach an inconsistency in PA, but in which we introduce an infinite descending sequence once we do. Since PA is in fact consistent, this ordering is of type ω, but we can't prove that in PA, since that would give a proof of the consistency of PA within itself. Further, from the hypothesis that transfinite induction holds for this ordering we can prove the consistency of PA. So the question is, what distinguishes the above natural well-ordering of type ϵ_0 from such "monsters". Evidently, recursiveness is not enough. Moreover, the fact that they can be used to prove the consistency of some formal system is not enough. Though natural well-orderings arise naturally in the proof theory of formal systems, the feeling is that what distinguishes such orderings are certain intrinsic mathematical properties that are independent of their possible use in proof-theoretical work. On the other hand it is just the use of natural well-orderings in proof theory that makes them important. When it is said that we have determined the ordinal of a formal system, just what is meant by that? To be sure, we can define the ordinal of a system in a way independent of particular questions of representation, as the sup of the ordinals of all provably recursive well-orderings (assuming the notion of well-ordering can be expressed in the system in one way or another). But that does not tell us in what form the ordinal of the system is to be determined. Now does it insure that we can use a determination of that ordinal to prove the consistency and other fundamental metamathematical properties of the system in question.

What we shall do in the following is describe a few systems of natural ordinal representation that have been of significance in proof-theoretical work, and try to say in general mathematical terms what is special about them. (All of this will be familiar to proof-theorists who have worked on ordinal analysis in the sense descended from the Gentzen-Schütte line.) Then we shall consider more speculative properties. Note from the outset that we have shifted the problem from—What constitutes a natural well-ordering qua pure well-ordering?—to the problem—What constitutes a natural structure $(\alpha, <, \mathbf{f})$ on an ordinal α?— where $\mathbf{f}$ is a sequence of functions under which α is closed, and α is the closure under $\mathbf{f}$ of $\{0\}$. The structure $(\alpha, <, \mathbf{f})$ determines a system S of ordinal representation and ordering $\leq_S$ as described at the beginning of this section, whose order-type $(\mathrm{mod} =_S)$ is just α. But now, since α is determined by $\mathbf{f}$, we may shift the problem once more to look at the members of the sequence $\mathbf{f}$ as functions on *arbitrary* ordinals and ask—What leads us to such $\mathbf{f}$ in the first place and what makes their properties have useful consequence for the associated system S just described?

For more details concerning various of these systems, the reader should begin with Schütte (1977) and Pohlers (1989) and then progress to Pohlers (1996) and Rathjen (1996).

1° **The Cantor system.** This is as described above, consisting of the single function $f = \lambda\xi, \eta.\omega^{\xi} + \eta$; we use χ (for 'Cantor') to denote this function. The closure under χ of $\{0\}$ is ϵ_0. In general, the ordinals closed under χ are the ϵ-numbers, which are the solutions α of $\omega^{\alpha} = \alpha$.

2° **The Veblen hierarchy.** Let f_0 be any unary normal function of ordinals. Define a sequence of functions f_α for $\alpha \neq 0$ by: f_α enumerates $\{\xi : f_\beta(\xi) = \xi \text{ for all } \beta < \alpha\}$. We call this the *Veblen hierarchy* over f_0. We can combine this into a single function $f = \lambda\xi, \eta.f_\xi(\eta)$. It is easily seen that if we take $f_0 = \lambda\xi.(1 + \xi)$ then the resulting f is just the function χ above. If we take $f_0 = exp_\omega = \lambda\xi.\omega^{\xi}$ then the resulting f is denoted ϕ. The closure under χ and ϕ of $\{0\}$ (or, equivalently under $+$ and ϕ) is the least α with $\phi_\alpha(0) = \alpha$ and is denoted Γ_0; in general such fixed points are called Γ-numbers (or strongly critical numbers). The step from a Veblen hierarchy $\lambda\xi, \eta.f_\xi(\eta)$ to the normal function $\lambda\xi.f_\xi(0)$ is called *diagonalization.*

In the following, let Ω_α be the initial ordinal of cardinal $\aleph_\alpha$. Ω_0 is then just ω and Ω_1 is the set Ω of all countable ordinals. In general, the sets Ω_α for $\alpha \neq 0$ are called the *(ordinal) number classes.* A system $\mathbf{f}$ on arbitrary ordinals is said to *preserve the number classes* if for each $\alpha \neq 0$, Ω_α is closed under $\mathbf{f}$. The Cantor function χ preserves the number classes. If a normal function f_0 preserves the number classes, then so also does the Veblen hierarchy over it. In particular, the system (χ, ϕ) preserves the number classes. The following systems, eventually designed to produce a system which preserves countable ordinals, are produced in a different way.

