Unfolding schematic systems*

Thomas Strahm**

January 14, 2016

Abstract

The notion of unfolding a schematic formal system was introduced by Feferman in 1996 in order to answer the following question: *Given a schematic system* S*, which operations and predicates, and which principles concerning them, ought to be accepted if one has accepted* S? After a short summary of precursors of the unfolding program, we survey the unfolding procedure and discuss the main results obtained for various schematic systems S, including non-finitist arithmetic, finitist arithmetic, feasible arithmetic, and theories of inductive definitions.

1 Introduction

The search for new axioms which are *exactly as evident and justified as those with which you have started* was already advocated by Gödel in his program for new axioms, see Gödel [24], p. 151:

> Let us consider, e.g., the concept of demonstrability. It is well known that, in whichever way you make it precise by means of a formalism, the contemplation of this very formalism gives rise to new axioms which are exactly as evident and justified as those with which you started, and this process of extension can be iterated into the transfinite. So there cannot exist any formalism which would embrace all these steps, but this does not exclude that all these steps (or at least all of them which give something new for the domain of propositions in which you are interested) could be described and collected together in some non-constructive way.

*To Solomon Feferman, with gratitude for his intellectual inspiration and friendship

**Institut für Informatik und angewandte Mathematik, Universität Bern, Neubrückstrasse 10, CH-3012 Bern, Switzerland. Email: strahm@iam.unibe.ch

The first very natural candidates for new axioms to be added to an arithmetical system S are proof-theoretic reflection principles, roughly stating that everything which is provable in S is correct. More precisely, the local reflection schema is the collection of sentences

$$(\mathsf{Rfn}_\mathsf{S}) \qquad \mathsf{Prov}_\mathsf{S}(\ulcorner A \urcorner) \to A$$

for A being a sentence in the language of S. The generalization of this schema to arbitrary formulas uniformly in their free parameters is called the uniform reflection schema of S, in symbols, RFN_S. As was shown by Turing [36], one may iterate the addition of reflection principles (and consistency statements) along Kleene's constructive ordinal notations $\mathcal{O}$ in order to define for each $a \in \mathcal{O}$ a formal system S_a by adding the reflection principle of the previous system at successor stages and taking the union of the previous systems at limit stages. Turing called the so-obtained progressions of a given system S *ordinal logics*. These were taken up in Feferman [10] and there renamed to *transfinite recursive progressions of axiomatic theories*. While Turing obtained completeness results for Π^0_1 sentences by iteration of the consistency or local reflection principle Rfn_S, Feferman showed that one gets completeness for all arithmetic statements by iteration of the uniform reflection principle RFN_S. Both completeness results were considered to be problematic because they depend on clever choices of ordinal notations which were not justified on previously accepted grounds. Indeed, ordinal logics are far from being invariant under the choice of ordinal notation: Feferman and Spector [19] have shown that there are paths through $\mathcal{O}$ whose progression is not even complete for Π^0_1 sentences. For more information on ordinal logics, see Feferman [17] and Franzen [23].

The crucial condition which was missing in the previous proposals is the one of *autonomy* which guarantees that one is only allowed to advance to a system S_a for an $a \in \mathcal{O}$ in case the wellfoundedness of a has been established in a system S_b with b smaller than a; see Kreisel [26] and Feferman [11]. Thus we are led to the study of *all principles of proof and ordinals which are implicit in given concepts*, see Kreisel [28]. The most influential series of results in the "autonomy program" concerns the study of the limits of predicative provability by Feferman [11] and Schütte [31] who independently determined the so-called Feferman-Schütte ordinal Γ_0 as the limiting number of predicativity. The first system proposed for an analysis of predicativity was autonomous ramified analysis. After its ordinal Γ_0 had been found, Feferman developed (autonomous) progressions of hyperarithmetical analysis based on the hyperarithmetic comprehension rule and the uniform reflection principle as well as the system IR for inductive-recursive analysis.

One objection with the above-described approaches in the implicitness program one may have is the inclusion of the notion of ordinal or wellordering, which is not prima-facie implicit in our conception of the natural numbers or arithmetic. In his search for "a more perspicuous system for predicativity", Feferman [13] came up with a natural system capturing ramified analysis in levels less than Γ_0 without presupposing any notion of ordinal at the outset. Crucial in this system is the fact that arithmetic is treated as a *schematic system* with *free* predicate P and induction in schematic form,

$$P(0) \wedge (\forall x)[P(x) \rightarrow P(x')] \rightarrow (\forall x)P(x),$$

together with a rule of predicate substitution (Subst),

$$(\mathsf{Subst}) \qquad A[P] \;\Rightarrow\; A[B/P],$$

expressing that the whenever we derive a statement $A[P]$ possibly containing the free predicate P, we also accept all its substitution instances by any formula B. We note that a crucial feature of schemata as understood here is their *openendedness*, i.e. "they are not conceived of as applying to a specific language whose stock of basic symbols is fixed in advance, but rather as applicable to any language which one comes to recognize as embodying meaningful basic notions" (Feferman [15], p. 9).

A general notion of reflective closure of an arbitrary schematic formal system S was proposed by Feferman in a lecture for a meeting in 1979 on the work of Kurt Gödel, which was only published in 1991, see Feferman [14]. The basic observation underlying the reflective closure procedure in [14] is that the informal reasoning about what is implicit in S makes use of a notion of truth for S, which then leads us to also reason about statements involving truth and so on. The technical apparatus of the reflective closure is governed by what has become to be known as the Kripke-Feferman axioms of partial truth, rooted in Kripke's semantic theory of truth, see Kripke [29] and the article of Cantini, Fujimoto, and Halbach in this volume. The main result obtained in Feferman [14] is that the full reflective closure of Peano arithmetic, $\mathsf{Ref}^*(\mathsf{PA})$, is proof-theoretically equivalent to predicative analysis $\mathsf{RA}_{<\Gamma_0}$, where $\mathsf{Ref}^*(\mathsf{PA})$ includes a suitable version of the substitution rule (Subst).

As Feferman writes in [18], the axiomatic theory of truth "as an engine for the explanation of reflective closure still has an air of artificiality": It is at least questionable whether the axioms of truth are exactly as evident as those of the given system S. Also, the fact that some amount of arithmetic is presupposed in order to describe the coding machinery is not very pleasing.

Given that a schematic system is formulated using function and predicate symbols in a given logical language, it is more attractive to expand and study the *operations* on individuals and predicates, which are implicit in the acceptance of S. This led Feferman [15] to his most recent proposal in the implicitness program, namely the notion of unfolding of an open-ended schematic system whose main aim is to answer the following question: *Given a schematic system* S*, which operations and predicates – and which principles concerning them – ought to be accepted if one has accepted* S*?*. The notion of unfolding has been applied to non-finitist-arithmetic (Feferman and Strahm [20]), finitist arithmetic (Feferman and Strahm [21]), feasible arithmetic (Eberhard and Strahm [9]), and theories of inductive definitions (Buchholtz [2]). The aim of this paper is to describe the unfolding procedure in detail and discuss the main results obtained for various schematic formal systems S.

2 The unfolding of non-finitist arithmetic

The aim of this section is to spell out the unfolding procedure in detail for the case of non-finitist arithmetic NFA, and state the main results obtained in Feferman and Strahm [20]. We follow the presentation in Feferman and Strahm [21].

The schematic system for classical non-finitist arithmetic, NFA, is defined as follows. Its basic operations on individuals with the constant 0 are successor, Sc, and predecessor, Pd; the basic logical operations are $\neg$, $\wedge$, and $\forall$. It is given by the following axioms, where we write as usual, x' for $\mathsf{Sc}(x)$:

(1) $x' \neq 0$

(2) $\mathsf{Pd}(x') = x$

(3) $P(0) \wedge (\forall x)[P(x) \rightarrow P(x')] \rightarrow (\forall x)P(x)$.

Here P is a free predicate variable, and the intention is to use the induction scheme (3) in a wider sense than is limited by the basic language of NFA or any language fixed in advance. Namely, one applies the general rule of substitution

(Subst) $$A[P] \;\Rightarrow\; A[B/P]$$

to any formulas A and B that arise in the process of unfolding NFA.

In a first step, we shall describe the unfolding of a schematic system S informally by stating some general methodological "pre-axioms". Then we shall spell out these axioms in all detail for S being the schematic system NFA.

Underlying the idea of unfolding for arbitrary S are general notions of (partial) *operation* and *predicate*, belonging to a universe V extending the universe of discourse of S. These have to be thought of as intensional entities, given by rules of computation and defining properties, respectively. Operations have to be considered as pre-mathematical in nature and not bound to any specific mathematical domain. They can apply to other operations as well as to predicates. Some operations are universal and are naturally self-applicable as a result, like the identity operation or the pairing operation, while some are partial and presented to us on specific mathematical domains only. Operations satisfy the laws of a *partial combinatory algebra* with pairing, projections, and definition by cases. Predicates are equipped with a membership relation $\in$ to express that given elements satisfy the predicate's defining property.

