Predicative Justification and Development of a Second Order Theory of Finite Sets (draft)

Paul J. Voda

Institute of Informatics, Comenius University Bratislava,
Mlynská dolina, 842 15 Bratislava, Slovakia,
email: voda@fmph.uniba.sk
homepage: www.fmph.uniba.sk/∼voda

1 Introduction

We will predicatively justify the induction axioms for arithmetical sets, which are the induction axioms of (the first-order) Peano Arithmetic, in the predicative second order theory of hereditarily finite sets *FSS*. The comprehension axioms of *FSS* permit the predicative formation of infinite classes of hereditarily finite sets. The predicative formation means that the infinite classes are comprehended without quantification over infinite classes.

We interprete into *FSS* the theory *FSI* which includes also the axiom of induction usable with the classes of *FSS*. This will show that the induction is consistent with *FSS*. We also show *FSI* equivalent to ACA_0.

The second order theory *FSI* was developed for its use in computer programming. For the motivation and more details see [8]. The predicative justification of *FSI* within *FSS* was inspired by S. Feferman and G. Hellman's paper *Challenges to Predicative Foundations of Arithmetic* [2].

2 Predicative Justification of Theory of Hereditarily Finite Sets and Classes

The theory *FSI* and an auxiliary theory *FSS* of finite sets and classes are formulated as first order set theories in the style of von Neumann Bernays and Gödel (NBG) see, [3].

2.1 The language of finite sets and classes. The (single-sorted) first order language $\mathcal{L}$ of our theories of finite sets and classes consists of the single non-logical symbol $\in$ of *class membership* with terms and formulas constructed in predicate calculus with identity in the usual way.

The variables of $\mathcal{L}$ are intended to range over classes definable by predicatively restricted formulas in the language of $\mathcal{L}$ (see Par. 2.2). The classes consist of *hereditarily finite sets*. Recall that the hereditarily finite sets can be generated from the empty set 0 by the binary operation $x \cup \{y\}$.

We say that the class u is a *finite set* - or just a *set* in the following - if $\exists v\, u \in v$. This last is abbreviated to $Fin(u)$. *Proper* classes are not finite. We attach to variable names no syntactic role but we will in the following use u and v, possibly with subscripts, in class positions (i.e. not in the context $u \in x$).

2.2 Restricted formulas. The formulas of $\mathcal{L}$ with all quantifiers (except those in $Fin(x)$) of the form $\exists x(Fin(x) \wedge \phi)$ and $\forall x(Fin(x) \to \phi)$ are *restricted* formulas. For a term τ without free x we treat the usual restricted quantifiers $\exists x{\in}\tau\, \phi$ and $\forall x{\in}\tau\, \phi$ as abbreviations for $\exists x(Fin(x) \wedge (x \in \tau \wedge \phi))$ and $\forall x(Fin(x) \to x \in \tau \to \phi)$. Note that predicate calculus proves the last two formulas equivalent to $\exists x(x \in \tau \wedge \phi)$ and $\forall x(x \in \tau \to \phi)$ respectively.

Axioms characterizing the classes are universal closures of:

$$\exists u \forall x(x \in u \leftrightarrow Fin(x) \wedge \phi[x, \overline{v}]) \tag{R-CA}$$

$$\forall x(x \in u \leftrightarrow x \in v) \to u = v \tag{Ext}$$

where $\phi[x, \overline{v}]$ is a restricted formula possibly with class parameters $\overline{v}$ and without u free.

Axioms characterizing the sets: are universal closures of:

$$\forall u\, u \notin s \to Fin(s) \tag{F_1}$$

$$\forall z(z \in s \leftrightarrow z \in x \vee z = y) \wedge Fin(x) \wedge Fin(y) \to Fin(s) \tag{F_2}$$

The induction axiom for sets is the universal closure of:

$$\exists s{\in}u \forall v\, v \notin s \wedge \forall x{\in}u \forall y{\in}u \exists s{\in}u \forall z(z \in s \leftrightarrow z \in x \vee z = y) \wedge Fin(x) \to x \in u \tag{Si}$$

Separation axiom is the universal closure of:

$$\forall z(z \in s \leftrightarrow z \in x \wedge z \in u) \wedge Fin(x) \to Fin(s) \tag{Sep}$$

Fig. 1. Axioms of $FS = \text{R-CA} + \text{Ext} + F_1 + F_2$, $FSI = FS + I$, and $FSS = FS + \text{Sep}$.

2.3 Axioms for classes. The axioms of *extensionality* (Ext) and of *restricted comprehension* (R-CA) given in Fig. 1 are concerned with the definition of classes.

In order to improve the readability we permit the extension of the language $\mathcal{L}$ by the introduction of new predicate and function symbols. These are introduced by a limited form of extensions by definitions (see, for instance, [6]). We admit the usual explicit definitions of predicate and functions symbols as well as the contextual definition of function symbols $f(\overline{x}) = y \leftrightarrow \phi[\overline{x}, y]$ provided we can prove the usual uniqueness condition and the existence condition in the *set form*: $\forall \overline{x} \exists y(Fin(y) \wedge \phi[\overline{x}, y])$. The formulas used in the definitions must be restricted.

We also admit the definition of *class abstractions*: $\{x|\,\phi[x,\overline{v}]\}$ where x is bound and $\overline{v}$ are free parameters. This class term is viewed as a new function symbol in the arguments $\overline{v}$ introduced by the definition:

$$\{x|\,\phi[x,\overline{v}]\} = u \leftrightarrow \forall x(x \in u \leftrightarrow Fin(x) \wedge \phi[x,\overline{v}]) \ .$$

Note that R-CA + Ext proves the existence condition by comprehension and the uniqueness condition by extensionality. We allow also explicit function definitions $f(\overline{x}) = \tau[\overline{x}]$ where the term τ may involve class terms.

For a term $\tau[x_1,\ldots,x_n,\overline{v}]$ with all free variables among the indicated ones we treat in the following the *generalized class abstraction* on the left as an abbreviation for the class abstraction on the right:

$$\{\tau[x_1,\ldots,x_n,\overline{v}]|\,\phi[x_1,\ldots,x_n,\overline{v}]\} \equiv \{y|\exists x_1(Fin(x_1) \wedge \cdots \wedge \exists x_n(Fin(x_n) \wedge y = \tau[x_1,\ldots,x_n,\overline{v}] \wedge \phi[x_1,\ldots,x_n,\overline{v}]) \cdots)\}.$$

Our form of defined symbols allows for their systematic elimination by *translation* from a formula ϕ of the extended language whereby we obtain its translated formula $\phi^* \in L$ such that R-CA + Ext $\vdash \phi \leftrightarrow \phi^*$. Every formula ϕ in the extended language can be thus viewed as defined by the formula ϕ^* in the language $\mathcal{L}$.

The translation is quite straightforward. Eliminate first the predicate and function symbols in the usual way (see [6]). Replace identities $s = t$ with class terms by $\forall x(Fin(x) \rightarrow (x \in c_1 \leftrightarrow x \in c_2))$. Replace $s \in t$ with s a class term by $\exists x(Fin(x) \wedge \phi \wedge x \in t)$ where ϕ is the formula eliminating the identity $x = s$. Replace formulas $y \in \{x|\,\phi[x]\}$ by $Fin(y) \wedge \phi[y]$. Note that the elimination of defined symbols introduces at most restricted quantifiers. We can thus freely use restricted formulas with defined symbols in comprehension.

2.4 Some defined symbols. We define the predicates of *class inclusion* by $u \subset v \leftrightarrow \forall x{\in}u\, x \in v$ and $u \subsetneq v \leftrightarrow u \subset v \wedge u \neq v$. We also define the standard class operators: $V = \{x|\,x = x\}$, $0 = \{x|\,x \neq x\}$, $\{x\} = \{y|\,y = x\}$, $u \cap v = \{x|\,x \in u \wedge x \in v\}$, $u \cup v = \{x|\,x \in u \vee x \in v\}$, $u \backslash v = \{x|\,x \in u \wedge x \notin v\}$, $U(u) = \{x|\,\exists a{\in}u\, x \in a\}$, and $P(u) = \{x|\,x \subset u\}$.

Note that R-CA + Ext proves $u \notin 0$, $(x \in V \leftrightarrow Fin(x))$, $u \subset V$, and

$$y \in \{x|\,\phi[x,\overline{u}]\} \leftrightarrow y \in V \wedge \phi[y,\overline{u}] \ .$$

2.5 Axioms for sets. We can define many classes with the class axioms of comprehension and prove many interesting properties with the help of the axiom of extensionality but we cannot prove the classes to be elements of another classes. For that we need the second group of axioms given in Fig. 1.