3° **Bachmann-Pfeiffer-Isles hierarchies.** The details are complicated to describe, but the idea originating with Bachmann (1950) is to use hierarchies in higher number classes to produce systems of notation which are then used to index hierarchies of functions preserving lower number classes. For example one uses the function $\exp_\Omega$ to generate an analogue of the Cantor hierarchy lifted to the Ω_2 number class. The first ordinal under which this is closed is $\epsilon_{\Omega+1}$, which I write as $\epsilon(\Omega+1)$

in the following. Now this is used to generate an extension of the Veblen hierarchy on $\Omega, \langle f_\alpha \rangle_{\alpha < \epsilon(\Omega+1)}$ where for $\alpha = \sup\{\alpha_\xi : \xi < \Omega\}$ of cofinality Ω in the lifted Cantor system, f_α is defined by diagonalization over all preceding functions: $f_\alpha(\xi) = g_\xi(0)$, where g_ξ is the f_β with index α_ξ. For α not of cofinality Ω, we use a fundamental sequence for it in the system for $\epsilon(\Omega + 1)$ to proceed as in the Veblen hierarchy under Ω. Pfeiffer extended this idea to the higher number classes and Isles extended it out to the first inaccessible. (See Isles 1970.) I shall refer to these all as Bachmann-style hierarchies. They give rise to very strong systems of ordinal representation for countable ordinals, which with some work can be shown to be recursive. In particular the system for ordinals up to $\phi_{\epsilon(\Omega+1)}$ was used by Howard to determine the ordinal of the system ID_1, and analogous higher Bachmann-style ordinal systems were used by Pohlers (cf. the history and references in Buchholz et al 1981) to determine the ordinals of systems of iterated inductive definitions ID_ν.

4° **Long hierarchies without fundamental sequences.** The Bachmann-style hierarchies of functions on Ω are *long* in the sense that they are indexed by ordinals going beyond Ω. Around 1970 I proposed a way of obtaining long hierarchies without fundamental sequences. Preliminary match-ups with ordinals generated by Bachmann-style systems were then obtained independently by Aczel and Weyrauch. None of this work was published. (Weyrauch's thesis, written in 1972, was not submitted until 1975.) Building on Aczel's notes, Jane Bridge in her 1972 Oxford thesis pushed the match-ups into higher inaccessibles. She made the first moves to establish recursiveness of the associated systems. Buchholz the gave a full treatment of recursiveness. (For detailed references and more of the history, see the introduction to Buchholz et al (1981) and Feferman (1987).)

What are now called the Feferman-Aczel systems allow one to start off with specific Ω_αs or functions enumerating such, etc. For example we could start with the single ordinal Ω or we could start with the function $\lambda\xi.\Omega_\xi$, or even larger inaccessible ordinals. For simplicity, and to bring out some of the issues most clearly, we consider the simplest case only, obtained by simply adding Ω. A sequence of functions θ_α for α an arbitrary ordinal is defined by recursion on α. Thus we assume given $\langle \theta_\xi \rangle_{\xi < \alpha}$ which we denote $\underline{\theta}_\alpha$. Next define the closure $C(\alpha, \gamma)$ under $\underline{\theta}_\alpha$ inductively as follows:

(i) $\gamma \cup \{0, \Omega\} \subseteq C(\gamma)$.

(ii) $\xi, \eta \in C(\alpha,\gamma) \Rightarrow \xi + \eta, \phi(\xi,\eta) \in C(\alpha,\gamma)$

(iii) $\xi < \alpha \,\&\, \xi \in C(\alpha,\gamma) \& \, \eta \in C(\alpha,\gamma) \Rightarrow \theta_\xi(\eta) \in C(\alpha,\gamma)$.

We say that γ is α-closed if $C(\alpha,\gamma) \cap \Omega_{\delta+1} = \gamma$ when γ belongs to the number class interval $[\Omega_\delta, \Omega_{\delta+1}]$. Then we define:

(iv) θ_α enumerates $\{\gamma : \gamma \text{ is } \alpha\text{−closed}\}$.