For the formulation of the full unfolding $\mathcal{U}(\mathsf{S})$ of any given schematic axiom system S, we have the following axioms.

1. The universe of discourse of S has associated with it an additional unary relation symbol, $\mathsf{U_S}$, and the axioms of S are relativized to $\mathsf{U_S}$.

2. Each n-ary function symbol f of S determines an element $f^\star$ of our partial combinatory algebra, with $f(x_1, \ldots, x_n) = f^\star x_1 \ldots x_n$ on U_S^n (or the domain of f in case f itself is given as a partial function).

3. Each relation symbol R of S together with $\mathsf{U_S}$ determines a predicate $R^\star$ with $R(x_1, \ldots, x_n)$ if and only if $(x_1, \ldots, x_n) \in R^\star$.

4. Operations on predicates, such as e.g. conjunction, are just special kinds of operations. Each logical operation l of S determines a corresponding operation $l^\star$ on predicates.

5. Sequences of predicates given by an operation f form a new predicate $\mathit{Join}(f)$, the disjoint union of the predicates from f.

Moreover, the free predicate variables $P, Q, \ldots$ used in the schematic formulation of S give rise to the crucial *rule of substitution* (Subst), according to which we are allowed to substitute any formula B for P in a previously recognized (i.e. derived) statement $A[P]$ depending on P.

The restriction $\mathcal{U}_0(\mathsf{S})$ of $\mathcal{U}(\mathsf{S})$ is obtained by dropping the axioms concerning predicates; $\mathcal{U}_0(\mathsf{S})$ is called the operational unfolding of S. Moreover, there is a natural intermediate predicate unfolding system $\mathcal{U}_1(\mathsf{S})$, which is simply $\mathcal{U}(\mathsf{S})$ without the predicate forming operation of *Join*.

The following spells out in detail the three unfolding systems $\mathcal{U}_0(\mathsf{S})$, $\mathcal{U}_1(\mathsf{S})$, and $\mathcal{U}(\mathsf{S})$ for $\mathsf{S} = \mathsf{NFA}$, the schematic system of non-finitist arithmetic introduced above. We begin with the operational unfolding $\mathcal{U}_0(\mathsf{NFA})$. Its language is first order, using variables $a, b, c, f, g, h, u, v, w, x, y, z \ldots$ (possibly with subscripts). It includes (i) the constant 0 and the unary function symbols Sc and Pd of NFA, (ii) constants for operations as individuals, namely sc, pd (successor, predecessor), k, s (combinators), $\mathsf{p}, \mathsf{p_0}, \mathsf{p_1}$ (pairing and unpairing), $\mathsf{d}, \mathsf{tt}, \mathsf{ff}$ (definition by cases, true, false), and e (equality), and (iii) a binary function symbol $\cdot$ for (partial) term application. Further, we have (iv) a unary relation symbol $\downarrow$ (defined) and a binary relation symbol $=$ (equality), as well as (v) a unary relation symbol N (natural numbers). In addition, we have a symbol $\bot$ for the false proposition. Finally, a stock of free predicate symbols $P, Q, R, \ldots$ of finite arities is assumed.[1]

The *terms* $(r, s, t, \ldots)$ of $\mathcal{U}_0(\mathsf{NFA})$ are inductively generated from the variables and constants by means of the function symbols Sc, Pd, as well as $\cdot$ for application. In the following we often abbreviate $(s \cdot t)$ simply as (st), st or sometimes also $s(t)$; the context will always ensure that no confusion arises. We further adopt the convention of association to the left so that $s_1 s_2 \ldots s_n$ stands for $(\ldots(s_1 s_2) \ldots s_n)$. Further, we put $t' := \mathsf{Sc}(t)$ and $1 := 0'$. We define general n-tupling by induction on $n \geq 2$ as follows:

$$(s_1, s_2) := \mathsf{p} s_1 s_2, \qquad (s_1, \ldots, s_{n+1}) := ((s_1, \ldots, s_n), s_{n+1}).$$

Moreover, we set $(s) := s$ and $() := 0$.

The *formulas* $(A, B, C, \ldots)$ of $\mathcal{U}_0(\mathsf{NFA})$ are inductively generated from the atomic formulas $\bot$, $s\downarrow$, $(s = t)$, $\mathsf{N}(s)$, and $P(s_1, \ldots, s_n)$ by means of negation $\neg$, conjunction $\wedge$, and universal quantification $\forall$. The remaining logical connectives and quantifiers are defined as usual by making use of classical logic.

The sequence notation $\bar{u}$ and $\bar{t}$ is used in order to denote finite sequences of variables and terms, respectively. Moreover, we write $t[\bar{u}]$ to indicate a sequence $\bar{u}$ of free variables possibly appearing in the term t; however, t may contain other variables than those shown by using this bracket notation. Further, $t[\bar{s}]$ is used to denote the result of simultaneous substitution of the terms $\bar{s}$ for the variables $\bar{u}$ in the term $t[\bar{u}]$. The meaning of $A[\bar{u}]$ and $A[\bar{s}]$ is

[1]The constants sc and pd as well as the relation symbol N are used instead of the symbols $\mathsf{Sc}^\star$, $\mathsf{Pd}^\star$, and $\mathsf{U}_{\mathsf{NFA}}$ mentioned in the informal description above.

understood accordingly. Finally, we shall also use the sequence notation $\bar{A}$ in order to denote a finite sequence $\bar{A} = A_1, \ldots, A_n$ of formulas.

$\mathcal{U}_0(\mathsf{NFA})$ is based on *partial* term application. Hence, it is not guaranteed that terms have a value, and $t\!\downarrow$ is read as "t is defined" or "t has a value". Accordingly, the *partial equality relation* $\simeq$ is introduced by

$$s \simeq t \;:=\; (s\!\downarrow \vee\, t\!\downarrow) \to (s = t).$$

Further, we shall use the following abbreviations concerning the predicate N for the natural numbers ($\bar{s} = s_1, \ldots, s_n$):

$$\begin{array}{rcl} \bar{s} \in \mathsf{N} & := & \mathsf{N}(s_1) \wedge \cdots \wedge \mathsf{N}(s_n), \\ (\exists x \in \mathsf{N})A & := & (\exists x)(x \in \mathsf{N} \wedge A), \\ (\forall x \in \mathsf{N})A & := & (\forall x)(x \in \mathsf{N} \to A). \end{array}$$

The logic of $\mathcal{U}_0(\mathsf{NFA})$ is the *classical logic of partial terms* LPT of Beeson [1], cf. also Feferman [12]. We recall that LPT embodies strictness axioms saying that all subterms of a defined compound term are defined as well. Moreover, if $(s = t)$ holds then both s and t are defined, and s is defined provided $\mathsf{N}(s)$ holds, and similarly for $P(\bar{s})$.

The axioms of $\mathcal{U}_0(\mathsf{NFA})$ are divided into three groups as follows.

I. Embedding of NFA

(1) The relativization of the axioms of NFA to the predicate N,[2]

(2) $(\forall x \in \mathsf{N})[\mathsf{Sc}(x) = \mathsf{sc}(x) \;\wedge\; \mathsf{Pd}(x) = \mathsf{pd}(x)]$.

II. Partial combinatory algebra, pairing, definition by cases

(3) $\mathsf{k}ab = a$,

(4) $\mathsf{s}ab\!\downarrow \;\wedge\; \mathsf{s}abc \simeq ac(bc)$,

(5) $\mathsf{p_0}(a,b) = a \;\wedge\; \mathsf{p_1}(a,b) = b$,

(6) $\mathsf{d}ab\,\mathsf{tt} = a \;\wedge\; \mathsf{d}ab\,\mathsf{ff} = b$.

III. Equality on the natural numbers N

(7) $(\forall x, y \in \mathsf{N})[\mathsf{e}xy = \mathsf{tt} \;\vee\; \mathsf{e}xy = \mathsf{ff}]$,

(8) $(\forall x, y \in \mathsf{N})[\mathsf{e}xy = \mathsf{tt} \;\leftrightarrow\; x = y]$.

[2] Note that this relativization also includes axioms such as $0 \in \mathsf{N}$ and $(\forall x \in \mathsf{N})(x' \in \mathsf{N})$.

Finally, crucial for the formulation of $\mathcal{U}_0(\mathsf{S})$ is the predicate substitution rule:

$$\text{(Subst)} \qquad A[\bar{P}] \;\Rightarrow\; A[\bar{B}/\bar{P}].$$

Here $\bar{P} = P_1, \ldots, P_m$ is a sequence of free predicate symbols possibly appearing in the formula $A[\bar{P}]$ and $\bar{B} = B_1, \ldots, B_m$ is a sequence of formulas. In the conclusion of this rule of inference, $A[\bar{B}/\bar{P}]$ denotes the formula $A[\bar{P}]$ with each subformula $P_i(\bar{t})$ replaced by $(\exists \bar{x})(\bar{t} = \bar{x} \wedge B_i[\bar{x}])$, where the length of $\bar{x}$ equals the arity of P_i.

