Klaus Aehlig

On Fragments of Analysis with Strengths of Finitely Iterated Inductive Definitions

HARLAND media

Bibliografische Information Der Deutschen Bibliothek
Die Deutsche Bibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detailierte bibliographische Daten sind im Internet über http://dnb.ddb.de abrufbar.

Dissertation, Universität München, 2003
1. Berichterstatter: Wilfried Buchholz
2. Berichterstatter: Helmut Schwichtenberg
Tag des Rigorosums: 18.07.2003

Herstellung: Books on Demand GmbH

Gedruckt auf alterungsbeständigem Papier nach ISO 9706 (säure-, holz- und chlorfrei).

Printed in Germany

ISBN 3-8330-1175-0

On Fragments of Analysis with Strengths of Finitely Iterated Inductive Definitions

Klaus Aehlig

Dissertation
an der Fakultät für Mathematik, Informatik und Statistik
der Ludwig-Maximilians-Universität München

vorgelegt am 28.04.2003

1. Berichterstatter: Wilfried Buchholz
2. Berichterstatter: Helmut Schwichtenberg
Tag des Rigorosums: 18.07.2003

Abstract

A fragment of strength and provably recursive functions of n iterated inductive definitions is obtained by restricting second order predicate logic in the following way. The underlying language allows only up to $n+1$ nested second order quantifications and those are in such a way, that no second order variable occurs free in the scope of another second order quantifier. The provably recursive functions are also characterized as the fragment of the polymorphic lambda-calculus with the same restrictions on type variables. Adding induction on arithmetical formulae to the fragments of second order predicate logic only affects the arithmetical consequences, while adding full induction increases the strength by one inductive definition. Finitely iterated inductive definitions are equivalent to the fragment of second order predicate logic restricted to a single predicate variable.

Zusammenfassung

Ein Fragment der Stärke und mit den beweisbar rekursiven Funktionen von n iterierten induktiven Definitionen erhält man durch die folgende Einschränkung der Prädikatenlogik zweiter Stufe. Die zugrunde liegende Sprache erlaubt nur bis zu $n+1$ geschachtelte zweitstufige Quantifikationen und diese sind dergestalt, daß keine zweitstufige Variable frei im Gültigkeitsbereich eines anderen zweitstufigen Quantors vorkommt. Die beweisbar rekursiven Funktionen werden auch charakterisiert als dasjenige Fragment des polymorphen lambda-Kalküls mit denselben Einschränkungen auf Typvariablen. Die Hinzunahme von Induktion über arithmetische Formeln zu den Fragmenten der Prädikatenlogik zweiter Stufe beeinflußt nur die arithmetischen Konsequenzen dieser Theorien, wohingegen die Hinzunahme voller Induktion die Stärke um eine induktive Definition erhöht. Endlich iterierte induktive Definitionen sind äquivalent zu demjenigen Fragment der Prädikatenlogik zweiter Stufe, das auf eine einzige Prädikatenvariable eingeschränkt ist.

Acknowledgments

As it is the case with almost every substantial piece of work, this thesis did not come into existence in complete isolation from the rest of the world. On the contrary, many people had a strong influence and without their support this thesis probably would not exist.

First of all I wish to thank my supervisor Wilfried Buchholz. He not only taught me proof theory, but also suggested the question that led to this thesis. The countless discussions I had with him and his numerous comments and suggestions greatly influenced this thesis in a very positive way.

I enjoyed discussing the ideas presented in this thesis with other colleagues. I wish to thank Thorsten Altenkrich, Thierry Coquand, and Wolfram Pohlers for helpful discussions.

I also thank all the colleagues in Munich, even though most of the joint work did not make its way into this thesis. It nevertheless created a pleasant and fruitful atmosphere and therefore I wish to thank Andreas Abel, Felix Joachimski, Jan Johannsen, Steffen Jost, Markus Latte, Ralph Matthes, and Monika Seisenberger.

Then I wish to thank all those, who carefully read the final draft of this thesis, that is, I wish to thank Andreas Abel, Johannes Oberreuter, and in particular Felix Joachimski, with whom I discussed almost every sentence in the present work.