We may think of the sequence of functions $\langle\theta_\alpha\rangle$ as a kind of long Veblen hierarchy using just the symbol for Ω. In general, for $\gamma \in [\Omega_\delta, \Omega_{\delta+1}]$, $C(\alpha,\gamma)$ will stretch beyond γ into the same number class interval to the least $\zeta \geq \gamma$ with $\zeta \notin C(\alpha,\zeta)$, but will have $\max(\aleph_0, \mathrm{card}(\gamma))$ elements in higher number classes. These will appear in stretches with gaps in between. However, it follows by the definition and this cardinality fact that each θ_α preserves each number class, just like the original Veblen hierarchy. Define $\theta(\alpha,\xi) = \theta_\alpha(\xi)$. Now we can form a notation system based on this sequence of functions, by closing $\{0\}$ under the constant Ω and the functions $+, \phi$, and θ, and intersecting the result with Ω. The least ordinal not thus obtained is $\theta_\alpha(0)$, where $\alpha = \Gamma_{\Omega+1}$ is the first Γ number beyond Ω. This system thus comprehends the Howard ordinal for ID_1.

5° **Systems with collapsing functions.** In order to make the verification of recursiveness for the systems based on long hierarchies of the sort just described less complicated, Buchholz made some simplifications in the definitions, such as the following which yields a notation system with the same ordinal as in 4°. This also shifts the attention from long hierarchies à la Veblen-Feferman-Aczel, to so-called *collapsing functions* such as the function ψ introduced by the following. These turn out to be crucial for current infinitary proof theory of subsystems of analysis and set theory. For each α, we define sets of ordinals $\mathrm{B}(\alpha)$ and numbers $\psi(\alpha)$ inductively as follows:

(i) $\{0, \Omega\} \subseteq B(\alpha)$.

(ii) $\xi, \eta \in B(\alpha) \Rightarrow \xi + \eta, \phi(\xi,\eta) \in B(\alpha)$.

(iii) $\xi < \alpha \,\&\, \xi \in B(\alpha) \Rightarrow \psi(\xi) \in B(\alpha)$

(iv) $\psi(\alpha) = \min\{\xi : \xi \notin B(\alpha)\}$.

Then the closure of $\{0\}$ under $\Omega, +, \phi$ and ψ, intersected with Ω, is the same as the ordinal described at the end of the preceding section. ψ is called a collapsing function since in this case *all* its values are countable. (There is also a relation with the Mostowski collapse.) See Pohlers (1989) for detailed treatment of this specific system. Schütte (1977), Ch. IX, gives a system using $\lambda\xi.\Omega_\xi$, and Jäger and Pohlers (1982) made use of a system with the further addition of the first inaccessible for the proof theory of $\sum_2^1$-AC+BI. Subsequently, Rathjen (1991) pushed this out to the use of the first Mahlo cardinal, and most recently (1995,1996) he has found applications of ordinal systems employing names for supercompact cardinals to the proof theory of $\prod_2^1$-CA and related systems of set theory.

General properties of such systems of functions. Suppose given a sequence $\mathbf{f}$ of functions defined for arbitrary ordinals, and suppose given an ordinal γ. We define $Cl(\mathbf{f};\gamma)$ to be the least set containing $\{0\} \cup \gamma$ and closed under $\mathbf{f}$. γ is said to be an $\mathbf{f}$-*inaccessible* if $Cl(\mathbf{f};\gamma) = \gamma$. The class of $\mathbf{f}$-inaccessibles is closed and unbounded, and so is enumerated by a normal function, denoted $\mathbf{f}'$; the passage from $\mathbf{f}$ to the adjunction $\mathbf{f} * \mathbf{f}'$ of this function is called the *critical process*. By Term($\mathbf{f}$, X) where X is a set of individual variables, is meant the set of terms generated from the symbol 0 and the variables in X by the formal function symbols for $\mathbf{f}$. The system of notation associated with $\mathbf{f}$ is then just Term($\mathbf{f}, \emptyset$). In Feferman (1968) I introduced the following notions:

I. $\mathbf{f}$ is said to be *complete* if $Cl(\mathbf{f};0)$ is an ordinal, i.e. fills up an initial segment of the ordinals. $\mathbf{f}$ is said to be *replete* if for any $\gamma, Cl(\mathbf{f};\gamma)$ is an ordinal, i.e. stretches γ to an initial segment of the ordinals.