We now turn to the full predicate unfolding $\mathcal{U}(\mathsf{NFA})$ and its restriction $\mathcal{U}_1(\mathsf{NFA})$.

The language of $\mathcal{U}(\mathsf{NFA})$ extends the language of $\mathcal{U}_0(\mathsf{NFA})$ by additional constants nat (natural numbers), eq (equality), pr_P (free predicate P), inv (inverse image), neg (negation), conj (conjunction), un (universal quantification), and join (disjoint unions). In addition, we have a new unary relation symbol Π for (codes of) predicates and a binary relation symbol $\in$ for expressing elementhood between individuals and predicates, i.e. satisfaction of those predicates by the given individuals. The *terms* of $\mathcal{U}(\mathsf{NFA})$ are generated as before but now taking into account the new constants. The *formulas* of $\mathcal{U}(\mathsf{NFA})$ extend the formulas of $\mathcal{U}_0(\mathsf{NFA})$ by allowing new atomic formulas of the form $\Pi(t)$ and $s \in t$.

The axioms of $\mathcal{U}(\mathsf{NFA})$ extend those of $\mathcal{U}_0(\mathsf{NFA})$, as follows.

IV. Basic axioms about predicates

(9) $\Pi(\mathsf{nat}) \wedge (\forall x)(x \in \mathsf{nat} \leftrightarrow \mathsf{N}(x))$,[3]

(10) $\Pi(\mathsf{eq}) \wedge (\forall x)(x \in \mathsf{eq} \leftrightarrow (\exists y)(x = (y, y)))$,

(11) $\Pi(\mathsf{pr}_P) \wedge (\forall \bar{x})((\bar{x}) \in \mathsf{pr}_P \leftrightarrow P(\bar{x}))$,

(12) $\Pi(a) \rightarrow \Pi(\mathsf{inv}(a, f)) \wedge (\forall x)(x \in \mathsf{inv}(a, f) \leftrightarrow fx \in a)$,

(13) $\Pi(a) \rightarrow \Pi(\mathsf{neg}(a)) \wedge (\forall x)(x \in \mathsf{neg}(a) \leftrightarrow x \notin a)$,

(14) $\Pi(a) \wedge \Pi(b) \rightarrow \Pi(\mathsf{conj}(a, b)) \wedge (\forall x)(x \in \mathsf{conj}(a, b) \leftrightarrow x \in a \wedge x \in b)$,

(15) $\Pi(a) \rightarrow \Pi(\mathsf{un}(a)) \wedge (\forall x)(x \in \mathsf{un}(a) \leftrightarrow (\forall y \in \mathsf{N})((x, y) \in a))$.

V. Join axiom

(16) $(\forall x \in \mathsf{N})\Pi(fx) \rightarrow \Pi(\mathsf{join}(f)) \wedge (\forall x)(x \in \mathsf{join}(f) \leftrightarrow J[f, x])$,

[3] Observe that nat is alternatively definable from the remaining predicate axioms by $x \in \mathsf{nat} \leftrightarrow (\exists y \in \mathsf{N})(x = y)$.

where $J[f,u]$ expresses that u is an element of the disjoint union of f over N, i.e.

$$J[f,u] \;:=\; (\exists y \in \mathsf{N})(\exists z)(u = (y,z) \wedge z \in fy).$$

In addition, $\mathcal{U}(\mathsf{NFA})$ contains the substitution rule (Subst), i.e. the rule $A[\bar{P}] \;\Rightarrow\; A[\bar{B}/\bar{P}]$, where now $\bar{B}$ denote arbitrary formulas in the language of $\mathcal{U}(\mathsf{NFA})$, but $A[\bar{P}]$ is required to be a formula in the language of $\mathcal{U}_0(\mathsf{NFA})$. This last restriction is due to the fact that predicates in general depend on the predicate parameters $\bar{P}$. Finally, we obtain an intermediate predicate unfolding system $\mathcal{U}_1(\mathsf{NFA})$ by omitting axiom (16), i.e., $\mathcal{U}_1(\mathsf{NFA})$ is just $\mathcal{U}(\mathsf{NFA})$ without the *Join* predicate.

To state the proof-theoretic strength of the three unfolding systems $\mathcal{U}_0(\mathsf{NFA})$, $\mathcal{U}_1(\mathsf{NFA})$, and $\mathcal{U}(\mathsf{NFA})$, as usual we let $\mathsf{RA}_{<\alpha}$ denote the system of ramified analysis in levels less than α. In addition, Γ_0 is the so-called Feferman-Schütte ordinal, which was identified in the early sixties as the limiting number of predicative provability. As in Feferman and Strahm [20] we obtain the following proof-theoretic equivalences. In particular, the full unfolding of non-finitist arithmetic is equivalent to predicative analysis.

Theorem 1 *We have the following proof-theoretic equivalences:*

1. $\mathcal{U}_0(\mathsf{NFA}) \;\equiv\; \mathsf{PA}$.
2. $\mathcal{U}_1(\mathsf{NFA}) \;\equiv\; \mathsf{RA}_{<\omega}$.
3. $\mathcal{U}(\mathsf{NFA}) \;\equiv\; \mathsf{RA}_{<\Gamma_0}$.

In each case we have conservation with respect to arithmetic statements of the system on the left over the system on the right.

Let us give a few indications with respect to the proofs of these equivalences. In order to show that $\mathcal{U}_0(\mathsf{NFA})$ contains PA, one first shows by using the canonical fixed point operator of the underlying partial combinatory algebra that each primitive recursive function F can be represented by a term t_F in the language of $\mathcal{U}_0(\mathsf{NFA})$. Then one needs to show that these terms are well-typed on the natural numbers N, namely that $t_F(\bar{x}) \in \mathsf{N}$ for each $(\bar{x}) \in \mathsf{N}$: here one uses induction which follows by one application of the substitution rule to axiom (3) of NFA. A further application of (Subst) thus shows that the usual formulation of PA is directly contained in the operational unfolding $\mathcal{U}_0(\mathsf{NFA})$ of NFA. Indeed, $\mathcal{U}_0(\mathsf{NFA})$ does not go beyond PA, as is seen be formalizing its standard recursion-theoretic model in PA, see e.g. [20].

The full unfolding of NFA, $\mathcal{U}(\mathsf{NFA})$, derives the schema of transfinite induction along each initial segment of the Feferman-Schuütte ordinal Γ_0: Whenever we can derive in $\mathcal{U}(\mathsf{NFA})$ transfinite induction $\mathsf{TI}(\prec, P)$ along a primitive recursive ordering $\prec$, then we may substitute for P any formula and thus derive the existence of the predicate corresponding to the hyperarithmetic hierarchy along $\prec$, relative to any initial predicate p. Thus, using standard arguments from predicative wellordering proofs, whenever $\mathcal{U}(\mathsf{NFA})$ derives transfinite induction up to α, it also does so up to $\varphi\alpha 0$, hence the lower bound Γ_0. This bound is sharp according to Feferman and Strahm [20], see also Strahm [33].

Recall that in the intermediate unfolding $\mathcal{U}_1(\mathsf{NFA})$, the join principle is not available. We can still justify finite levels of the ramified analytical hierarchy, corresponding to the proof-theoretic ordinal $\varphi 20$, which is also the ordinal of the subsystem of second order arithmetic based on arithmetic comprehension and the bar rule, see Rathjen [30]. Indeed, in $\mathcal{U}_1(\mathsf{NFA})$, each application of the substitution rule lets us step from α to ε_α.

Let us close this section by mentioning that the original formulation of unfolding in Feferman [15] made use of a background theory of typed operations with general Least Fixed Point operator. The present formulation is a simplification of this approach. The upper bound computation in Feferman and Strahm [20] was done for this original formulation; it is worth mentioning that the proof-theoretic analysis of the unfolding of NFA in the present formulation is somewhat simpler and more elegant, since leastness for the fixed point operator is not present. A further difference is that predicates in the original formulation of unfolding were modeled as propositional functions using a truth predicate.

3 The unfolding of finitist arithmetic

In this section we describe the unfolding of two schematic systems for finitist arithmetic, namely FA and FA plus a form of the Bar rule BR. The main results are that all three unfolding systems for FA are equivalent to Primitive Recursive Arithmetic PRA, while the three unfoldings of $\mathsf{FA} + \mathsf{BR}$ reach precisely the strength of Peano arithmetic PA. These two characterizations of finitism are in accord with two prominent views about the limits of finitist reasoning due to Tait [35] and Kreisel [27]. In the sequel we follow Feferman and Strahm [21].

The logical operations of the basic schematic system FA of finitist arithmetic are restricted to $\wedge$, $\vee$, and $\exists$. In order to reason from such statements to new such statements given the above restriction of the logical operations of FA, we make use of a sequent formulation of our calculus, i.e. the statements proved are sequents Σ of the form $\Gamma \to A$, where Γ is a finite sequence (possibly empty) of formulas, and A may also be the false proposition $\bot$. Moreover, induction must now be given as a rule of inference involving such sequents. Accordingly, the basic axioms and rules of FA are as follows:

(1) $x' = 0 \to \bot$

(2) $\mathsf{Pd}(x') = x$

(3) $\dfrac{\Gamma \to P(0) \qquad \Gamma,\, P(x) \to P(x')}{\Gamma \to P(x)}$.