With the *empty* class 0 and with the *insertion* operation $u \cup \{v\}$ the *set* axioms F_1 and F_2 are equivalent in R-CA + Ext to $0 \in V$ and $x \in V \wedge y \in V \rightarrow x \cup \{y\} \in V$ respectively. Note that if $x \in V$ then $\{x\} \in V$ because $\{x\} = 0 \cup \{x\}$.

2.6 Plan. We are primarily interested in the theory *FSI* which is obtained from $FS = \text{R-CA} + \text{Ext} + F_1 + F_2$ by adjoining the *induction axiom for sets* Si. The axiom is given in the language $\mathcal{L}$ in Fig. 1 and it is in the extended language as follows:

$$0 \in u \land \forall x{\in}u \forall y{\in}u\, x \cup \{y\} \in u \to V \subset u\ . \tag{1}$$

The induction axiom asserts that any class u closed under the set operations 0 and $x \cup \{y\}$ contains all sets. In other words, the sets are (barring the nonstandard models) exactly the hereditarily finite sets.

A good argument can be made by predicativists and constructivists that the induction axiom (1), because it is of second order, is impredicative, i.e. it refers to the totality of all subsets of the hereditarily finite sets (see [5, 2]). The early predicativists (Poincaré, Weyl) accepted the second-order induction axiom for natural numbers because Poincaré claimed that every construction of natural numbers will have to involve induction in some form. Modern investigations (see for instance [2]) show that this is not the case.

In this section we will take a similar approach and we will justify *FSI* in terms of a seemingly weaker theory $FSS = FS + \text{Sep}$, where the *separation* axiom Sep (see Fig. 1 for its formulation in the language $\mathcal{L}$) can be given as

$$x \in V \to x \cap u \in V\ , \tag{2}$$

The theory *FSS* can be convincingly argued to be predicative because the only form of quantification involved in the definition of new objects by comprehension is over the inductively generetated domain of hereditarily finite sets.

For a given class L the *restriction to* L of the axioms F_1, F_2, Sep, and Si are in that order the following formulas:

$$0 \in L \tag{3}$$

$$x \in L \land y \in L \to x \cup \{u\} \in L \tag{4}$$

$$u \subset L \land 0 \in u \land \forall x{\in}u \forall y{\in}u\, x \cup \{y\} \in u \to L \subset u\ . \tag{5}$$

We will define in Thm. 2.11 a class H (intended to consist of hereditarily finite sets) such that *FSS* proves the restrictions to H of the axioms F_1, F_2, and Si. We will in Thm. 2.12 show the relative consistency of *FSI* with *FSS* by interpreting *FSI* in *FSS*.

2.7 Inductive classes. We will establish in Lemma 2.8 sufficient conditions for a class L to be *inductive*, i.e. to satisfy 2.6(5). For that we need the following definitions.

The class $u^- = \{x | x \in u \lor x \subsetneq u\}$ is the class of *predecessors* of the class u. A class u is *super-transitive*, in symbols $\mathit{Strans}(u)$, if its elements are closed under predecessors, i.e

$$\mathit{Strans}(u) \leftrightarrow \forall x{\in}u\, x^- \subset u\ .$$

The class $Str = \{a|\, Strans(a)\}$ consists of super-transitive sets. Note that a super-transitive class is transitive in the usual sense of set theory because $x \subset x^-$.

Our first condition on the class L is that any $x \in L$ be in some super-transitive set, i.e. $x \in U(Strans)$.

The induction principle for L can be proved from the principle of *super-regularity* holding for a non-empty class u if u is *super-regular*, i.e. if $\exists m{\in}u\, m^- \cap u = 0$. Thus m is minimal in u because none of its predecessors is in u. Note that a super-regular class is regular in the usual sense of set theory because for the minimal m we have $m \subset m^-$ and so $m \cap u = 0$.

Our second condition on $x \in L$ is that x be *grounded*, i.e. all sets containing x are super-regular. This can be expressed as $x \in Grnd$ where

$$Grnd = \{x|\, \forall a{\in}V(x \in a \to \exists m{\in}a\, m^- \cap a = 0)\}\ .$$

2.8 Lemma *For any class L such that*

$$FSS \vdash L \subset U(Str) \cap Grnd\ . \tag{1}$$

FSS proves the induction priniciple 2.6(5).

Proof. Assume (1), take any $u \subset L$, s.t. $0 \in u$ and

$$\forall x{\in}u \forall y{\in}u\, x \cup \{y\} \in u\ . \tag{2}$$

Suppose by the way of contradiction that $w \in L$ and $w \notin u$ for some w. From (1) we get $w \in a$ for some $a \in Str$ and so $w \in a \backslash u \in V$ by Sep. Now $w \in Grnd$ by (1) and so there is a $m \in a\backslash u$ s.t. $m^- \cap (a\backslash u) = 0$. For any $z \in m^-$ we have $z \in a$ by the super-transitivity of a and so $z \in u$. Hence $m^- \subset u$.

Since $0 \in u$ and $m \notin u$, we have $m \neq 0$ and so $y \in m$ for some y. For $x = m\backslash\{y\}$ we have $x \in V$ by Sep and $m = x \cup \{y\}$. Thus $y \in m^-$ and hence $y \in u$. Also $x \subsetneq m$ and hence $x \in m^-$ which implies $x \in u$. From $x, y \in u$ we now obtain the contradiction $m = x \cup \{y\} \in u$ by (2). □

2.9 Shortening of *Str*. Lemma 2.8 suggests that $H \subset U(Str)$ and so we are looking for a subclass of $U(Str)$ satisfying the restrictions of F_1, F_2 and Sep. *FSS* apparently cannot prove the restriction of F_2 to $U(Str)$. This will be seen in the proof of Lemma 2.10 where *FSS* can prove for $x \in a \in Str$, $y \in b \in Str$, and $c = a \cup b \cup (a\, \dot\cup \{y\})$ that $x \cup \{y\} \in c$ and $Strans(c)$ but probably not that c is a set.

The proof uses the auxiliary class $u\,\dot\cup\, v = \{x \cup v|\, x \in u\}$ for which *FSS* proves the straightforward property:

$$x, y \in V \wedge y \notin x \to (x \cup \{y\})^- = x^- \cup \{x\} \cup \{y\} \cup ((P(x)\backslash\{x\})\,\dot\cup\{y\})\ . \tag{1}$$

To remedy the situation we will employ the trick used in PA by Solovay under the name of *cut shortening* whose idea can be traced to Gentzen (see, for instance, [4]). We will shorten Str to a class $Str \cap K$ which will be proved closed under unions. The class K is defined from an auxiliary class J:

$$J = \{a|\, \forall c{\in}V\; a \cup c \in V\}$$
$$K = \{a|\, a \in J \wedge \forall c{\in}J\; a \dot\cup c \in J\}$$

2.10 Lemma *FSS proves the restrictions of F_1 and F_2 to $U(Str \cap K)$.*

Proof. We will first prove in FSS the following properties of J:

$$0 \in J \tag{1}$$
$$x \in V \to \{x\} \in J \tag{2}$$
$$a \in J \wedge b \in J \to a \cup b \in J\ . \tag{3}$$

(1): We take any $c \in V$ and obtain $0 \cup c = c \in V$. (2): For any $x, c \in V$ we have $\{x\} \cup c \in V$. (3): When $a, b \in J$ then we take any $c \in V$ and, since $b \cup c \in V$, we have $(a \cup b) \cup c = a \cup (b \cup c) \in V$ from $a \in J$.

FSS proves the following properties of K:

$$\{0\} \in K \tag{4}$$
$$a \in K \wedge b \in K \to a \cup b \in K \tag{5}$$
$$a \in K \wedge y \in V \to a \dot\cup \{y\} \in K\ . \tag{6}$$

(4): We have $\{0\} \in J$ by (2). Take any $c \in J$ and get $\{0\} \dot\cup c = \{c\} \in J$ by (2). (5): When $a, b \in K$ we have $a, b \in J$ and $a \cup b \in J$ by (3). We take any $c \in J$ and get $(a \cup b) \dot\cup c = (a \dot\cup c) \cup (b \dot\cup c) \in J$ by (3). (6): When $a \in K$, $y \in V$ then, since $a, \{y\} \in J$, we get $a \dot\cup \{y\} \in J$. We now take any $c \in J$ and we have $(a \dot\cup \{y\}) \dot\cup c = a \dot\cup (c \cup \{y\}) \in J$ because $a \in K$ and $c \cup \{y\} \in J$ by (2)(3).

We are now ready to prove the restrictions of F_1, F_2 to $U(Str \cap K)$. For F_1 we note that the set $\{0\}$ has no predecessors and so $\{0\} \in Str$. Thus $0 \in \{0\} \in Str \cap K$ by (4).