II. $\mathbf{f}$ is said to be *effective*, if the natural ordering of terms in Term($\mathbf{f}, \emptyset$) is recursive. $\mathbf{f}$ is said to be *effectively relatively categorical (e.r.c.)* if whenever X is a finite set of variables and σ is an assignment of values to the members of X in the class of $\mathbf{f}$-inaccessibles, then the natural ordering of Term($\mathbf{f}$, X) is recursive uniformly in the order diagram of σ.

The properties of completeness and effectiveness are not in general preserved, when one passes from a system $\mathbf{f}$ to the system consisting of $\mathbf{f}$ and $\mathbf{f}'$. But it was proved op. cit. that the properties of being replete and being e.r.c. are preserved under transfinite iteration of the critical process. As examples, the systems (χ) and (χ, ϕ) or $(+, \phi)$ are replete and e.r.c.

The work Feferman (1968) preceded the introduction of long hierarchies described in 4° above and later simplified in 5°. The notions treated there do not apply directly, because the ordinals generated have gaps, but that is taken care of by a slight modification of the definitions: for $\gamma \in [\Omega_\delta,\ \Omega_{\delta+1}]$, we define $Cl_0(\mathbf{f}, \gamma)$ to be the intersection of $Cl(\mathbf{f}, \gamma)$ with $\Omega_{\delta+1}$. Then $\mathbf{f}$ is said to be replete if $Cl_0(\mathbf{f}, \gamma)$ is an ordinal. Also, the $\mathbf{f}$-inaccessibles are defined to be those γ such that $\gamma \notin Cl_0(\mathbf{f}, \gamma)$ and $\mathbf{f}$ is said to be e.r.c. as before. It is morally certain (I have not checked the details) that the systems $(+, \phi, \theta)$ of 4° and $(+, \phi, \psi)$ of 5° are replete and e.r.c. What's new in these situations is that we have an *auxiliary set of terms* using symbols from higher number classes when forming the system $Cl_0(\mathbf{f}, \emptyset)$, namely just those naming the elements of $Cl(\mathbf{f}, \emptyset)$ which lie beyond Ω. This use of auxiliaries is a bit mysterious, and we shall return to it below. But let us look more speculatively at what would be considered **good properties** of such systems produced by $\mathbf{f}$'s which are both replete and e.r.c.; these are admittedly vague.

(1) **Normal forms.** There is somehow a naturally described subset of terms which denote all the ordinals generated in $Cl_0(\mathbf{f}, \emptyset)$, no two of which denote the same ordinal. Can we give some theoretical criteria for what constitute normal forms? Presumably, each is given by a term built up by constituent terms representing smaller arguments, but this is not enough. As shown in Feferman (1968), e.r.c. systems determine functors on inaccessibles of a system, which preserve limits. Girard (1981) defined a notion of *dilator*, which is a functor on orderings that preserves limits and pull-backs. This notion corresponds somehow to having a unique system of representation. So a candidate for an $\mathbf{f}$ whose associated system has normal forms is one which is e.r.c. and whose associated functor is a dilator.

(2) **Retracing functions.** Once one discerns normal forms, one can define (what Kreisel called) *retracing functions* which tell how each ordinal is built up from smaller ordinals. For example, in the Cantor system (χ) for ϵ_0, we associate with each $\gamma < \epsilon_0$, for which $\gamma \neq 0$, ordinals α, β and a positive integer k with $\gamma = \omega^\alpha \cdot k + \beta$ and $\beta < \omega^\alpha$. If this system is extended by the function $\lambda\xi.\epsilon_\xi$, the retracing functions depend on whether $\omega^\gamma = \gamma$ or not; in the latter case, we use the preceding retracing functions, while in the latter case, we associate α with γ where $\gamma = \epsilon_\alpha$—and so on for further iterations of the critical process.

(3) **Replacing transfinite recursion by ordinary recursion.** In some suitable sense, once one has retracing functions, further functions

which are defined by "ordinary" transfinite recursions on the given ordinal and under which it is closed, can be defined by finite recursions instead. Example: on the Cantor system for ϵ_0 we can define ordinal multiplication by a finite recursion.