The substitution rule (Subst) may now generalized to incorporate sequent inference rules; the corresponding (meta) rule is called (Subst′) and will be spelled out in detail below.

Let us begin by describing the *operational unfolding* $\mathcal{U}_0(\mathsf{FA})$ of finitist arithmetic FA. That system tells us which operations from and to natural numbers, and which principles concerning them, ought to be accepted if we have accepted FA. It is seen that Skolem's system PRA of primitive recursive arithmetic is contained in $\mathcal{U}_0(\mathsf{FA})$. Indeed, the operational and even the full unfolding of finitist arithmetic do not go beyond PRA.

Large parts of the unfolding systems for FA and NFA are identical. Therefore, we shall confine ourselves in the sequel to mentioning the main differences in the specification of the unfolding systems for FA, beginning with its operational unfolding.

The *terms* of $\mathcal{U}_0(\mathsf{FA})$ are the same as the terms of $\mathcal{U}_0(\mathsf{NFA})$. Recall that FA is based on the logical operations $\wedge$, $\vee$, and $\exists$. Accordingly, the *formulas* of $\mathcal{U}_0(\mathsf{FA})$ are generated from the atomic formulas $\bot$, $s\downarrow$, $(s = t)$, $\mathsf{N}(s)$, and $P(\bar{s})$ by means of $\wedge$, $\vee$, and $\exists$; here P denotes an arbitrary free predicate variable of appropriate arity.

The underlying calculus of $\mathcal{U}_0(\mathsf{FA})$ is a Gentzen-type sequent system based on sequents of the form $\Gamma \to A$ for Γ being a finite sequence of formulas in the language of $\mathcal{U}_0(\mathsf{FA})$. In case Γ is empty, we shall write A for $\to A$. The logical axioms and rules of inference are the standard ones: apart from identity axioms, rules for $\bot$, cut and structural rules, these include the usual Gentzen-type rules for $\wedge$ and $\vee$ as well as introduction of $\exists$ on the left and on the right in the form

$$\frac{\Gamma \to A[t] \wedge t\downarrow}{\Gamma \to (\exists x)A[x]}, \qquad \frac{\Gamma,\, A[u] \to B}{\Gamma,\, (\exists x)A[x] \to B} \quad (u \text{ fresh})$$

Note that quantifiers range over defined objects only. Moreover, defined terms can be substituted for free variables according to the following rule of inference; here $\Gamma[t]$ stands for the sequence $(B[t] : B[u] \in \Gamma)$.

$$\frac{\Gamma[u] \;\rightarrow\; A[u]}{\Gamma[t],\, t{\downarrow} \;\rightarrow\; A[t]}$$

Finally, the equality and strictness axioms of our underlying logic of partial terms are given a Gentzen-style formulation in the obvious way.

The non-logical axioms and rules of $\mathcal{U}_0(\mathsf{FA})$ include the relativization of the axioms and rules of FA to the predicate N in the expected manner, as well as suitable formulations of the axioms (2)–(8) of $\mathcal{U}_0(\mathsf{NFA})$. We shall not spell out these axioms again, but instead give an example how to reformulate axiom (4) about the s combinator in our new setting. This now breaks into the following two axioms,

$$\mathsf{s}ab{\downarrow} \quad \text{and} \quad \mathsf{s}abc{\downarrow} \vee ac(bc){\downarrow} \;\rightarrow\; \mathsf{s}abc = ac(bc).$$

What is still missing in $\mathcal{U}_0(\mathsf{FA})$ is a suitable version of the *substitution rule* (Subst), which is central to all unfolding systems. In order to fit this into our Gentzen-style setting, (Subst) has to be formulated in a somewhat more general form. For that purpose, we let $\Sigma, \Sigma_1, \Sigma_2, \ldots$ range over sequents in the language of $\mathcal{U}_0(\mathsf{FA})$. A *rule of inference* for such sequents has the general form

$$\frac{\Sigma_1, \Sigma_2, \ldots, \Sigma_n}{\Sigma},$$

which we simply abbreviate by $\Sigma_1, \Sigma_2, \ldots, \Sigma_n \Rightarrow \Sigma$ in the sequel; we also allow n to be 0, i.e. rules with an empty list of premises are possible. As usual we call a rule of inference $\Sigma_1, \Sigma_2, \ldots, \Sigma_n \Rightarrow \Sigma$ *derivable* from a collection of axioms and rules $\mathcal{T}$ (all in Gentzen-style), if the sequent Σ is derivable from $\mathcal{T} \cup \{\Sigma_1, \Sigma_2, \ldots, \Sigma_n\}$.

In the following $\bar{P} = P_1, \ldots, P_m$ denotes a finite sequence of free predicate symbols of finite arity and $\bar{B} = B_1, \ldots, B_m$ a corresponding sequence of formulas in the language of $\mathcal{U}_0(\mathsf{FA})$. If $\Sigma[\bar{P}]$ is a sequent possibly containing the free predicates $\bar{P}$, then as above $\Sigma[\bar{B}/\bar{P}]$ denotes the sequent $\Sigma[\bar{P}]$ with each subformula of the form $P_i(\bar{t})$ replaced by $(\exists \bar{x})(\bar{t} = \bar{x} \wedge B[\bar{x}])$, where the length of $\bar{x}$ is equal to the arity of P_i.

We are now ready to state our (meta) substitution rule (Subst'). Its meaning is as follows: whenever the axioms and rules of inference at hand allow us to show that the rule $\Sigma_1, \Sigma_2, \ldots, \Sigma_n \Rightarrow \Sigma$ is *derivable*, then we can adjoin each of its substitution instances $\Sigma_1[\bar{B}/\bar{P}], \Sigma_2[\bar{B}/\bar{P}], \ldots, \Sigma_n[\bar{B}/\bar{P}] \Rightarrow \Sigma[\bar{B}/\bar{P}]$ as a new rule of inference to $\mathcal{U}_0(\mathsf{FA})$, for $B_i[\bar{x}]$ being formulas in the language of $\mathcal{U}_0(\mathsf{FA})$.[4] Symbolically,

$$(\mathsf{Subst}') \qquad \frac{\Sigma_1, \Sigma_2, \ldots, \Sigma_n \Rightarrow \Sigma}{\Sigma_1[\bar{B}/\bar{P}],\, \Sigma_2[\bar{B}/\bar{P}], \ldots, \Sigma_n[\bar{B}/\bar{P}] \;\Rightarrow\; \Sigma[\bar{B}/\bar{P}]}$$

It is not difficult to see that the primitive recursive functions can be introduced and proved total in $\mathcal{U}_0(\mathsf{FA})$. Indeed, the argument described in the previous section is readily seen to be formalizable in $\mathcal{U}_0(\mathsf{FA})$, since induction on N is available for equations and formulas of the form $t(\bar{x}) \in \mathsf{N}$. Thus, Primitive Recursive Arithmetic is interpretable in $\mathcal{U}_0(\mathsf{FA})$, see [21] for details.

The *full unfolding* $\mathcal{U}(\mathsf{FA})$ of finitist arithmetic FA is an extension of the operational unfolding $\mathcal{U}_0(\mathsf{FA})$ and is used, in addition, to answer the question of which operations on and to predicates, and which principles concerning them, are to be accepted if one has accepted FA. We shall see that $\mathcal{U}(\mathsf{FA})$ does not go beyond primitive recursive arithmetic PRA in proof-theoretic strength.

The language of $\mathcal{U}(\mathsf{FA})$ is an extension of the language of $\mathcal{U}_0(\mathsf{FA})$. It includes, in addition, the constants nat (natural numbers), eq (equality), pr_P (free predicate P), inv (inverse image), conj (conjunction), disj (disjunction), $(\exists)$ (existential quantification), and join (disjoint unions). Moreover, as above, we have a new unary relation symbol Π for (codes of) predicates and a binary relation symbol $\in$ for the elementhood relation. The *terms* of $\mathcal{U}(\mathsf{FA})$ are built as before. The *formulas* of $\mathcal{U}(\mathsf{FA})$ extend the formulas of $\mathcal{U}_0(\mathsf{FA})$ by allowing the new atomic formulas $\Pi(t)$ and $s \in t$.