Finally I wish to thank my friends. Their encouragement and support in various ways helped me to get through with the labor and frustrations connected with writing a thesis. I am particularly grateful to Johannes Oberreuter and also wish to thank Thomas Fischbacher, Marc Schäfer, and my fellow students Andreas Barth, Birgit Huber, Christoph Lenzen, and Stefan Ufer.

This work was supported by the Graduiertenkolleg „Logik in der Informatik“, financed by the Deutsche Forschungsgemeinschaft.

Contents

Preliminaries

Introduction and Related Work

Analysis is a theory concerned with real numbers. Real numbers can also be considered as sets of natural numbers and as such are, as all set-like concepts in mathematics, usually introduced by considering the set $\{x \in \mathbb{N} \mid A(x)\}$ of all natural numbers x with property $A(x)$. The philosophical justification of such a definition depends on the way the formula A is constructed. If the formula A consists only of clearly defined and previously understood functions and predicates (like $+$, $\cdot$, $=$, $<$) and quantifiers range only over the natural numbers, it can be argued that all the notions involved in this definition (including the notion of a natural number) have been understood beforehand. If, on the other hand, the formula A contains analytic quantification, that is, quantification over sets of natural numbers, the justification is very vague — while defining a set of natural numbers we already require knowledge of *all* sets of natural numbers.

Therefore a study of analysis as a formal system seems desirable in order to provide a better justification, at least for fragments of it. The first formalization is due to Takeuti [40] who introduced an extension of Gentzen's calculus LK of classical predicate logic [19, 20] and conjectured the "Hauptsatz", that is, the cut elimination property, for this system. The latter implies the consistency of the theory of real numbers. Takeuti [41] also studied fragments of analysis with Π^1_1-comprehension.

One constructively more justified way to define sets of natural numbers is by means of inductive definitions. Informally, inductive definitions are presented by giving some rules for generating elements and stating that an element should only belong to the inductively defined set if it is generated according to the rules. In other words, one considers the smallest set closed under the given rules. A good overview over inductive definitions and their use can be found in Aczel's handbook article [1]. The simplest such example is the definition of the natural numbers as the smallest set closed under zero and successor. Such finitary inductive definitions are well understood [36]. General inductive definitions have been used by Brouwer for a constructive theory of ordinals [9]. A recursive formulation was given by Church and Kleene [16]. An overview over proof theoretical aspects of inductive definitions can be found in the lecture notes volume by Buchholz, Feferman, Pohlers, and Sieg [12].

These non-finitely branching definitions can be justified particularly well, if the given rules depend on the "previously constructed elements" only in a strictly positive way. From a set theoretic point of view however, it would suffice that this dependency is monotone. In fact, known proof theoretic techniques reduce systems of monotone, or at least positive, inductive definitions to strictly positive ones. As a side effect of our investigations we will obtain another proof of that reduction.

Given the different levels of a priori justification of second order arithmetic and strictly positive inductive definitions, it is a natural goal to find fragments of second order arithmetic that can be understood by means of inductive definitions. By recursion theoretic methods Feferman [18] showed a strong form of proof theoretical equivalence, the so called "ω, T-equivalence" for various fragments of analysis and corresponding systems of iterated inductive definitions. In particular, for $\nu < \epsilon_0$ of the form $\nu = \omega^{\gamma+1}$ the systems $(\Pi^1_1 - \mathrm{CA})_{<\nu} + \mathrm{BI}$

(less than ν iterated Π^1_1-comprehension plus bar induction) and ID^c_ν (ν iterated inductive definitions) are ω, T-equivalent. And for $\nu \leq \epsilon_0$ of the form $\nu = \omega^\gamma$ with γ a limit ordinal the systems $(\Pi^1_1 - \mathrm{CA})_{<\nu} + \mathrm{BI}$ and $(\Pi^1_1 - \mathrm{CA})_{<\nu}$ and $\mathrm{ID}^c_{<\nu}$ are ω, T-equivalent.