For F_2 we assume $x \in a \in Str \cap K$ and $y \in b \in Str \cap K$ for some a, b. We consider the class $c = a \cup b \cup (a \dot\cup \{y\})$ which is in K by (6) and (5). We have $x \cup \{y\} \in a \dot\cup \{y\} \subset c$ and so it suffices to establish $Strans(c)$. Thus we take any $z \in c$ and, in order to prove $z^- \subset c$, we consider three cases. If $z \in a$ then $z^- \subset a \subset c$. The case $z \in b$ is similar. If $z \in a \dot\cup \{y\}$, then $z = x_1 \cup \{y\} \in V$ for some $x_1 \in a$. If $y \in x_1$ then $z = x_1$ and we have $x_1{}^- \subset a \subset c$. Otherwise $y \notin x_1$ and, since $x_1{}^- \cup \{x_1\} \cup \{y\} \subset c$, it suffices by 2.9(1) to establish $(P(x_1) \backslash \{x_1\}) \dot\cup \{y\} \subset c$. Thus take any $x_2 \in V$ s.t. $x_2 \subsetneq x_1$. We have $x_2 \in x_1{}^- \subset a$ and hence $x_2 \cup \{y\} \in a \dot\cup \{y\} \subset c$. □

2.11 Theorem *FSS proves the restrictions of F_1, F_2, and Si to the super-transitive class $H = U(Str \cap K) \cap Grnd$.*

Proof. We first prove some properties of $Grnd$:

$$Strans(Grnd) \tag{1}$$
$$x \in V \wedge x^- \subset Grnd \to x \in Grnd\ . \tag{2}$$

(1): Take any $x \in Grnd$ and any z, a such that $z \in x^-$ and $z \in a \in V$. We wish to prove a super-regular which will prove $z \in Grnd$ and hence $x^- \subset Grnd$. Toward that end we consider two cases. If $x \in a$ we are done. Otherwise we consider the super-regular $a \cup \{x\}$. Thus for some $m \in a \cup \{x\}$ we have and $m^- \cap (a \cup \{x\}) = 0$. From the last we get $m \neq x$ and also $m^- \cap a = 0$. But then $m \in a$ and m is minimal in a.

(2): Assume the antecedent and take any $a \in V$ s.t. $x \in a$. If $x^- \cap a = 0$ then x is minimal in a. Otherwise $z \in x^- \cap a$ for some a. Thus $z \in Grnd$ and a is super-regular.

We are now ready to prove the restrictions of the set axioms to H. For F_1 we have $0 \in U(Str \cap K)$ by the restricted F_1 of Lemma 2.10. If $0 \in a \in V$ then 0 is minimal in a because $0^- \cap a = 0 \cap a = 0$. Thus also $0 \in Grnd$.

For F_2 we take any $x, y \in H$ and consider two cases. If $y \in x$ we have $x \cup \{y\} = x \in H$. Otherwise $y \notin x$. Since $x, y \in U(Str \cap K)$, we have $x \cup \{y\} \in U(Str \cap K)$ by the restricted F_2 of Lemma 2.10. In order to derive $x \cup \{y\} \in H$ it suffices to derive $x \cup \{y\} \in Grnd$ and for that it suffices by (2) to prove $(x \cup \{y\})^- \subset Grnd$. We have $x, y \in Grnd$ and $x^- \subset Grnd$ by (1). Thus by 2.9(1) it suffices to prove $b \subset Grnd$ for $b = (P(x) \backslash \{x\}) \dot\cup \{y\}$. So take any $z \in b$ and $a \in V$ s.t. $z \in a$. We have $z = x_1 \dot\cup \{y\}$ for some $x_1 \in P(x) \backslash \{x\}$, i.e. such that $x_1 \subsetneq x$. We wish to prove a super-regular and toward that end we consider two cases If some z_1 is in $((x \cup \{y\})^- \backslash b) \cap a$ then $z_1 \in Grnd$ by the above which means that a is super-regular.

Otherwise $(x \cup \{y\})^- \cap a \subset b$. We now consider the class $c = \{x_1 | x_1 \subsetneq x \wedge x_1 \cup \{y\} \in a\}$. The class is not empty as $x_1 \in c$. Moreover, $x_1 \in x^- \subset Grnd$ and, since $x \in U(Str \cap K)$, we have $x \in d$ for some $d \in Str$. Thus $c \subset P(x) \backslash \{x\} \subset x^- \subset d$ and so $c \in V$ by Sep which means that c is super-regular, i.e. $m^- \cap c = 0$ for some $m \in c$. We have $m \subsetneq x$ and $m \cup \{y\} \in a$. It cannot be the case that some $z_2 \in (m \cup \{y\})^- \cap a$ because then $z_2 \in (x \cup \{y\})^- \cap a \subset b$. Hence $z_2 = x_2 \cup \{y\}$ for some $x_2 \subsetneq m$ which leads to a contradiction $x_2 \in m^- \cap c$. Thus $m \cup \{y\}$ is minimal in a which proves a super-regular.

For Si we have $H \subset U(Str) \cap Grnd$ and so H is inductive by Lemma 2.8.

For the proof of the super-transitivity of H we take any $x \in H$. Then $x \in Grnd$ and $x^- \subset Grnd$ by (1). Also $x \in a \in Str \cap K$ for some a. Thus $x^- \subset a$ and hence $x^- \subset U(Str \cap K)$. □

2.12 Theorem *FSI is interpretable in FSS.*

Proof. We will now consider the set-theoretical structure $\mathcal{H} = \langle P(H) \cup H, \in \rangle$ where H is the (ZFC) set of the hereditarily finite sets. We will show by provability within *FSS*, i.e. by finitary means, that $\mathcal{H}$ is a model of *FSI*. This is done by the interpretation (see [6]) of *FSI* in *FSS* in such a way that for every sentence $\phi \in \mathcal{L}$ such that $FSS \vdash \phi_H$ we will have $\mathcal{H} \models \phi$. Here ϕ_H is a formula of the extended language of *FSS* called the *interpretation of* ϕ. The domain of the interpretation is $U_H(u) \leftrightarrow u \subset H$. For a formula $\phi \in \mathcal{L}$ we obtain the interpreting formula ϕ_H by restricting every quantifier in ϕ to U_H, i.e. by replacing every $\exists u \psi$ and $\forall u \psi$ by $\exists u(U_H(u) \wedge \psi)$ and $\forall u(U_H(u) \to \psi)$ respectively.

The formula $\exists u\, x \in u$, which is abbreviated to $Fin(x)$, has the interpretation $\exists u(U_H(u) \wedge x \in u)$ which FSS proves equivalent to $x \in H$. We can thus without the loss of generality identify the interpretations ϕ_H of restricted formulas ϕ with formulas obtained by replacing every $Fin(u)$ in ϕ by $x \in H$. Thus, for instance, the formula $(\exists x(Fin(x) \wedge \phi))_H$ is $\exists x{\in}H\, \phi_H$.

For a formula $\phi[u_1, \ldots, u_n]$ with all free variables indicated we define $\phi^{(H)}$ as the formula $u_1 \subset H \wedge \ldots u_n \subset H \to \phi_H$. We abbreviate this to $\overline{u} \subset H \to \phi_H[\overline{u}]$ and extend the notation also to sentences ϕ where $\phi^{(H)}$ stands for ϕ_H.

We will now interprete in $\mathcal{H}$ by working in FSS the axioms of FSI. Since the axioms of FSI are the closures of formulas R-CA, Ext, F_1, F_2, or Si from Fig 1, it is sufficient to prove $\phi^{(H)}$ for any such formula. The reader will see that the proofs in FSS of the interpretations of axioms F_1, F_2 and Si basically rephrase the corresponding properties of H from Thm. 2.11. We will heavily use the super-transitivity of H.

For a restricted ϕ the formula $(\text{R-CA})^{(H)}$ is:

$$\overline{v} \subset H \to \exists u(U_H(u) \wedge \forall x(U_H(x) \to (x \in u \leftrightarrow x \in H \wedge \phi_H[x, \overline{v}]))) \ .$$

Define $u = \{x |\, x \in H \wedge \phi_H[x, \overline{v}]\}$. Clearly $u \subset H$ and the equivalence holds for any x not just such that $x \subset H$.

The formula $\text{Ext}^{(H)}$ is:

$$u, v \subset H \to \forall x(U_H(x) \to (x \in u \leftrightarrow x \in v)) \to u = v \ .$$

Take any $x \in u$. We have $x \in H$ and hence $x \subset H$. Thus $x \in v$ by the second assumption. This proves $u \subset v$. We prove $v \subset u$ similarly.

The formula $F_1^{(H)}$ is $s \subset H \to \forall u(U_H(u) \to u \notin s) \to s \in H$. If $x \in s$ then $x \in H$ and so $x \subset H$. But we then get a contradiction $x \notin s$ from the second assumption. Thus $s = 0 \in H$ by the restriction of F_1 to H.