(4) **Maximality.** Put in other terms, a *good system* $\mathbf{f}$ *for an ordinal* α is one which gives a representation system for α and which is *maximal* in the sense that everything that can be defined by "ordinary" transfinite recursion on the ordinal can be replaced by a finite recursion. Two different systems for the same ordinal which are maximal in this sense should be interchangeable.

The mysterious role of the higher number classes and other auxiliaries. In some sense, all we are doing in building these systems of functions is iterating the critical process, which simply means that we are adding at each "stage" the first unnamed ordinal at that stage. Unlike 1° and 2°, what happens in 4° and 5° is that there are big gaps in the ordinals represented, and that what we are really after is the first segment with no gaps, and it's that that we keep closing up. Now admittedly, the collapsing function ψ is of proof-theoretical use to us outside of this segment, but if we are *just* after this initial segment, the use of ordinals from higher number classes appears mysterious. In Feferman (1970) I suggested a different way of getting at the same segment, via higher types over Ω, rather than higher number classes. Then the notion of repleteness can be extended to higher types as a notion of hereditary repleteness, and one can think of iteration as a functional in each type ≥ 2 which iterates through Ω any functional of next lower type. So as we ascend in types we can deal with the iteration of the critical process, its hyper-iteration etc. Now (as I conjectured there and was proved by Weyrauch in his thesis) the ordinal obtained in this way when one starts with the function $\lambda\xi(1+\xi)$ and uses iteration of the critical process, hyper-iteration, etc. through all finite types, is just the Howard ordinal. At my further suggestion, Aczel (1972) extended this to transfinite types, but he concluded that this approach would be limited relatively low down in the Bachmann hierarchies, though of course well beyond the Howard ordinal. This limitation depended on a specific way of passing to transfinite types, and it still seems to me that there may be alternative ways of doing so which move us beyond the Aczel limit. One advantage to looking at things this way is that one can equally well consider hereditarily replete functionals and iteration of the critical process over ω_1^{CK}, to obtain directly from the set-up that the ordinal segment generated is recursive. Currently, for systems of style 4° or 5° this requires a special argument in each case. But perhaps here in any case, we're barking up the wrong tree, since the collapsing functions on higher ordinals seem to be needed in any case.

III. "Large" "cardinals"

We observe that various notions of large cardinals: inaccessible, Mahlo, indescribables, and still higher, have reasonable analogues in admissible set theory and in the construction of systems of ordinal function as in section II. The problem is whether the reason for the success of such analogues can be located in the existence of a more general framework in which each of these appears as a special case. I believe the prospects for doing this in the case of ordinary (impredicative) set theory and admissible set theory are very good, and I have suggested how that might be done in my paper (already available) for the Gödel '96 conference following soon on the heels of this one. So I won't repeat any details here, but simply say that the idea is to reformulate both kinds of set theory by the *addition* of variables for *partial functions* whose domain will in general classes (possibly all of V), and the addition of specific constants, functions and functionals: in the case of admissible set theory, these are the constants 0 and ω, unordered pair, union and the characteristic function of the $\in$ relation, together with functionals S and R for Separation and Replacement respectively, with axioms:

(S) $\quad \forall x \in a[f(x)\downarrow] \rightarrow S(f,a)\downarrow \wedge \forall x[x \in S(f,a) \leftrightarrow x \in a \wedge f(x) = 0].$

(R) $\quad \forall x \in a[f(x)\downarrow] \rightarrow R(f,a)\downarrow \wedge \forall y[y \in R(f,a) \leftrightarrow \exists x \in a(f(x) = y)].$

For impredicative set theory we also add the power set function and a functional for universal quantification, which allows us to assign a characteristic function to every definable class. Both systems also have the Axiom Scheme of Foundation. Now to the extent that large cardinal notions can be formulated within the common language, we see how to generalize the two situations. So the obvious thing to look for now is an appropriate set theory, still weaker than admissible, which is just what we need to establish the above kinds of hierarchies of ordinal functions as in 4° and 5° and in which we can interpret the least uncountable ordinal as ω_1^{CK}. The essential difference has to be that we allow for *collapse*. In admissible set theory, we can't use functions defined in terms of higher number classes to yield "countable" ordinals. It has been shown by Rathjen and Schlüter (see Pohlers (1996) p. 186 for references) that the Ω_α in the various systems of ordinal representation with collapsing functions indicated above, may be interpreted as the admissible ordinals ω_α, though there is no principled basis for this conclusion in recursion theory on admissible sets as currently developed. So the problem is, whether there is a weaker form of the theory in which each "regular" ordinal is regular with respect to *arbitrary* functions produced, not just those defined in a restricted way on each "number class".