As a special case of Feferman's result ID^c_ω is equivalent to $(\Pi^1_1 - \mathrm{CA}) + \mathrm{BI}$. In this thesis we will be concerned with systems of *finitely* iterated inductive definitions, so the limit of the systems under consideration is $\mathrm{ID}^c_{<\omega}$.

A system of strictly positive finitely iterated inductive definitions has been investigated by Martin-Löf [31], however without obtaining a proof theoretical reduction, as the proof of normalizations used the very means of the system under consideration.

In order to obtain corresponding fragments of analysis, systems of iterated inductive definitions are embedded naturally into second order arithmetic and the image of this embedding is characterized tightly enough, so that cut elimination for this fragment can be proved by means of finitely iterated inductive definitions. We provide a characterization by means of a restricted language, rather than just taking the image itself. The proof method for cut elimination is Buchholz' Ω-rule [11], which has also been used by Buchholz and Schütte for fragments of analysis [13]. As the formalization of the Ω-rule is a strictly positive inductive definition, we obtain as a side effect another proof of the reduction of systems of monotone inductive definitions to systems of strictly positive inductive definitions.

Apart from the consistency strength (say over Primitive Recursive Arithmetic) an important proof theoretic aspect of formal systems is that of the provably recursive functions, which can also be characterized by type systems for the λ-calculus [4]. As the provably recursive functions of full analysis are characterized [23] by the functions typable in the polymorphic λ-calculus [22, 37] we obtain a hierarchy of fragments of the polymorphic λ-calculus. The lowest level of this hierarchy is the "finitary subsystem of the polymorphic lambda-calculus" of Altenkirch and Coquand [2].

Outline of the Contents

This thesis is organized as follows. Section 1 introduces the involved formal systems and shows the corresponding proof theoretical equivalences. As the fragments of analysis do *not* have induction on the natural numbers in first place, we will then (in Section 2) discuss the consequences of adding induction principles. We will show a positive and a negative result. On the one hand, adding full induction on the natural numbers takes us up one level in the hierarchy of iterated inductive definitions. On the other hand, adding induction only for arithmetical formulae just affects the arithmetical consequences. Neither the consistency strength nor the provably recursive functions are affected. We will then (in Section 3) define a family of fragments of the polymorphic λ-calculus and show that they precisely characterize the provably recursive functions of the systems of finitely iterated inductive definitions. In the final Section 4 we give a particularly simple characterization of the limit of the systems under consideration: $\mathrm{ID}^c_{<\omega}$ corresponds to the fragment of second order arithmetic with the language restricted to a single set variable. Thus, in a precise sense, analysis can be understood by finitely many inductive definitions, if one restricts oneself to speak about one set at a time.

Meta Theory

The meta theory for this thesis is Primitive Recursive Arithmetic, PRA for short. That is, if we claim something to hold we claim it to hold as a theorem of PRA. More often than not, however, our theorems will read as "The following can be proved in the theory...". As all the theories under consideration in some way host Peano Arithmetic and hence reflect our meta theory, it is not only reasonable to talk about equiconsistency results, but also amounts to a stronger statement to claim that a proposition holds in the meta theory, rather than just in a particular object theory.

Notational Conventions

When introducing a symbol as a notation for some sort of objects, this always silently implies that indices, primes, and accents such as tildes, bars, hats, and checks may be added to derive further notations for the same sort of objects.

As usual, finite lists of notations differing only in successive indices, standing for the same sort of syntactical entities such as variables, terms, and formulae are abbreviated by putting an arrow over the notation for these entities. So, for example, $\vec{t}$ is short for $t_1, \dots, t_n$. This may as well include the empty list, if $n = 0$.

As usual, a list of variables following a binder is an abbreviation for successive binding of all these variables. So $\lambda\vec{x}.t$ is short for $\lambda x_1 \lambda x_2 \dots \lambda x_n.t$.