The formula $F_2^{(H)}$ is:

$$s, x, y \subset H \to \forall z(U_H(z) \to (z \in s \leftrightarrow z \in x \vee z = y)) \wedge x \in H \wedge y \in H \to s \in H \ .$$

From $y \subset H$ we get $y \in s$ from the second assumption. If $z \in x \subset H$ then $z \in H$. Thus $z \subset H$ and so $z \in s$ from the same assumption. Hence $x \cup \{y\} \subset s$. Vice versa, if $z \in s \subset H$ then $z \subset H$ and we have $z \in x$ or $z = y$ from the assumption again. Hence $s \subset x \cup \{y\}$. But then $s = x \cup \{y\} \in H$ by the restriction of F_2 to H.

For the proof of the formula $\text{Si}^{(H)}$ we note that under the assumption $u \subset H$ FSS proves

$$(\exists x{\in}u\, \phi)_H \leftrightarrow (\exists x(Fin(x) \wedge x \in u \wedge \phi))_H \leftrightarrow \exists x(x \in H \wedge x \in u \wedge \phi_H) \leftrightarrow \exists x{\in}u\, \phi_H \ .$$

Interpretation of restricted universal quantifiers is similar. The interpretation of Si is thus equivalent to

$$\begin{aligned} &s, u \subset H \wedge \exists s{\in}u \forall v(U_H(v) \to v \notin s) \wedge \\ &\qquad \forall x{\in}u \forall y{\in}u \exists s{\in}u \forall z(U_H(z) \to (z \in s \leftrightarrow z \in x \vee z = y)) \wedge x \in H \to x \in u \ . \end{aligned}$$

We have $\forall v(U_H(v) \to v \notin s)$ for some $s \in u \subset H$ which implies $s \subset H$. We now derive $s = 0$ just like with $F_1^{(H)}$ above. Take any $x_1, y \in u \subset H$ thus $x_1, y \subset H$. Then $\forall z(U_H(z) \to (z \in s \leftrightarrow z \in x_1 \vee z = y))$ for some $s \in u$. Since also $s \subset H$ we get $s = x_1 \cup \{y\}$ just like with $F_2^{(H)}$ above. Hence $H \subset u$ by the restriction of Si to H. This means $x \in u$. $\square$

2.13 Connection to the set theory. We have originally planned to interpret in *FSS* the theory $FSS + V = H$ instead of *FSI* whereby we would have established the consistence of the 'axiom of sets' $V = H$ with *FSS*. The difference is not substantial as already the theory *FS* can prove

$$V = H \wedge \mathrm{Sep} \leftrightarrow \mathrm{Si} \ .$$

We have abandoned our original plan because the interpetation of $V = H$ in the setup from Thm. 2.12 amounts to a straightforward prooof that the interpretation V_H of V within the domain U_H is the class H and a more complicated proof that the interpretation V_H of H takes H to itself. The last would involve absoluteness arguments needed for the interpretations of auxiliary classes involved in the definition of H. We would need to prove $Grnd_H = H$, $J_H = K_H = H$, $Str_H = Str \cap H$, $(u \subset H \to U_H(u) = U(u))$ and $u, v \subset H \to u \cap_H v = u \cap v$. Hence

$$H_H = (U(Str \cap K) \cap Grnd)_H = U_H(Str \cap H) \cap_H Grnd_H = U(Str \cap H) \cap H = H \ .$$

By proving $V = H$ consistent with *FSS* we would be in a situation similar to when ZF is proved consistent with the axiom of constructibility $V = L$. Both axioms are completeness axioms because they offer induction over sets. The sets are hereditarily finite in the case of $V = H$ and constructible sets in the case of $V = L$. The latter axiom is not considered by set theoreticians to be 'true' in the platonic domain of sets. Although its consequences (axiom of choice, continuum hypothesis) are quite fruitful logicians are reluctant to exclude non-constructible sets from their domain. This despite the fact that the axiom of constructibility is independent of ZFC because ZFC cannot prove without additional axioms that nonconstructible sets exist.

On the other hand, $V = H$ offers the just right kind of restriction of sets (i.e those satisfying *Fin*) to the hereditarily finite sets. To admit as finite additional sets is intuitively seen to be wrong. Moreover, we will see in the following section that the induction over sets permits their identification with natural numbers whereby *FSI* will be shown of the same strength as the subsystem ACA_0 of second order arithmetic.

3 Arithmetic in *FSI*

The standard set-theoretical development of natural numbers would now define the natural numbers as (finite) ordinals in the style of von Neumann. These are generated from 0 by the successor operation $x \cup \{x\}$. This approach is not suitable to us because we wish that all sets, not just the von Neumann ordinals, be natural numbers.

We will endow the class V of hereditarily finite sets with the well-order $<$ of type ω imposed on V by the Ackermann's bijection $\pi : V \mapsto N$ satisfying

$$\pi(0) = 0 \qquad \pi(x \cup \{y\}) = \begin{cases} \pi(x) & \text{if } y \in x \\ 2^{\pi(y)} + \pi(x) & \text{otherwise .} \end{cases}$$

Toward that end we wish to introduce into *FSI* the relation $<$ satisfying:

$$x < y \leftrightarrow x \in V \wedge y \in V \wedge \exists s {\in} y \backslash x \forall t {\in} x \backslash y \, t < s$$

which informally says that for $x, y \in V$ we have $x < y$ iff the largest bit in which x and y differ is set in y and not in x.

3.1 Two-sorted language for *FSI*. Since most of our development in this section will be within the class V, it will be of advantage to simplify the statement of definitions and theorems by introducing special variables for sets. This means that we effectively move from a single-sorted theory into a two-sorted theory. From now on we will use the lower case letters a, b, ..., x, y, z, possibly subscripted, to designate the variables ranging over finite sets and the upper class letters A, B, ..., X, Y, Z, possibly subscripted, to designate the variables ranging over classes. The two sorts are not disjoint as sets continue to be also classes. This arrangement is similar to that in the set theory VBG with classes [3]. Incidentally, the strength of *FSI* is the same as that of VBG minus the axiom of infinity. This parallels the well-known result of Ackermann [1] when he interpreted ZFC minus axiom of infinity into PA.

A formula ϕ in this two-sorted language has a natural translation into the formula $\phi^\star$ of the single-sorted language $\mathcal{L}$. The translation is obtained by the systematic renaming of the class variables in ϕ to new variables of $\mathcal{L}$ and by restricting all set quantifers $\exists x \psi$ and $\forall x \psi$ in ϕ to the form $\exists x(\mathit{Fin}(x) \wedge \psi)$ and $\forall x(\mathit{Fin}(x) \rightarrow \psi)$ respectively.

It is not difficult to see that for every theorem $\phi[x_1, \ldots, x_n]$ proved in the two sorted-language where $x_1, \ldots x_n$ include all free set variables of ϕ the single-sorted theory *FSI* proves the theorem

$$x_1 \in V \wedge \cdots x_n \in V \rightarrow \phi^\star[x_1, \ldots, x_n] \ .$$

The reader will note that the scheme R-CA of comprehension axioms is in the two-sorted language $\exists X \forall x(x \in X \leftrightarrow \phi[x])$ and the induction axiom Si is $0 \in X \wedge \forall x {\in} X \forall y {\in} X \, x \cup \{y\} \in X \rightarrow x \in X$. Note that the two-sorted *FSI* proves $y \in \{x | \, \phi[x]\} \leftrightarrow \phi[y]$. The axioms F_1 and F_2 are 0 and $x \cup \{y\} \in V$ respectively.

The two-sorted axiom of extensionality is $\forall x(x \in X \leftrightarrow x \in Y) \rightarrow X = Y$ and its translation $\forall x(\mathit{Fin}(x) \rightarrow (x \in u \leftrightarrow x \in v)) \rightarrow u = v$ is proved in the single-sorted *FSI* by assuming the antecedent. We wish to prove $\forall x(x \in u \leftrightarrow x \in v)$ and then use Ext. So we take any x, and consider two cases. If $\mathit{Fin}(x)$ than we get $x \in u \leftrightarrow x \in v$ from the assumption. If $\neg Fin(x)$ then $x \notin u$ and $x \notin v$ and so $x \in u \leftrightarrow x \in v$ again.