The axioms of $\mathcal{U}(\mathsf{FA})$ extend those of $\mathcal{U}_0(\mathsf{FA})$. In addition, we have the obvious defining axioms for the basic predicates of $\mathcal{U}(\mathsf{FA})$. These include straightforward reformulations using sequents of the axioms (9)–(12) and (14) of $\mathcal{U}(\mathsf{NFA})$ as well as the expected axiom about existentially quantified

[4] Observe that derivability of rules is a dynamic process as we unfold FA. In particular, new rules of inference obtained by (Subst') allow us to establish new derivable rules, to which in turn we can apply (Subst'). In particular, the usual rule of induction

$$\frac{\Gamma \,\rightarrow\, A[0] \quad \Gamma,\, u \in \mathsf{N},\, A[u] \,\rightarrow\, A[u']}{\Gamma,\, v \in \mathsf{N} \,\rightarrow\, A[v]}$$

is an immediate consequence of (Subst') applied to rule (3) of FA. Moreover, the substitution rule in its usual form as stated in Section 2,

$$(\mathsf{Subst}) \qquad \frac{\Sigma[\bar{P}]}{\Sigma[\bar{B}/\bar{P}]}$$

is readily seen to be an admissible rule of inference of $\mathcal{U}_0(\mathsf{FA})$.

predicates; see Feferman and Strahm [21]. Further, axiom (16) of $\mathcal{U}(\mathsf{NFA})$ concerning join is now stated in terms of suitable inference rules; this is due to the absence of universal quantification in the framework of finitist arithmetic, see [21] for details.

Finally, $\mathcal{U}(\mathsf{FA})$ of course also includes the substitution rule (Subst') which we have spelled out for $\mathcal{U}_0(\mathsf{FA})$. The formulas $\bar{B}$ to be substituted for $\bar{P}$ are now in the language of $\mathcal{U}(\mathsf{FA})$; the rule in the premise of (Subst'), however, is required to be in the language of $\mathcal{U}_0(\mathsf{FA})$. This last restriction is imposed as before since predicates may depend on the free relation symbols $\bar{P}$. The intermediate unfolding system $\mathcal{U}_1(\mathsf{FA})$ for FA is obtained by dropping the rules about join.

It is shown in Feferman and Strahm [21] that all three unfolding systems for FA do not go beyond PRA in strength. This is obtained via a suitable recursion-theoretic interpretation into the subsystem Σ_1-IA of PA with induction on the natural numbers restricted to Σ_1 formulas; the latter system is known to be a Π_2 conservative extension of PRA by the well-known Mints-Parsons-Takeuti theorem, see Sieg [32] for a simple proof. The embedding essentially models the applicative axioms by means of partial recursive function application and the predicates by Σ_1 definable properties, where some special attention is required in order to validate the generalized substitution rule (Subst'). Thus we can summarize:

Theorem 2 $\mathcal{U}_0(\mathsf{FA})$*,* $\mathcal{U}_1(\mathsf{FA})$ *and* $\mathcal{U}(\mathsf{FA})$ *are all proof-theoretically equivalent to primitive recursive arithmetic* PRA*.*

In the remainder of this section we shall discuss an extension of FA by a Bar Rule BR and, correspondingly, three unfolding systems of $\mathsf{FA} + \mathsf{BR}$, all of strength Peano arithmetic.

Informally speaking, the Bar Rule BR says that if $\prec$ is a partial ordering provably satisfying $\mathsf{NDS}(\prec)$ (no infinite descending sequence property for $\prec$) then the principle $\mathsf{TI}(\prec, P)$ of transfinite induction on $\prec$ holds for arbitrary predicates P. It is sufficient to restrict this to provably decidable linear orderings $\prec$ in the natural numbers, with 0 as least element. But further restrictions have to be made in order to fit a version of BR to the language of FA. First of all, the statement that a given function f on N is descending in the $\prec$ relation, as long as it is not 0, is universal, so cannot be expressed as a formula of our language. Instead, we add a new function constant symbol f interpreted as an arbitrary (or "anonymous") function, and require that we establish a rule, $\mathsf{NDS}(\prec, \mathsf{f})$, that allows us to infer from the hypotheses that $\mathsf{f} : \mathsf{N} \to \mathsf{N}$ and that $\mathsf{f}(u') \prec \mathsf{f}(u)$ as long as $\mathsf{f}(u) \neq 0$ ('u' a free variable) the conclusion $(\exists x \in \mathsf{N})(\mathsf{f}(x) = 0)$. In addition, we must modify $\mathsf{TI}(\prec, P)$, since its standard formulation for a unary predicate P is of the form:

$$(\forall x)[(\forall u \prec x)P(u) \to P(x)] \to (\forall x)P(x).$$

Again, the idea is to treat this as a rule of the form:

$$\text{from} \quad (\forall u)[u \prec x \to P(u)] \to P(x) \quad \text{infer} \quad P(x).$$

But we still need an additional step to reformulate the hypothesis of this rule in the language of FA. For atomic A, B write $A \supset B$ for $(\neg A \vee B)$. Then the hypothesis is implied by

$$[t_1 \prec x \supset P(t_1)] \wedge \cdots \wedge [t_m \prec x \supset P(t_m)] \to P(x),$$

where the t_i are terms that have been proved to be defined. Now it may be that we cannot prove that $t_i\!\downarrow$ until we know that certain of its subterms $s_1, \ldots, s_n$ are defined and satisfy

$$[s_1 \prec x \supset P(s_1)] \wedge \cdots \wedge [(s_n \prec x \supset P(s_n)],$$

and so on. Indeed, as we shall see, that is necessary to establish closure under nested recursion on the $\prec$ ordering. This leads to the precise statement of BR in the language of FA augmented by a new function symbol f as follows.[5]

The rule $\mathsf{NDS}(\mathsf{f}, \prec)$ says that for each possibly infinite descending chain f w.r.t. $\prec$ there is an x such that $\mathsf{f}x = 0$. Formally, it is given as follows:

$$\frac{\begin{array}{c} u \in \mathsf{N} \;\to\; \mathsf{f}u \in \mathsf{N}, \\ u \in \mathsf{N},\, \mathsf{f}u \neq 0 \;\to\; \mathsf{f}(u') \prec \mathsf{f}u, \\ u \in \mathsf{N},\, \mathsf{f}u = 0 \;\to\; \mathsf{f}(u') = 0 \end{array}}{(\exists x \in \mathsf{N})(\mathsf{f}x = 0)}$$

Next, the bar rule BR is spelled out in detail for the case of nesting level two and a predicate with one parameter. The general case for nesting of arbitrary level and number of parameters is analogous.

Let $\bar{s^r} = s^r_1, \ldots, s^r_n$ and $\bar{s^p} = s^p_1, \ldots, s^p_n$ be sequences of terms of length n, and let $\bar{t^r} = t^r_1, \ldots, t^r_m$ and $\bar{t^p} = t^p_1, \ldots, t^p_m$ be sequences of terms of length m. The superscripts 'r' and 'p' stand for recursion and parameter, respectively.

The bar rule BR now reads as follows. Whenever we have derived the four premises

[5] In the formulation of the rules below we use a binary relation $\prec$ whose characteristic function is given by a closed term $t_\prec$ for which $\mathcal{U}_0(\mathsf{FA})$ proves $t_\prec : \mathsf{N}^2 \to \{0, 1\}$. We write $x \prec y$ instead of $t_\prec xy = 0$ and further assume that $\prec$ is a linear ordering with least element 0, provably in $\mathcal{U}_0(\mathsf{FA})$.

(1) $\mathsf{NDS}(\mathsf{f},\prec)$

(2) $x, y \in \mathsf{N} \;\rightarrow\; \bar{s^r} \in \mathsf{N} \wedge \bar{s^p} \in \mathsf{N}$

(3) $x, y \in \mathsf{N},\; \bigwedge_i [s_i^r \prec x \supset P(s_i^r, s_i^p)] \;\rightarrow\; \bar{t^r} \in \mathsf{N} \wedge \bar{t^p} \in \mathsf{N}$

(4) $x, y \in \mathsf{N},\; \bigwedge_i [s_i^r \prec x \supset P(s_i^r, s_i^p)],\; \bigwedge_j [t_j^r \prec x \supset P(t_j^r, t_j^p)] \;\rightarrow\; P(x, y)$

we can infer $x, y \in \mathsf{N} \;\rightarrow\; P(x, y)$.[6]

The new unfolding system $\mathcal{U}_0(\mathsf{FA} + \mathsf{BR})$ is the extension of $\mathcal{U}_0(\mathsf{FA})$ by this rule.

One of the crucial observations is that whenever we have derived $\mathsf{NDS}(\mathsf{f},\prec)$ in $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$, for a specific ordering $\prec$, then we can use the bar rule BR in order to justify function definitions by *nested* recursion along $\prec$, see Feferman and Strahm [21] for details.

Theorem 3 *Assume that* $\mathsf{NDS}(\mathsf{f},\prec)$ *is derivable in* $\mathcal{U}_0(\mathsf{FA} + \mathsf{BR})$*. Then* $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$ *justifies nested recursion along* $\prec$*.*

In the following let us assume that for each ordinal $\alpha < \varepsilon_0$ we have a standard primitive recursive wellordering $\prec_\alpha$ of ordertype α. Further, let us write $\mathsf{NDS}(\mathsf{f},\alpha)$ for $\mathsf{NDS}(\mathsf{f},\prec_\alpha)$. The crucial ingredient of the argument to show that $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$ derives $\mathsf{NDS}(\mathsf{f},\alpha)$ for each $\alpha < \varepsilon_0$ is the famous result by Tait [34] that nested recursion on $\omega\alpha$ entails ordinary recursion on ω^α or, more useful in our setting, nested recursion on $\omega\alpha$ entails $\mathsf{NDS}(\mathsf{f},\omega^\alpha)$.