First order substitutions, denoted by σ, are finite lists of pairs of variables and terms with the variables, that is, the first components of the pairs, distinct. The list $(x_1, t_1), \dots, (x_n, t_n)$ is abbreviated by $[\vec{x}{:=}\vec{t}]$, again including the case of the empty list if $n = 0$. A substitution is applied to a term or a formula by replacing simultaneously the free variables x_i by t_i, renaming bound variables wherever necessary to avoid free variables of the t_i getting bound in the result of the substitution. In other words, all our substitutions are to be understood as "capture-free" in the usual sense. It should be noted that, up to α-equivalence, this is a well-defined notion. We hereby adopt the convention that we identify α-equal terms or formulae, unless explicitly stated otherwise. The application of a substitution σ to a term t or a formula A is denoted by postfixing its notation, as in $t\sigma$ or $A\sigma$. This notation has the narrowest possible binding; parentheses are added wherever necessary to guaranty unique reading.

We use A, B, C as notations for formulae. When displaying variables of a formula as in $A(\vec{x})$ this should signify that these variables, among others, may occur in A. This notation does not imply that these variables actually do occur free and the list $\vec{x}$ does not necessarily exhaust all the free variables of A. The purpose of such a display of variables is to distinguish certain variables so that later $A(\vec{t})$ can be used as a shorthand for the substitution $A[\vec{x}{:=}\vec{t}]$.

We also use $\mathcal{A}$ as notation for formulae, however with the silent assumption that there is one distinguished variable, as if this formula had already been displayed as $\mathcal{A}(x)$ for some x.

Second order variables are denoted by $\mathfrak{X}, \mathfrak{Y}, \dots$. All our theories will have second order variables (sometimes also called "set variables") of arity one only. In atomic propositions $\mathfrak{X}(t)$ the parentheses are usually omitted if they can be reconstructed uniquely, as in $\mathfrak{X}t$.

Second order substitutions, denoted by Σ, are finite lists of pairs of distinct second order variables and formulae with a distinguished first order variable. So, officially, second order substitutions are lists of pairs, where the second component is a pair again. We often use a centered dot as in $A(\cdot, y)$ to denote the distinguished variable. If a variable has to be distinguished then A is short for $A(\cdot)$. The latter presupposes that A has been, maybe implicitly, displayed as $A(x)$ for some x. The substitution $(\mathfrak{X}_1, \mathcal{A}_1), \ldots, (\mathfrak{X}_n, \mathcal{A}_n)$ is abbreviated by $[\vec{\mathfrak{X}}{:=}\vec{\mathcal{A}}]$. It is applied to a formula A by replacing simultaneously every atom of the form $\mathfrak{X}_i t$, with $\mathfrak{X}_i$ free, by $\mathcal{A}_i(t)$. Again this is to be done in a capture-free way. Fortunately we will only have to consider formulae where no second order variable occurs free in the scope of another; hence no renaming of second order variables will be necessary.

When displaying second order variables we use the same conventions as for first order variables. Moreover, we use $\mathfrak{A}$ and $\mathfrak{B}$ as notations for formulae with the silent assumption that there is a distinguished second order variable and a distinguished first order variable. In other words, we treat $\mathfrak{A}$ and $\mathfrak{B}$ as if they had been displayed as $\mathfrak{A}(\mathfrak{X}, x)$ for some $\mathfrak{X}$ and x.

We make use of various notations of recursion theory [33], most prominently, the coding $\langle k_1, \ldots, k_n \rangle$ of lists of natural numbers and the i'th projection $(n)_i$ of a natural number n, thought of as the code of a list. That is, $(\langle k_1, \ldots, k_n \rangle)_i = k_{i+1}$. We assume a primitive recursive such coding. Moreover, by $\{e\}(k)$ we denote the "Kleene bracket": the value, if it exists, of the e'th partial recursive function at argument k. The notation $\{e\}(k) = \ell$ presupposes that $\{e\}(k)$ is defined. We note that $\{e\}(k) = \ell$ can be expressed by a Σ_1-formula in the language of arithmetic.