We will define in Par. 3.10 and in Thm. 3.15 functions operating over sets by contextual definitions $f(\overline{x}, \overline{Y}) = z \leftrightarrow \phi[\overline{x}, \overline{Y}, z]$. This can be done when $FSI \vdash \exists! z \phi[\overline{x}, \overline{Y}, z]$. Since the new function symbols must be defined also for proper classes, we consider such definitions as abbreviations for

$$f(\overline{X}, \overline{Y}) = Z \leftrightarrow X_1 \in V \wedge \cdots \wedge X_n \in V \wedge \phi[\overline{x}, \overline{Y}, Z] \vee \\ \neg(X_1 \in V \wedge \cdots \wedge X_n \in V) \wedge Z = 0 \ .$$

3.2 Induction theorem. *For every formula $\phi[x]$ without class quantifiers FSI proves the principles of* set induction *and of* $\in$-induction *on x with $\phi[x]$ which are in that order as follows:*

$$\phi[0] \wedge \forall x \forall y (\phi[x] \wedge \phi[y] \rightarrow \phi[x \cup \{y\}]) \rightarrow \phi[x] \tag{1}$$

$$\forall x (\forall y {\in} x\, \phi[y] \rightarrow \phi[x]) \rightarrow \phi[x] \ . \tag{2}$$

Proof. (1): By Si with $u := \{x |\, \phi[x]\}$. (2): Assume the antecedent. The conclusion follows from $\forall z {\in} x\, \phi[z]$ instantiated with $x := \{x\}$ and $z := x$. The last formula is proved by set induction (1) on x In the base case there is nothing to prove. In the inductive case we wish to prove $\forall z {\in} x \cup \{y\}\, \phi[z]$. We thus take any $z \in x \cup \{y\}$ and consider two cases. If $z \in x$ we get $\phi[z]$ from IH: $\forall z {\in} x\, \phi[z]$. If $z = y$ then, since $\forall z_1 {\in} y\, \phi[z_1]$ is IH, we intantiate the assumption with $x := z$ to obtain $\phi[z]$. □

3.3 Theorem (Closure properties of basic class operators on V). *FSI proves that $x \cap X$, $x \cup y$, $x \dot{\cup} y$, $x \backslash u$, $P(x)$, and $U(x)$ are sets.*

Proof. All classes are proved sets by set induction on x. Take, for instance, $P(x) \in V$. In the base case we have $P(0) = \{0\} \in V$ and in the inductive case $P(x \cup \{y\}) = P(x) \cup (P(x) \dot{\cup} \{y\}) \in V$ by IH and by the closure of V under $\cup$ and $\dot{\cup}$. □

3.4 Ordered pairs. Ordered pairs are defined by the well-known construction of Kuratowski-Weiner:

$$\langle u, v \rangle = \{u\} \cup \{\{u\} \cup \{v\}\} \ .$$

FSI proves the basic properties of ordered pairs in the usual way:

$$\langle x, y \rangle \in V \tag{1}$$

$$\langle x_1, y_1 \rangle = \langle x_2, y_2 \rangle \rightarrow x_1 = y_1 \wedge y_1 = y_2 \ . \tag{2}$$

3.5 Course of values sequences for $<$. For the explicit definition of $<$ we need to encode the finitely many pairs $t < s$ needed in the computation of $x < y$ from the above recurrence into a *course of values* set. This is done with the class M consisting of course of values sets:

$$M = \{m |\, \forall x \forall y (\langle x, y \rangle \in m \rightarrow \exists s {\in} y \backslash x \forall t {\in} x \backslash y\, \langle t, s \rangle \in m)\} \ .$$

FSI proves the following:

$$0 \in M \ . \tag{1}$$
$$m \in M \wedge n \in M \rightarrow m \cup n \in M \tag{2}$$
$$\exists m{\in}M \forall t{\in}a(\langle t,s\rangle \in U(M) \rightarrow \langle t,s\rangle \in m) \ . \tag{3}$$

(1): Trivial.

(2): If $\langle x,y\rangle \in m \cup n$ then if $\langle x,y\rangle \in m$ there is a $z \in y\backslash x$ s.t. for all $t \in x\backslash y$ we have $\langle t,s\rangle \in m$ and hence $\langle t,s\rangle \in m \cup n$. The case when $\langle x,y\rangle \in n$ is similar.

(3): This permits the collection of pairs $\langle t,s\rangle \in U(M)$ such that $t \in a$ into a single set $m \in M$. The property is proved by set induction on a. In the base case take $m = 0$. In the inductive case where we wish to prove the property for $a \cup \{x\}$ we obtain a set $m_1 \in M$ s.t. $\forall t{\in}a(\langle t,s\rangle \in U(M) \rightarrow \langle t,s\rangle \in m_1)$ by IH. In order to find a set m collecting from $a \cup \{x\}$ we consider two cases. If $\langle x,s\rangle \in m_2 \in M$ for some m_2 we take $m = m_1 \cup m_2 \in M$ by (2). Otherwise we just take $m = m_1$.

3.6 The definition of $<$. We explicitly define the relations $x < y$ and $x \leq y$ as follows:

$$x < y \leftrightarrow \langle x,y\rangle \in U(M)$$
$$x \leq y \leftrightarrow x < y \vee x = y \ .$$

We will now prove that $<$ satisfies the recurrence:

$$x < y \leftrightarrow \exists s{\in}y\backslash x \forall t{\in}x\backslash y \, t < s \ . \tag{1}$$

In the direction $\rightarrow$ assume $x < y$ and obtain $\langle x,y\rangle \in m$ for some $m \in M$. From the definition of M there is an $s \in y\backslash x$ s.t. $\forall t{\in}x\backslash y \, \langle t,s\rangle \in m$. But then $\forall t{\in}x\backslash y \, t < s$ from the definition of $<$.

In the direction $\leftarrow$ take any $x, y, s \in y\backslash x$ s.t. $\forall t{\in}x\backslash y \, t < s$. From the definition of $<$ we have $\forall t{\in}x\backslash y \, \langle t,s\rangle \in U(M)$ and instantiating 3.5(3) with $a = x\backslash y$ yields an $m_1 \in M$ s.t. $\forall t{\in}x\backslash y(\langle t,s\rangle \in U(M) \rightarrow \langle t,s\rangle \in m_1)$. Hence

$$\forall t{\in}x\backslash y \, \langle t,s\rangle \in m_1 \ . \tag{$*$}$$

We set $m := m_1 \cup \{\langle x,y\rangle\}$ and prove $m \in M$ by taking any x_1, y_1 s.t. $\langle x_1,y_1\rangle \in m$. Then if $\langle x_1,y_1\rangle \in m_1$ we have

$$\exists s{\in}y_1\backslash x_1 \forall t{\in}x_1\backslash y_1 \, \langle t,s\rangle \in m_1$$

because $m_1 \in M$. Hence the same holds with m instead of m_1. If $\langle x_1,y_1\rangle = \langle x,y\rangle$ then $x_1 = x$ and $y = y_1$ and we get from $(*)$ $\forall t{\in}x_1\backslash y_1 \, \langle t,s\rangle \in m$ for the above s s.t. $s \in y_1\backslash x_1$. This ends the proof of $m \in M$. Since $\langle x,y\rangle \in m$, we get $x < y$ from the definition of $<$.

3.7 Theorem *The relation $<$ orders V with the minimum 0 and no maximum, i.e. FSI proves:*

$$0 \leq x \tag{1}$$
$$\exists y\, x < y \tag{2}$$
$$x \not< x \tag{3}$$
$$x < y \land y < z \to x < z \tag{4}$$
$$x < y \lor x = y \lor y < x \ . \tag{5}$$

FSI also proves that all bits in a set are less than the set and every non-empty set has has a maximum bit set to one:

$$y \in x \to y < x \tag{6}$$
$$a \neq 0 \to \exists x{\in}a \forall y{\in}a\, y \leq x \ . \tag{7}$$

Proof. The main property used in the following proofs is the recurrence 3.6(1). We prove first the property (6) which follows from $\forall y{\in}x\, y < x$ proved by $\in$-induction on x. So take any $y \in x$. We have $y \in x \backslash y$ and for any $t \in y \backslash x$ we have $t < y$ by IH. Hence $y < x$.

(1): If $0 \neq x$ take any $s \in x$ and, since $s \in x \backslash 0$ and trivially $\forall t{\in}0 \backslash x\, t < s$, we have $0 < x$.

(2): Take any x. For $y = \{x\}$ we have $x < y$ (6).

(3): If it were the case that $x < x$ we would have $x \backslash x = 0$ non-empty.

For the surprisingly subtle and complex proof of the transitivity (4) and linearity (5) of $<$ we temporarily abbreviate by $\mathit{Trlin}[x, y, z]$ the formula:

$$(x < y \land y < z \to x < z) \land (x < y \lor x = y \lor y < x)$$

and prove by set induction on a an auxiliary property asserting the existence of the maximum bit x set to one in a non-empty set a whose elements satisfy transitivity and linearity:

$$a \neq 0 \land \forall x{\in}a \forall y{\in}a \forall z{\in}a\ \mathit{Trlin}[x, y, z] \to \exists x{\in}a \forall y{\in}a\, y \leq x \ . \tag{8}$$

The base case is trivial. In the inductive case assume (8) as IH and prove the property for $a \cup \{z\}$ by assuming

$$\forall x{\in}a \cup \{z\} \forall y{\in}a \cup \{z\} \forall z{\in}a \cup \{z\}\ \mathit{Trlin}[x, y, z] \ , \tag{$*$}$$

and considering two cases. If $a = 0$ then for the proof of $\exists x{\in}\{z\} \forall y{\in}\{z\}\, y \leq x$ it suffices to set $x := z$. If $a \neq 0$ then the second assumption of (8) is satisfied from ($*$) and so we have $\forall y{\in}a\, y \leq x_1$ for some $x_1 \in a$. By ($*$) one of $z \leq x_1$ or $x_1 < z$ holds. If the former then we set $x := x_1$ and get $\forall y{\in}a \cup \{z\}\, y \leq x_1$. If the latter then we set $x := z$. We take $y \in a \cup \{z\}$. If $y \in a$ then $y \leq x_1 < z$ and we get $y < z$ from ($*$). If $y = z$ then we have $z \leq z$ trivially. This ends the proof of (8).