Theorem 4 *Provably in* $\mathcal{U}_0(\mathsf{FA} + \mathsf{BR})$*, nested recursion along* $\omega\alpha$ *entails* $\mathsf{NDS}(\mathsf{f},\omega^\alpha)$*.*

Clearly, $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$ proves $\mathsf{NDS}(\mathsf{f},\omega 2)$ and hence we have nested recursion along $\omega 2$, which in turn entails $\mathsf{NDS}(\mathsf{f},\omega^2)$; further, nested recursion on ω^2 gives us $\mathsf{NDS}(\mathsf{f},\omega^\omega)$ and thus nested recursion along $\omega^\omega = \omega(\omega^\omega)$. Then we can derive $\mathsf{NDS}(\mathsf{f},\omega^{\omega^\omega})$ and so on.

The upshot is that $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$ derives $\mathsf{NDS}(\mathsf{f},\omega_n)$ for each natural number n, where as usual we set $\omega_0 = \omega$ and $\omega_{n+1} = \omega^{\omega_n}$.

Corollary 5 *We have for each* $\alpha < \varepsilon_0$ *that* $\mathcal{U}_0(\mathsf{FA}+\mathsf{BR})$ *derives* $\mathsf{NDS}(\mathsf{f},\alpha)$*.*

It is not difficult to see that this lower bound is sharp, see Feferman and Strahm [21].

[6] In the formulation of this rule, we have used the shorthand $r \prec x \supset A$ for the formula $t_\prec rx = 1 \vee A$.

Corollary 6 $\mathcal{U}_0(\mathsf{FA} + \mathsf{BR})$ *is proof-theoretically equivalent to Peano arithmetic* PA*.*

Even the full unfolding system with bar rule, $\mathcal{U}(\mathsf{FA}+\mathsf{BR})$, does not go beyond Peano arithmetic in strength.

Theorem 7 $\mathcal{U}_0(\mathsf{FA} + \mathsf{BR})$*,* $\mathcal{U}_1(\mathsf{FA} + \mathsf{BR})$*, and* $\mathcal{U}(\mathsf{FA} + \mathsf{BR})$ *are all proof-theoretically equivalent to Peano arithmetic* PA*.*

4 The unfolding of feasible arithmetic

The aim of this section is to discuss the concept of unfolding in the context of a natural schematic system FEA for *feasible arithmetic*. We shall sketch various unfoldings of FEA and indicate their relationship to weak systems of explicit mathematics and partial truth. We follow Eberhard and Strahm [9].

Let us first introduce the basic schematic system FEA of feasible arithmetic. Its intended universe of discourse is the set $\mathbb{W} = \{0, 1\}^*$ of finite binary words and its basic operations and relations include the binary successors S_0 and S_1, the predecessor Pd, the initial subword relation $\subseteq$, word concatenation $\circledast$ as well as word multiplication $\boxtimes$.[7] The logical operations of FEA are conjunction ($\wedge$), disjunction ($\vee$), and bounded existential quantification ($\exists^{\leq}$). As in the case of finitist arithmetic FA, the statements proved in FEA are sequents of formulas in the given language, i.e. implication is allowed at the outermost level.

The language of FEA contains a countably infinite supply $\alpha, \beta, \gamma, \ldots$ of variables (possibly with subscripts). These variables are interpreted as ranging over the set of binary words $\mathbb{W}$. We have a constant ϵ for the empty word, three unary function symbols $\mathsf{S}_0, \mathsf{S}_1, \mathsf{Pd}$ and three binary function symbols $\circledast$, $\boxtimes$, $\subseteq$.[8] Terms are defined as usual and are denoted by $\sigma, \tau, \ldots$. Further, there is the binary predicate symbol $=$ for equality, and an infinite supply $P, Q, \ldots$ of free predicate letters.

The atomic formulas of FEA are of the form $(\sigma = \tau)$ and $P(\sigma_1, \ldots, \sigma_n)$. The formulas are closed under $\wedge$ and $\vee$ as well as under bounded existential quantification. In particular, if A is formula, then $(\exists \alpha \leq \tau)A$ is formula as well, where τ is not allowed to contain α. Further, as usual for theories of words, we use $\sigma \leq \tau$ as an abbreviation for $1 \boxtimes \sigma \subseteq 1 \boxtimes \tau$, thus expressing

[7] Given two words w_1 and w_2, the word $w_1 \boxtimes w_2$ denotes the length of w_2 fold concatenation of w_1 with itself.

[8] We assume that $\subseteq$ defines the characteristic function of the initial subword relation. Further, we employ infix notation for these binary function symbols.

that the length of σ is less than or equal to the length of τ. As before, we use $\bar{\alpha}, \bar{\sigma}$, and $\bar{A}$ to denote finite sequences of variables, terms, and formulas, respectively.

FEA is formulated as a system of sequents Σ of the form $\Gamma \rightarrow A$, where Γ is a finite sequence of formulas and A is a formula. Hence, we have the usual Gentzen-type logical axioms and rules of inference for our underlying restricted language, see Eberhard and Strahm [9]. The non-logical axioms of FEA state the usual defining equations for the function symbols of the language $\mathcal{L}$, see, e.g., Ferreria [22] for similar axioms. Finally, we have the schematic induction rule formulated for a free predicate P as follows:

$$\frac{\Gamma \rightarrow P(\epsilon) \qquad \Gamma, P(\alpha) \rightarrow P(\mathsf{S}_i(\alpha)) \quad (i = 0, 1)}{\Gamma \rightarrow P(\alpha)}$$

In the various unfolding systems of FEA introduced below, we shall be able to substitute an arbitrary formula for an arbitrary free predicate letter P. Let us now quickly review the *operational unfolding* $\mathcal{U}_0(\mathsf{FEA})$ of FEA. It tells us which operations from and to individuals, and which principles concerning them, ought to be accepted if one has accepted FEA.

In the operational unfolding, we make these commitments explicit by extending FEA by a partial combinatory algebra. Since it represents any recursion principle and thus any recursive function by suitable terms, it is expressive enough to reflect any ontological commitment we want to reason about. Using the notion of *provable totality*, we single out those functions and recursion principles we are actually committed to by accepting FEA.

The language of $\mathcal{U}_0(\mathsf{FEA})$ is an expansion of the language of FEA including new constants k, s, p, $\mathsf{p_0}$, $\mathsf{p_1}$, d, tt, ff, e, ϵ, $\mathsf{s_0}$, $\mathsf{s_1}$, pd, $\mathsf{c}_\subseteq$, $*$, $\times$, and an additional countably infinite set of variables $x_0, x_1, \ldots$.[9] The new variables are supposed to range over the universe of operations and are usually denoted by $a, b, c, x, y, z, \ldots$. The terms $(r, s, t, \ldots)$ are inductively generated from variables and constants by means of the function symbols of FEA and the application operator $\cdot$. We use the usual abbreviations for applicative terms as before. We have $(s = t)$, $s{\downarrow}$ and $P(\bar{s})$ as atoms. The formulas $(A, B, C, \ldots)$ are built from the atoms as before using $\vee, \wedge$ and the bounded existential quantifier, where as above the bounding term is a term of FEA not containing the bound variable. Finally, we write $s \leq \tau$ for $(\exists \beta \leq \tau)(s = \beta)$.[10]

[9] These variables are syntactically different from the FEA variables $\alpha_0, \alpha_1, \ldots$.

[10] It is important to note that we do not have a predicate W for binary words in our language, since this would allow us to introduce (hidden) unbounded existential quantifiers via formulas of the form $\mathsf{W}(t)$. Thus it is necessary to have two separate sets of variables for words and operations, respectively.

The axioms and rules of $\mathcal{U}_0(\mathsf{FEA})$ are now spelled out in the expected manner, see Eberhard and Strahm [9] for details. In particular, $\mathcal{U}_0(\mathsf{FEA})$ includes the (meta) substitution rule (Subst'). Next we want to show that the polynomial time computable functions can be proved to be total in $\mathcal{U}_0(\mathsf{FEA})$. We call a function $F : \mathbb{W}^n \to \mathbb{W}$ provably total in a given axiomatic system, if there exists a closed term t_F such that (i) t_F defines F pointwise, i.e. on each standard word, and, moreover, (ii) there is a FEA term $\tau[\alpha_1, \dots, \alpha_n]$ such that the assertion

$$t_F(\alpha_1, \dots, \alpha_n) \leq \tau[\alpha_1, \dots, \alpha_n]$$

is provable in the underlying system. Thus, in a nutshell, F is provably total iff it is provably and uniformly bounded.