For a rather different kind of generalization, but in the same spirit of seeking notions of large set or large cardinal applicable to a variety of "environments", see Griffor and Rathjen (1996).

References

P. Aczel (1972), Describing ordinals using functionals of transfinite type, *JSL 37*, 35-47.

H. Bachmann (1950), Die Normalfunktionen und das Problem des ausgezeichneten Folgen von Ordnungszahlen, *Viertelj. Nat. Ges. Zürich 95*, 5-37.

W. Buchholz, S. Feferman, W. Pohlers and W. Sieg (1981), *Iterated inductive definitions and subsystems of analysis: Recent proof-theoretical studies, Lecture Notes in Maths. 897.*

S. Feferman (1968), Systems of predicative analysis II. Representations of ordinals, *JSL 33*, 193-220.

S. Feferman (1969), Set-theoretical foundations of category theory (with an appendix by G. Kreisel), in *Reports of the Midwest Category Seminar III, Lecture Notes in Maths. 106*, 201-247.

S. Feferman (1970), Hereditarily replete functionals over the ordinals, in *Intuitionism and Proof Theory*, 289-301.

S. Feferman (1974), Some formal systems for the unlimited theory of structures and categories, (unpublished MS, Abstract in *J. Symbolic Logic 39*, 374-375).

S. Feferman (1975), A language and axioms for explicit mathematics, in *Algebra and Logic, Lecture Notes in Maths. 450*, 87-139.

S. Feferman (1977), Categorical foundations and foundations of category theory, in *Logic, Foundations of Mathematics and Computability Theory, Vol. I*, 149-169.

S. Feferman (1979), Constructive theories of functions and classes, in *Logic Colloquium '78*, 159-224.

S. Feferman (1987), Proof theory: A personal report, Appendix to G. Takeuti, *Proof Theory, 2nd edn.*, 447-485.

S. Feferman (1996), Gödel's program for new axioms, Why, where, how and what?, *Gödel '96, Lecture Notes in Logic 6*, 3-22.

J.-Y. Girard (1981), $\prod^1_2$-logic, Part I: Dilators, *Annals Math. Logic 21*, 75-219.

E. R. Griffor and M. Rathjen (1996), Inaccessibility in constructive set theory and type theory, Univ. of Leeds 1996 Preprint Series No. 19.

D. Isles (1970), Regular ordinals and normal forms, in *Intuitionism and Proof Theory*, 339-361.

G. Jäger and W. Pohlers (1982), Eine beweistheoretische Untersuchung von (Δ^1_2-CA)+(BI) und verwandter Systeme, *Sitzungsberichte Bayer. Akad. Wiss.*

R. Jensen (1969), On the consistency of a slight (?) modification of Quine's *New Foundations*, in *Words and Objections: Essays on the Work of W.V.O.Quine*, 278-291.

S. MacLane (1961), Locally small categories and the foundations of mathematics, in *Infinitistic Methods*, 25-43.

W. Pohlers (1989), *Proof Theory: An Introduction, Lecture Notes in Maths. 1407*

W. Pohlers (1996), Pure proof theory. Aims, methods and results, *Bull. Symbolic Logic 2*, 159-188.

M. Rathjen (1991), Proof theoretic analysis of KPM, *Archive for Math. Logic 30*, 377-403.

M. Rathjen (1995), Recent advances in ordinal analysis: $\prod^1_2$-CA and related systems, *Bull. Symbolic Logic 1*, 468-485.

M. Rathjen (1996), The higher infinite in proof theory, Univ. of Leeds Preprint Series No. 20 (to appear in Proc. *Logic Colloquium '95).*

A. Schlüter (1994), *Zur Mengenexistenz in formalen Theorien der Mengenlehre*, Ph.D. Thesis, Münster.

K. Schütte (1977), *Proof Theory.*