Both properties (4), (5) follow directly from $\forall x \forall z\, \mathit{Trlin}[x,y,z]$. We prove the last by $\in$-induction on y where we take any y, assume as IH $\forall w{\in}y\, \forall x \forall z\, \mathit{Trlin}[x,w,z]$, and take any x and z.

For the proof of the first conjunct of $\mathit{Trlin}[x,y,z]$ we also assume $x < y < z$ and we wish to prove $x < z$. From the recurrence we have $s \in y \backslash x$ and $t \in z \backslash y$ such that

$$\forall a{\in}x\backslash y\, a < s \wedge \forall a{\in}y\backslash z\, a < t\ . \tag{$\dagger$}$$

Since $s \in y$ we get $\forall x \forall z\, \mathit{Trlin}[x,s,z]$ by IH. The case $t = s$ from $\mathit{Trlin}[t,s,t]$ leads to the contradiction $s \notin y$ and so we have either $t < s$ or $s < t$.

If $t < s$ then we cannot have $s \notin z$ because then $s \in y\backslash z$ and hence $s < t$ by ($\dagger$). $t < s < t$ leads to the contradiction $t < t$ by $\mathit{Trlin}[t,s,t]$. Thus $s \in z\backslash x$ and in order to derive $x < z$ it suffices to prove $\forall a{\in}x\backslash z\, a < s$. For that we take any $a \in x\backslash z$ and by way of contradiction assume not $a < s$. From $\mathit{Trlin}[a,s,t]$ we thus have $s \leq a$ and we consider two cases. If $a \in y$ then $a \in y\backslash z$ and so $a < t$ by ($\dagger$). Now, $\mathit{Trlin}[s,a,t]$ by IH (since $a \in y$) and so from $s \leq a < t$ we obtain $s < t$. From $t < s < t$ and $\mathit{Trlin}[t,s,t]$ we then get a contradiction $t < t$. The second case is $a \notin y$ and then $a \in x\backslash y$ and hence $a < s$ by ($\dagger$). Thus $a < s \leq a$ and $\mathit{Trlin}[a,s,a]$ leads to the contradiction $a < a$. Both cases led to a contradiction and so $a < s$ and thus $x < z$ holds.

If $s < t$ then we cannot have $t \in x$ because then $t \in x\backslash y$ and hence $t < s$ by ($\dagger$). $t < s < t$ leads to the contradiction $t < t$ by $\mathit{Trlin}[t,s,t]$. Thus $t \in z\backslash x$ and in order to derive $x < z$ it suffices to prove $\forall a{\in}x\backslash z\, a < t$. For that we take any $a \in x\backslash z$ and consider two cases. If $a \in y$ then $a \in y\backslash z$ and we get $a < t$ from ($\dagger$). If $a \notin y$ then $a \in x\backslash y$ and hence $a < s$ by ($\dagger$). From $\mathit{Trlin}[a,s,t]$ and $a < s < t$ we get $a < t$ again.

For the proof of the second conjunct of $\mathit{Trlin}[x,y,z]$ we assume $x \neq y$ and derive $x < y$ or $y < x$ by considering two cases.

If $y\backslash x = 0$ then we must have $s \in x\backslash y$ for some s because $x \neq y$. But then $\forall t{\in}y\backslash x\, t < s$ holds trivially and we have $y < x$.

If $y\backslash x \neq 0$ then

$$\forall x_1 \forall m{\in}y\backslash x \forall z\, \mathit{Trlin}[x_1,m,z] \tag{$\sharp$}$$

holds by IH because of $m \in y$. We can thus use (8) with $a := y\backslash x$ to get the maximum $m \in y\backslash x$ s.t. $\forall a{\in}y\backslash x\, a \leq m$. We now consider two cases.

If there is a $t \in x\backslash y$ such that $m < t$ then we derive $y < x$ by proving $\forall a{\in}y\backslash x\, a < t$. So take any $a \in y\backslash x$. We have $a \leq m < t$ and, since $\mathit{Trlin}[a,m,t]$ holds by ($\sharp$), we get $a < t$.

If for all $t \in x\backslash y$ we have not $m < t$ then, since $\mathit{Trlin}[t,m,t]$ by ($\sharp$), we have $t \leq m$. It cannot be the case that $t = m$ because then both $m \in x$ and $m \notin x$. Thus $\forall t{\in}x\backslash y\, t < m$. This and $m \in y\backslash x$ proves $x < y$.

(7): This directly follows from (8) by (4), and (5). □

3.8 Theorem *FSI proves that the class* $[x] = \{y \mid y < x\}$ of predecessors of *x is a set, i.e.* $[x] \in V$.

Proof. We work in *FSI* where we first prove an auxiliary property:

$$a \subset [x] \to [a \cup \{x\}] \subset P([x]) \cup ([a] \dot{\cup} \{x\}) \ . \tag{1}$$

by assuming $a \subset [x]$ and $y < a \cup \{x\}$. Thus for some $s \in (a \cup \{x\}) \backslash y$ we have

$$\forall t {\in} y \backslash (a \cup \{x\})\, t < s \ . \tag{$*$}$$

Note that $s \leq x$ because $s \in a \cup \{x\}$ and we have $a \subset [x]$. We now consider two cases.

If $x \in y$ then for $y_1 := y \backslash \{x\}$ we have $y = y_1 \cup \{x\}$ and if we succeed in proving $y_1 < a$ we will have $y = y_1 \cup \{x\} \in [a] \dot{\cup} \{x\}$. We have $s \in a \backslash y_1$ and for any $t \in y_1 \backslash a$ we have $t \in y \backslash (a \cup \{x\})$ and hence $t < s$ by $(*)$. This proves $y_1 < a$.

If $x \notin y$ then we wish to prove $y \subset [x]$ which will give us $y \in P([x])$. Thus take any $t \in y$ and consider two cases. If $t \in a$ then $t \in [x]$ from $a \subset [x]$. If $t \notin a$ then, since $t \in y \backslash a = y \backslash (a \cup \{x\})$, we get $t < s \leq x$ by $(*)$ and hence $t \in [x]$.

This ends the proof of (1). Incidentally, the property holds also with $=$ instead of the second $\subset$ For the proof of $\supset$ one has to show shows $P([x]) \subset [a \cup \{x\}]$ and $[a] \dot{\cup} \{x\} \subset [a \cup \{x\}]$.

The theorem follows from $\forall z {\in} x^- \, [z] \in V$ by taking $x := \{x\}$ and $z := x$. We prove the last by set induction on x. In the base case there is nothing to prove. We thus assume $\forall z {\in} x^- \, [z] \in V$ and $\forall z {\in} y^- \, [z] \in V$ as IHs and we wish to prove $\forall z {\in} (x \cup \{y\})^- \, [z] \in V$. If $y \in x$ then the property follows from the first IH. If $y \notin x$ then we take any $z \in (x \cup \{y\})^-$ and by 2.9(1) consider four cases. If $z \in x^-$ then we have $[z] \in V$ from the first IH. If $z = x$ then if $x = 0$ we have $[0] = 0 \in V$. Otherwise we use 3.7(7) to get the maximum bit m of x set to one. From the first IH we have $[m], [a] \in V$ for $a = x \backslash \{m\} \subsetneq x$. Since $a \subset [m]$ by the maximality of m, we get $[x] = [a \cup \{m\}] \subset P([m]) \cup ([a] \dot{\cup} \{m\})$ by (1). The last class is in V by Thm. 3.3 and, since the theorem justifies separation, we have also $[x] \in V$. When $z = y$ we prove $[y] \in V$ similarly. The fourth case is when $z = x_1 \cup \{y\}$ for some $x_1 \subsetneq x$. Thus $[x_1] \in V$ by the first IH. We now consider two subcases. If y is the maximum bit set to one in z we have $x_1 \subset [y] \in V$ and we use (1) again to derive $[z] \in V$. Otherwise, there is an $m \in x_1$ such that $y < m$. Since $m < x_1$ by 3.7(6), we have $[z] = [x_1] \in V$. □

3.9 Theorem (Least number principle and complete induction).

$$\mathit{FSI} \vdash X \neq 0 \to \exists x {\in} X\, [x] \cap X = 0 \tag{1}$$

$$\mathit{FSI} \vdash \forall x([x] \subset X \to x \in X) \to V \subset X \ . \tag{2}$$

Proof. The least number principle (1) is first proved for sets by set induction on a with $\phi[a]$ which stands for the formula (1) with X replaced by a. In the base case we have $\phi[0]$ trivially. In the inductive case we wish to prove $\phi[a \cup \{z\}]$. For that we consider two cases. If $a = 0$ then since, $[z] \cap \{z\} = 0$, we get $\phi[\{z\}]$ by taking $x := z$. If $a \neq 0$ then from the IH $\phi[a]$ we have an $x_1 \in a$ s.t. $[x_1] \cap a = 0$ and we consider two cases. If $x_1 \leq z$ then, since $[x_1] \cap \{z\} = 0$, we have $[x_1] \cap (a \cup \{z\}) = 0$ and so we get $\phi[a \cup \{z\}]$ by taking $x := x_1$. If $z < x_1$ then $[z] \cap a = 0$, $[z] \cap \{z\} = 0$ and so it suffices to take $x := z$.