We use Cobham's characterization of the polynomial time computable functions (cf. [5, 4]): starting off from the initial functions of FEA and arbitrary projections, the polynomial time computable functions can be generated by closing under composition and bounded recursion. In order to show closure under bounded recursion, assume that F is defined by bounded recursion with initial function G and step function H_i ($i = 0, 1$), where τ is the corresponding bounding polynomial.[11] By the induction hypothesis, G and H_i are provably total via suitable terms t_G and t_{H_i}. Using the recursion or fixed point theorem of the partial combinatory algebra, we find a term t_F which provably in $\mathcal{U}_0(\mathsf{FEA})$ satisfies the following recursion equations for $i = 0, 1$:

$$\begin{aligned} t_F(\bar{\alpha}, \epsilon) &\simeq t_G(\bar{\alpha}) \mid \tau[\bar{\alpha}, \epsilon], \\ t_F(\bar{\alpha}, \mathsf{s}_i(\beta)) &\simeq t_{H_i}(t_F(\bar{\alpha}, \beta), \bar{\alpha}, \beta) \mid \tau[\bar{\alpha}, \mathsf{s}_i(\beta)] \end{aligned}$$

Here $\mid$ is the usual truncation operation such that $\alpha|\beta$ is α if $\alpha \leq \beta$ and β otherwise. Now fix $\bar{\alpha}$ and let $A[\beta]$ be the formula $t_F(\bar{\alpha}, \beta) \leq \tau[\bar{\alpha}, \beta]$ [12] and simply show $A[\beta]$ by induction on β. Thus F is provably total in $\mathcal{U}_0(\mathsf{FEA})$.

Next we shall describe the full predicate unfolding $\mathcal{U}(\mathsf{FEA})$ of FEA. It tells us, in addition, which predicates and operations on predicates ought to be accepted if one has accepted FEA. By accepting FEA one implicitly accepts an equality predicate and operations on predicates corresponding to the logical operations of FEA. Finally, we may accept the principle of forming the predicate for the disjoint union of a (bounded) sequence of predicates given by an operation.

[11] We can assume that only functions built from concatenation and multiplication are permissible bounds for the recursion.

[12] Recall that by expanding the definition of the $\leq$ relation, the formula $A[\beta]$ stands for the assertion $(\exists \gamma \leq \tau[\bar{\alpha}, \beta])(t_F(\bar{\alpha}, \beta) = \gamma)$.

The language of $\mathcal{U}(\mathsf{FEA})$ is an extension of the language of $\mathcal{U}_0(\mathsf{FEA})$ by new individual constants id (identity), inv (inverse image), con (conjunction), dis (disjunction), leq (bounded existential quantifier), and j (bounded disjoint unions); further new constants are of the form pr_P as combinatorial representations of free predicates. Finally, we have a new unary relation symbol Π in order to single out the predicates we are committed to as well as a binary relation symbol $\in$ for elementhood of individuals in predicates. The terms are generated as before but now taking into account the new constants. The formulas of $\mathcal{U}(\mathsf{FEA})$ extend the formulas of $\mathcal{U}_0(\mathsf{FEA})$ by allowing new atomic formulas of the form $\Pi(t)$ and $s \in t$.

The axioms of $\mathcal{U}(\mathsf{FEA})$ extend those of $\mathcal{U}_0(\mathsf{FEA})$ by the expected axioms about predicates, see Eberhard and Strahm [9] for details. Further, the full unfolding $\mathcal{U}(\mathsf{FEA})$ includes axioms stating that a bounded sequence of predicates determines the predicate of the disjoint union of this sequence. We use a set of three inference rules to express the join principle, see [9] for details. The rules of inference of $\mathcal{U}_0(\mathsf{FEA})$ are also available in $\mathcal{U}(\mathsf{FEA})$. In particular, $\mathcal{U}(\mathsf{FEA})$ contains the generalized substitution rule (Subst'): the formulas $\bar{B}$ to be substituted for $\bar{P}$ are now in the language of $\mathcal{U}(\mathsf{FEA})$; as above the rule in the premise of (Subst'), however, is required to be in the language $\mathcal{U}_0(\mathsf{FEA})$. This concludes the description of the predicate unfolding $\mathcal{U}(\mathsf{FEA})$ of FEA. We shall turn to its proof-theoretic strength at the end of this section.

Let us next discuss an alternative way to define the full unfolding of FEA. The truth unfolding $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$ of FEA makes use of a truth predicate T which reflects the logical operations of FEA in a natural and direct way. We shall see that the full predicate unfolding $\mathcal{U}(\mathsf{FEA})$ is directly contained in $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$.[13]

As before, we want to make the commitment to the logical operations of FEA explicit. This is done by introducing a truth predicate for which truth biconditionals defining the truth conditions of the logical operations hold. The axiomatization of the truth predicate relies on a coding mechanism for formulas. In the applicative framework, this is achieved in a very natural way by using new constants designating the logical operations of FEA. The language of $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$ extends the one of $\mathcal{U}_0(\mathsf{FEA})$ by new individual constants $\dot{=}$, $\dot{\wedge}$, $\dot{\vee}$, $\dot{\exists}$, as well as constants of the form pr_P. In addition, It includes a new unary relation symbol T. The terms and formulas are defined in the expected manner. Moreover, we shall use infix notation for $\dot{=}$, $\dot{\wedge}$ and $\dot{\vee}$.

The axioms of $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$ extend those of $\mathcal{U}_0(\mathsf{FEA})$ by the following axioms

[13]Recall that in Feferman's original definition of unfolding in [15], a truth predicate is used in order to describe the full unfolding of a schematic system.

about the truth predicate T:

$(\dot{=})$	$\mathsf{T}(x \dot{=} y)$	$\leftrightarrow$	$x = y$
$(\dot{\wedge})$	$\mathsf{T}(x \dot{\wedge} y)$	$\leftrightarrow$	$\mathsf{T}(x) \wedge \mathsf{T}(y)$
$(\dot{\vee})$	$\mathsf{T}(x \dot{\vee} y)$	$\leftrightarrow$	$\mathsf{T}(x) \vee \mathsf{T}(y)$
$(\dot{\exists})$	$\mathsf{T}(\dot{\exists}\alpha x)$	$\leftrightarrow$	$(\exists \beta \leq \alpha)\mathsf{T}(x\beta)$
(pr_P)	$\mathsf{T}(\mathsf{pr}_P(\bar{x}))$	$\leftrightarrow$	$P(\bar{x})$

The generalized substitution rule (Subst') can be stated in a somewhat more general form for $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$, see Eberhard and Strahm [9] for a detailed discussion. It is easy to see that the full predicate unfolding $\mathcal{U}(\mathsf{FEA})$ is contained in the truth unfolding $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$. The argument proceeds along the same line as the embedding of weak explicit mathematics into theories of truth in Eberhard and Strahm [8].

In Eberhard and Strahm [9] it is shown how to determine a suitable upper bound for $\mathcal{U}(\mathsf{FEA})$ and $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$ thus showing that their provably total functions are indeed computable in polynomial time. There one proceeds via the weak truth theory T_{PT} introduced in Eberhard and Strahm [8] and Eberhard [6, 7], whose detailed and very involved proof-theoretic analysis is carried out in [7]. Thus we have:

Theorem 8 *The provably total functions of* $\mathcal{U}_0(\mathsf{FEA})$*,* $\mathcal{U}(\mathsf{FEA})$*, and* $\mathcal{U}_{\mathsf{T}}(\mathsf{FEA})$ *are exactly the polynomial time computable functions.*

5 The unfolding of one inductive definition

Our last example for illustrating the unfolding program stems from a natural schematic system for arithmetical inductive definitions. We shall see that its unfolding corresponds to a generalization $\Psi(\Gamma_{\Omega+1})$ of Γ_0, which we shall describe below. The main result of this section is due to Buchholtz [2].

Let ID_1 be the usual system for one inductive definitions. In order to formulate it in schematic form as an extension of NFA, we have a new predicate constant $P_{\mathcal{A}}$ for each arithmetical operator form $\mathcal{A}[P, x]$ in which P occurs only positively. Then we obtain a schematic version of ID_1 as follows, with P denoting a free predicate variable:

(1) $(\forall x)(\mathcal{A}[P_{\mathcal{A}}, x] \rightarrow P_{\mathcal{A}}(x))$

(2) $(\forall x)(\mathcal{A}[P, x] \rightarrow P(x)) \rightarrow (\forall x)(P_{\mathcal{A}}(x) \rightarrow P(x))$.

The full unfolding $\mathcal{U}(\mathsf{ID}_1)$ of ID_1 is now defined according to the procedure described in detail for the case of NFA, with the only exception that the join axiom is formulated in a somewhat more general form: the family of predicates to which the join operation is applied, is not restricted to be indexed by the natural numbers N but by an arbitrary predicate p.

In order to describe the result about the strength of $\mathcal{U}(\mathsf{ID}_1)$ obtained in Buchholtz's thesis [2], let us review some basic ordinal theory needed to calibrate the proof-theoretic ordinal of $\mathcal{U}(\mathsf{ID}_1)$. Let Ω stand for $\aleph_1$. Then the sets $B_\Omega(\alpha)$ and ordinals $\Psi_\Omega(\alpha)$ are defined recursively as follows: $B_\Omega(\alpha)$ is the closure of $\{0, \Omega\}$ under $+$, the binary Veblen function φ, and $(\xi \mapsto \Psi_\Omega(\xi))_{\xi<\alpha}$; moreover,

$$\Psi_\Omega(\alpha) \simeq \min\{\xi < \Omega : \xi \notin B(\alpha)\}.$$

It is seen that $\Psi_\Omega(\alpha)$ is always defined and, hence, denotes an ordinal smaller than Ω. Finally, let $\Gamma_{\Omega+1}$ be the least Γ number beyond Ω. In the following, we are interested in the ordinal $\Psi_\Omega(\Gamma_{\Omega+1})$, which we simply denote by $\Psi(\Gamma_{\Omega+1})$. The main result of Buchholtz' thesis is the following

Theorem 9 *The proof-theoretic ordinal of* $\mathcal{U}(\mathsf{ID}_1)$ *is* $\Psi(\Gamma_{\Omega+1})$.

More recently, a number of natural systems have been identified whose proof-theoretic ordinal is $\Psi(\Gamma_{\Omega+1})$, see Buchholtz, Jäger, and Strahm [3]. Basically, those systems arise from natural systems of second order arithmetic of strength Γ_0 by allowing one generalized inductive definition at the bottom level, resulting in analogues of Δ^1_1 comprehension, Σ^1_1 choice and dependent choice, always with substitution rule, as well as Friedman's arithmetical transfinite recursion.

6 Concluding remarks

One of the motivations in Feferman [15] for studying the unfolding of a schematic system S was to explicate some ideas that were initiated by Gödel regarding axioms for hierarchies of inaccessible and Mahlo cardinals. Gödel [24], p. 182 writes that "these axioms show clearly, not only that the axiomatic system of set theory as known today is incomplete, but also that it can be supplemented without arbitrariness by new axioms which are only the natural continuation of those set up so far".

Feferman [15] proposes a number of schematic systems for impredicative and admissible set theory. He further introduces a schematic reflection principle, so-called *Downward Reflection*, expressing that *whatever holds in the universe of sets already holds in arbitrary large transitive sets*. This principle entails a form of Bernays' downward second order reflection principle, from which the existence of hierarchies of Mahlo cardinals follows.

The unfolding systems for set theory mentioned above may also be directly expressed in the language of Feferman's operational set theory OST (cf. [16, 25]). We refer the reader to Feferman [18] for a number of interesting conjectures regarding various unfoldings of OST.

References

[1] Beeson, M. J. *Foundations of Constructive Mathematics: Metamathematical Studies.* Springer, Berlin, 1985.

[2] Buchholtz, U. *Unfolding of systems of inductive definitions.* PhD thesis, Stanford University, 2013.

[3] Buchholtz, U., Jäger, G., and Strahm, T. Theories of proof-theoretic strength $\Psi(\Gamma_{\Omega+1})$. In *Concepts of Proof in Mathematics, Philosophy, and Computer Science*, D. Probst and P. Schuster, Eds., vol. 6 of *Ontos Mathematical Logic*. De Gruyter, to appear.

[4] Clote, P. Computation models and function algebras. In *Handbook of Computability Theory*, E. Griffor, Ed. Elsevier, 1999, pp. 589–681.

[5] Cobham, A. The intrinsic computational difficulty of functions. In *Logic, Methodology and Philosophy of Science II.* North Holland, Amsterdam, 1965, pp. 24–30.

[6] Eberhard, S. *Weak Applicative Theories, Truth, and Computational Complexity.* PhD thesis, Universität Bern, 2013.

[7] Eberhard, S. A feasible theory of truth over combinatory algebra. *Annals of Pure and Applied Logic 165*, 5 (2014), 1009–1033.

[8] Eberhard, S., and Strahm, T. Weak theories of truth and explicit mathematics. In *Logic, Construction, Computation*, U. Berger, H. Diener, P. Schuster, and M. Seisenberger, Eds. Ontos Verlag, 2012, pp. 157–184.

[9] Eberhard, S., and Strahm, T. Unfolding feasible arithmetic and weak truth. In *Unifying the Philosophy of Truth*, D. Achourioti, H. Galinon, K. Fujimoto, and J. Martinez, Eds. Springer, 2015, pp. 153–167.

[10] Feferman, S. Transfinite recursive progressions of axiomatic theories. *Journal of Symbolic Logic 27* (1962), 259–316.

[11] Feferman, S. Systems of predicative analysis. *Journal of Symbolic Logic 29*, 1 (1964), 1–30.

[12] Feferman, S. A language and axioms for explicit mathematics. In *Algebra and Logic*, J. Crossley, Ed., vol. 450 of *Lecture Notes in Mathematics*. Springer, 1975, pp. 87–139.

[13] Feferman, S. A more perspicuous system for predicativity. In *Konstruktionen vs. Positionen I*. de Gruyter, Berlin, 1979, pp. 87–139.

[14] Feferman, S. Logics for termination and correctness of functional programs. In *Logic from Computer Science*, Y. N. Moschovakis, Ed., vol. 21 of *MSRI Publications*. Springer, Berlin, 1991, pp. 95–127.

[15] Feferman, S. Gödel's program for new axioms: Why, where, how and what? In *Gödel '96*, P. Hájek, Ed., vol. 6 of *Lecture Notes in Logic*. Springer, Berlin, 1996, pp. 3–22.

[16] Feferman, S. Operational set theory and small large cardinals. *Information and Computation 207* (2009), 971–979.

[17] Feferman, S. Turing's Thesis: Ordinal logics and oracle computability. In *Alan Turing: His Work and Impact*, S. B. Cooper and J. van Leeuwen, Eds. Elsevier, 2013, pp. 145–150.

[18] Feferman, S. The operational perspective: three routes. In *Advances in Proof Theory*, R. Kahle, T. Strahm, and T. Studer, Eds., Progress in Computer Science and Applied Logic. Birkhäuser, to appear.

[19] Feferman, S., and Spector, C. Incompleteness along paths in progressions of theories. *Journal of Symbolic Logic 27* (1962), 383–390.

[20] Feferman, S., and Strahm, T. The unfolding of non-finitist arithmetic. *Annals of Pure and Applied Logic 104*, 1–3 (2000), 75–96.

[21] Feferman, S., and Strahm, T. Unfolding finitist arithmetic. *Review of Symbolic Logic 3*, 4 (2010), 665–689.

[22] Ferreira, F. Polynomial time computable arithmetic. In *Logic and Computation, Proceedings of a Workshop held at Carnegie Mellon University, 1987*, W. Sieg, Ed., vol. 106 of *Contemporary Mathematics*. American Mathematical Society, Providence, Rhode Island, 1990, pp. 137–156.

[23] Franzén, T. Transfinite progressions: a second look at completeness. *Bulletin of Symbolic Logic 10*, 3 (2004), 367–389.

[24] Gödel, K. *Collected Works Vol. II,* S. Feferman et al. Eds. Oxford University Press, New York, 1990.

[25] Jäger, G. On Feferman's operational set theory OST. *Annals of Pure and Applied Logic 150*, 1–3 (2007), 19–39.

[26] Kreisel, G. Ordinal logics and the characterization of informal concepts of proof. In *Proceedings International Congress of Mathematicians, 14–21 August 1958*. Cambridge University Press, Cambridge, 1960, pp. 289–299.

[27] Kreisel, G. Mathematical logic. In *Lectures on Modern Mathematics*, T. L. Saaty, Ed., vol. 3. Wiley, New York, 1965, pp. 95–195.

[28] Kreisel, G. Principles of proof and ordinals implicit in given concepts. In *Intuitionism and Proof Theory*, A. Kino, J. Myhill, and R. E. Vesley, Eds. North Holland, Amsterdam, 1970, pp. 489–516.

[29] Kripke, S. Outline of a theory of truth. *Journal of Philosophy 72*, 19 (1975), 690–716.

[30] Rathjen, M. The role of parameters in bar rule and bar induction. *Journal of Symbolic Logic 56*, 2 (1991), 715–730.

[31] Schütte, K. Eine Grenze für die Beweisbarkeit der transfiniten Induktion in der verzweigten Typenlogik. *Archiv für Mathematische Logik und Grundlagen der Mathematik 7* (1964), 45–60.

[32] Sieg, W. Herbrand analyses. *Archive for Mathematical Logic 30*, 5+6 (1991), 409–441.

[33] Strahm, T. The non-constructive μ operator, fixed point theories with ordinals, and the bar rule. *Annals of Pure and Applied Logic 104*, 1–3 (2000), 305–324.

[34] Tait, W. Nested recursion. *Mathematische Annalen 143* (1961), 236–250.

[35] Tait, W. Finitism. *Journal of Philosophy 78* (1981), 524–546.

[36] Turing, A. Systems of logic based on ordinals. *Proc. London Math. Soc. 2nd ser. 45*, I (1939), 161–228.