We now prove $\phi[X]$ for which we assume $a \in X$ for some a. If $[a] \cap X = 0$ then it suffices to take $x := a$ in $\phi[X]$. If $[a] \cap X \neq 0$ then, since $[a] \cap X \in V$ by Thm. 3.8 and separation, we get $[x] \cap ([a] \cap X) = 0$ for some $x \in X$ s.t. $x < a$ from $\phi[[a] \cap X]$. But then $[x] \subset [a]$ and so it suffices to take $x := x$ because:

$$[x] \cap X = ([x] \cap [a]) \cap X = [x] \cap ([a] \cap X) = 0 \ .$$

The principle of complete induction (2) directly follows from $\phi[\{x | x \notin X\}]$. □

3.10 Successor function. The class $A = \{y | x < y\}$ is not empty and so there is a $y \in A$ such that $[y] \cap A = 0$ by the least number principle. For any z s.t. $x < z$ we have $z \notin [y]$, since $z \in X$, and hence $y \leq z$. *FSI* thus proves:

$$\exists y (x < y \wedge \forall z (x < z \rightarrow y \leq z))$$

Since *FSI* also proves that any such y is unique we can introduce the *successor* function $S(x)$ by

$$S(x) = y \leftrightarrow x < y \wedge \forall z (x < z \rightarrow y \leq z) \ .$$

3.11 Theorem *FSI proves*

$$S(x) \in V \tag{1}$$
$$x < S(x) \tag{2}$$
$$x < y \rightarrow S(x) \leq y \tag{3}$$
$$S(x) = S(y) \rightarrow x = y \tag{4}$$
$$0 < y \rightarrow \exists x \, S(x) = y \ . \tag{5}$$

Proof. (1), (2), and (3) follow directly from the definition of S.

For (4) we observe that if $S(x) = S(y)$ and, say $x < y$, then we get a contradiction $S(y) = S(x) \leq y$ with (2). Similarly it cannot be the case that $y < x$.

(5): Take any $y > 0$. Since $0 \in [y]$, we have $[y] \neq 0$. We have $[y] \in V$ by Thm. 3.8 and so $[y]$ has a maximum x by 3.7(7). From $x \in [y]$ we get $x < y$ and hence $S(x) \leq y$. Vice versa, since $x < S(x)$, we have $S(x) \notin [y]$ by the maximality of x and hence $y \leq S(x)$. Thus $S(x) = y$. □

3.12 Theorem (Mathematical induction). *For every formula $\phi[x]$ without class quantifiers FSI proves:*

$$0 \in X \wedge \forall x{\in}X \, S(x) \in X \rightarrow V \subset X \tag{1}$$
$$\phi[0] \wedge \forall x (\phi[x] \rightarrow \phi[S(x)]) \rightarrow \phi[x] \ . \tag{2}$$

Proof. The formula form (2) of the principle of mathematical induction follows from the class form (1) by taking $X := \{x \mid \phi[x]\}$.

We prove (1) by assuming its antecedent and by by using X with complete induction 3.9(2) where we take any x, assume $[x] \subset X$, and in order to prove $x \in X$ we consider two cases. If $x = 0$ we use the assumption $0 \in X$. If $x > 0$ then $S(y) = x$ for some y by 3.11(5). Since $y < S(y) = x$, we have $y \in [x] \subset X$. □

3.13 Finite functions. A class F is a *function*, in symbols $Fn(F)$, if

$$Fn(F) \leftrightarrow \forall a{\in}F \exists x \exists y\, a = \langle y, x\rangle \wedge \forall x \forall y_1 \forall y_2 (\langle y_1, x\rangle{\in}F \wedge \langle y_2, x\rangle{\in}F \rightarrow y_1 = y_2)\ .$$

Finite functions are functions which are sets. We define the class *Fun* of finite functions as $\mathit{Fun} = \{f \mid Fn(f)\}$.

We will also need a binary operation $F{\restriction}x$ *restricting* the function F to arguments below x. The restriction operation is defined as follows:

$$F{\restriction}x = \{\langle y, x_1\rangle \mid x_1 < x \wedge \langle y, x_1\rangle \in F\}\ .$$

FSI clearly proves $Fn(F) \rightarrow F{\restriction}x \in \mathit{Fun}$.

Functions F are *applied* to arguments x by a binary operation $F\text{'}x$ defined as follows:

$$F\text{'}x = y \leftrightarrow Fn(F) \wedge \langle y, x\rangle \in F \vee \neg Fn(F) \wedge y = 0 \vee \forall y\, \langle y, x\rangle \notin F \wedge y = 0\ .$$

3.14 Set funtions and their restrictions. Let F be a function symbol with arguments $x, \overline{y}, \overline{Y}$ introduced into *FSI* such that $\mathit{FSI} \vdash F(x, \overline{y}, \overline{Y}) \in V$. We call F a *set function*. The *restriction of F to x* is a function with the same arguments defined as follows:

$$F^{[x]}(\overline{y}, \overline{Y}) = \{\langle F(x_1, \overline{y}, \overline{Y}), x_1\rangle \mid x_1 < x\}\ .$$

One can easily see that this is a set function yielding finite functions: $F^{[x]}(\overline{y}, \overline{Y}) \in \mathit{Fun}$ and for all $z < x$ we have $F(z, \overline{y}, \overline{Y}) = F^{[x]}(\overline{y}, \overline{Y})\text{'}y$.

3.15 Theorem (Course of values recursion). *Let $\phi[f, x, \overline{y}, \overline{Y}, z]$ with all free variables indicated be a a formula without class quantifiers such that $\mathit{FSI} \vdash \exists! z \phi[f, x, \overline{y}, \overline{Y}, z]$. We can extend FSI by definitions with a new set function symbol $F(x, \overline{y}, \overline{Y})$ such that the extended theory proves:*

$$F(x, \overline{y}, \overline{Y}) = z \leftrightarrow \phi[F^{[x]}(\overline{y}, \overline{Y}), x, \overline{y}, \overline{Y}, z]\ . \tag{1}$$

Proof. In order to simplify the following proof in *FSI* we omit the parameters $\overline{y}, \overline{Y}$ which the reader can easily insert back. We first define an auxiliary set function Φ and a predicate symbol Q:

$$\Phi(f, x) = z \leftrightarrow \phi[f, x, z]$$
$$Q(f, x) \leftrightarrow f \in \mathit{Fun} \wedge \forall y{<}x\, f\text{'}y = \Phi(f{\restriction}y, y)\ .$$

We intend to define the function F as follows:

$$F(x) = z \leftrightarrow \exists f(Q(f,x) \wedge \Phi(f{\restriction}x, x) = z) \ . \tag{2}$$

Its existence condition follows from (5) and the uniqueness condition from (4). The properties are as follows:

$$Q(f,x) \wedge y < x \to Q(f,y) \tag{3}$$
$$Q(f,x) \wedge Q(g,x) \to f{\restriction}x = g{\restriction}x \tag{4}$$
$$\exists f\, Q(f,x) \tag{5}$$

(3): Straightforward.

(4): By complete induction with the class $X = \{x|\, Q(f,x) \wedge Q(g,x) \to f{\restriction}x = g{\restriction}x\}$. So assume the antecedent and take any $y < x$. Since $Q(f,y)$ and $Q(g,y)$ by (3), we get

$$f\text{'}y \overset{Q(f,x)}{=} \Phi(f{\restriction}y, y) \overset{IH}{=} \Phi(g{\restriction}y, y) \overset{Q(g,x)}{=} g\text{'}y$$

and hence $f{\restriction}x = g{\restriction}x$.

(5): By induction on x. In the base case it suffices to take $f = 0$. In the inductive case take any x and assume $Q(f_1, x)$ for some f_1 from IH. We set

$$f = f_1{\restriction}x \cup \{\langle \Phi(f_1{\restriction}x, x), x\rangle\}$$

and prove $Q(f, S(x))$. We clearly have $f \in \mathit{Fun}$. For any $z < x$ we have $f\text{'}z = f_1\text{'}z$ and hence $f{\restriction}x = f_1{\restriction}x$. We now take any $y < S(x)$. Thus $y \leq x$ by 3.11(3) and we consider two cases: If $y < x$ then $f\text{'}y = f_1\text{'}y \overset{Q(f_1,x)}{=} \Phi(f_1{\restriction}y, y) = \Phi(f{\restriction}y, y)$. If $y = x$ then $f\text{'}x = \Phi(f_1{\restriction}x, x) = \Phi(f{\restriction}x, x)$.

With the existence and uniqeness for F proved we introduce F by (2), then its restriction function $F^{[x]}$, and then we prove:

$$Q(f,x) \to f{\restriction}x = F^{[x]} \ . \tag{6}$$

Assume $Q(f,x)$. Take any $y < x$ and derive $Q(f,y)$ by (3). We then have $f\text{'}y \overset{Q(f,x)}{=} \Phi(f{\restriction}y, y) \overset{\text{def. } F}{=} F(y)$. The conclusion follows because $f{\restriction}x, F^{[x]} \in \mathit{Fun}$.

We are now ready to prove the recurrence (1) for which we have $F(x) = z$ iff $Q(f,x)$ and $\Phi(f{\restriction}x, x) = z$ for some f iff $Q(f,x)$ and $\phi[f{\restriction}x, x, z]$ for some f iff, by (6), $Q(f,x)$ and $\phi[F^{[x]}, x, z]$ for some f iff, by (5), $\phi[F^{[x]}, x, z]$. □

3.16 Addition function. Consider the following formula:

$$\phi[f,x,y,z] \equiv x = 0 \wedge z = y \vee \exists x_1(S(x_1) = x \wedge z = S(f\text{'}x_1) \ .$$

We clearly have $\mathit{FSI} \vdash \exists! z \phi[f,x,y,z]$ and so we define the binary function $\mathit{Add}(x,y)$, which we abbreviate to $x + y$, by course of values recursion to satisfy: $x + y = z \leftrightarrow \phi[\mathit{Add}^{[x]}(y), x, y, z]$. FSI then proves:

$$0 + y = y \tag{1}$$
$$S(x) + y = S(x+y) \ . \tag{2}$$

For instance, we have $S(x) + y = S(Add^{[x]}(y)\text{'}x) = S(x + y)$. *FSI* then proves by mathematical induction the associativity and commutativity of addition as well as $S(x) = x + 1$ where $1 = S(0)$.

3.17 Multiplication function. Consider the following formula:

$$\phi[f, x, y, z] \equiv x = 0 \land z = 0 \lor \exists x_1(S(x_1) = x \land z = f\text{'}x_1 + y \ .$$

We clearly have $FSI \vdash \exists! z\phi[f, x, y, z]$ and so we define the binary function $Mull(x, y)$, which we abbreviate to $x{\cdot}y$, by course of values recursion to to satisfy: $x{\cdot}y = z \leftrightarrow \phi[Mul^{[x]}(y), x, y, z]$. *FSI* then proves:

$$0{\cdot}y = 0 \tag{1}$$

$$S(x){\cdot}y = x{\cdot}y + y \ . \tag{2}$$

FSI then proves by mathematical induction the associativity, commutativity and distributivity of multiplication.

3.18 Second order arithmetic. Ssecond order arithmetic (see, for instance, [7]) is usually formulated with two sorts. The sort of *natural numbers* has four function symbols: 0, 1, +, and $\cdot$ as well as the two relation symbols $x < y$ and $x = y$. The *number variables* are $a, b, \ldots, x, y, z, \ldots$. The sort of *classes* (usually called sets) has *class variables* $X, Y, Z, \ldots$ and the atomic formulas are $\tau \in \sigma$ with τ a numeric term (constructed in the standard way in the sort of natural numbers) and σ a class variable. There is no identity over the sort of classes. Formulas are built as usual with two kinds of quantifiers: $\exists x \forall x, \exists X, \forall X$.

Thus every formula of second order arithmetic is a formula of extended *FSI*.

The axioms of the full second order arithmetic are given in three groups as universal closures of the following formulas:

1. number axioms:

$$\begin{array}{rc} x + 1 \neq 0 & x + 1 = y + 1 \to x = y \\ x + 0 = x & x + (y + 1) = (x + y) + 1 \\ x{\cdot}0 = 0 & x{\cdot}(y + 1) = (x{\cdot}y) + x \\ x \not< 0 & x < y + 1 \leftrightarrow x < y \lor x = y \end{array}$$

2. induction axiom

$$0 \in X \forall x(x \in X \to x + 1 \in X) \to \forall x\, x \in X$$

3. comprehension scheme:

$$\exists X \forall x(x \in X \leftrightarrow \phi[x])$$

where $\phi[x]$ is a formula without X free.

The subsystem ACA_0 of *arithmetical comprehension* restricts the formulas $\phi[x]$ in the comprehension scheme to those without class quantifiers.

3.19 Theorem *Every theorem of ACA_0 is a theorem of FSI.* □

3.20 Interpretation of *FSI* in ACA_0. We now informally sketch an interpretation of *FSI* in the subsystem ACA_0 of second order arithmetic. The last has a strict two-sort language where the sorts of numbers and classes are disjoint. In ACA_0 there is no identity over classes and no membership relation over sets.

The interpretation is done by a translation ϕ^* of a two-sorted formula ϕ of *FSI* into the language of ACA_0 which we extend by a number-theoretical relation $x \in_1 y$ holding iff the x-th bit of y (counting the least significant bit as 0) is set to one. We then extend ACA_0 with a two place number-theoretical function symbol $x \cup \{y\}$ by

$$x \cup \{y\} = z \leftrightarrow y \in_1 x \wedge z = x \vee y \notin_1 x \wedge z = 2^y + x$$

and prove in ACA_0 the following theorem:

$$0 \in X \wedge \forall x \forall y (x \in_1 X \wedge y \in_1 X \rightarrow x \cup \{y\} \in_1 X \rightarrow x \in_1 X \ . \tag{1}$$

We finally define by comprehension in ACA_0 the class of all natural numbers $V = \{x | \, x = x\}$

The translation ϕ^* is achieved in two stages by replacing in ϕ all atomic formulas of *FSI* involving set-membership as follows:

$$(x \in y)^* \equiv x \in_1 y \qquad (x \in Y)^* \equiv x \in Y$$
$$(X \in y)^* \equiv \exists x (X = x \wedge x \in_1 y) \qquad (X \in Y)^* \equiv \exists x (X = x \wedge x \in Y)$$

and then replacing all identities as follows:

$$(x = y)^* \equiv x = y \qquad (X = Y)^* \equiv \forall z (z \in X \leftrightarrow z \in Y)$$
$$(X = y)^* \equiv \forall z (z \in X \leftrightarrow z \in_1 x) \qquad (x = Y)^* \equiv \forall z (z \in_1 x \leftrightarrow z \in Y) \ .$$

Under this translations the axioms of *FSI* in the form from Par. 3.1 are derivable as theorems of ACA_0. Specifically, formulas R-CA become instances of comprehension in the extended ACA_0. The axiom of extensionality becomes a tautology. The axioms $0 \in V$ and $x \cup \{y\} \in V$ become theorems $0 \in V$ and $x \cup \{y\} \in V$. Finally, the axiom of set-induction becomes (1).

References

1. W. Ackermann, *Die Wiederspruchsfreiheit der allgemainen Mengenlehre*, Mathematische Annalen vol 100, 1937.
2. S. Feferman, G. Hellman, *Challenges to Predicative Foundations of Arithmetic*, in *Between Logic and Intuition Essays in Honor of Charles Parsons* G. Sher and R. Tieszen, eds., Cambridge Univ. Press, Cambridge, 2000.
3. K. Gödel, *The consistency of the continuum hypothesis*, Annals of mathematics studies no. 3. Princeton University Press, Princeton, N.J. 1940.
4. P. Hájek, P. Pudlák, *Metamathematics of first-order arithmetic*, Springer, 1991.

5. C. Parsons, *The impredicativity of induction*, in M. Detlefsen, ed., Proof, Logic and Formalization Routledge London, 1992.
6. J. Shoenfield, *Mathematical Logic*, Addison-Wesley, 1967.
7. S. Simpson, *Subsystems of Second Order Arithmetic*, Springer, 1999.
8. P. J. Voda, *Grundlagenstreit in the Theory of Programming*, www.fmph.uniba.sk/∼voda, 2003.