We also presuppose some canonical arithmetization (also called Gödel numbering referring to the first use of this technique by Gödel [24]) of all our syntactical entities which we denote by $\ulcorner\urcorner$. We assume this Gödel numbering to be such that all the usual operations on syntactical entities are primitive recursive on the codes. In accordance with our convention that we identify α-equal formulae we assume that the codes are invariant under α-equivalence. More precisely, if A and A' are α-equal this implies $\ulcorner A \urcorner = \ulcorner A' \urcorner$. This can, for example, be achieved by using de Bruijn indices [10] for bound variables while keeping names for free ones [26]. We use dotted variables $\dot{x}$ within the Gödel brackets $\ulcorner\urcorner$ to signify that the object coded by x, rather than the code of x itself should be inserted at this position. That is, $\ulcorner A(\dot{x}) \urcorner$ is short for $\texttt{Subst}(\ulcorner A(x) \urcorner, \ulcorner x \urcorner, x)$ where $\texttt{Subst}$ is the primitive recursive substitution function on codes. It should be recalled that the implicit substitution in the notation $A(x)$ is assumed to be capture free. Since our Gödel codes are invariant under α-equivalence this is a well-defined notion. For example, $\ulcorner \lambda x.\dot{x} \urcorner$ is a term in the language of arithmetic with the variable x free. More precisely, $\ulcorner \lambda x.\dot{x} \urcorner = \texttt{Subst}(\ulcorner \lambda y.x \urcorner, \ulcorner x \urcorner, x)$. The insertion of a syntactical object at such an open position is assumed to be capture-free: for $x = \ulcorner x \urcorner$ we have that $\ulcorner \lambda x.\dot{x} \urcorner$ is $\ulcorner \lambda y.x \urcorner$ rather than the code of the identity combinator. And for $x = \ulcorner y \urcorner$ we have $\ulcorner \lambda x.\dot{x} \urcorner = \ulcorner \lambda x.y \urcorner$.

We assume that all our codings are standard, that is, they have the usual monotonicity properties. In particular, we assume that the code of a list is bigger than the code of every of its entries and the Gödel number of a formula or a term is bigger than the Gödel number of every proper subformula or subterm.

We use the dot-notation as follows. A dot stands for a pair of parentheses opening at the dot and extending as far to the right as syntactically possible.

1 Proof Theoretic Analysis

The well-known systems of iterated inductive definitions are embedded into second order arithmetic and the range of this embedding is carefully characterized. Then we analyze these fragments locally within systems of strictly positive iterated inductive definitions. As a side effect, we obtain another proof of the reduction of systems of positive (and monotone) inductive definitions to strictly positive ones.

1.1 The Systems ID_n^c, ID_n^- and ID_n^* of n-fold Iterated Inductive Definitions

Theories of inductive definitions can be motivated from the observation of set theory that every monotone operator has a least fixed point. More precisely, let $\Upsilon\colon \mathfrak{P}(\mathbb{N}) \to \mathfrak{P}(\mathbb{N})$ be monotone, that is, assume $M \subset N \subset \mathbb{N}$ implies $\Upsilon(M) \subset \Upsilon(N)$. Then one defines $\mathrm{I}_\Upsilon = \bigcap\{X \in \mathfrak{P}(\mathbb{N}) \mid \Upsilon(X) \subset X\}$ and notes that $\Upsilon(X) \subset X$ implies $\mathrm{I}_\Upsilon \subset X$ for every $X \subset \mathbb{N}$. Since Υ is monotone we have $\forall X(\Upsilon(X) \subset X \to \Upsilon(\mathrm{I}_\Upsilon) \subset \Upsilon(X) \subset X)$ and hence $\Upsilon(\mathrm{I}_\Upsilon) \subset \bigcap\{X \mid \Upsilon(X) \subset X\} = \mathrm{I}_\Upsilon$. Again by monotonicity of Υ we get $\Upsilon(\Upsilon(\mathrm{I}_\Upsilon)) \subset \Upsilon(\mathrm{I}_\Upsilon)$ and this implies $\mathrm{I}_\Upsilon \subset \Upsilon(\mathrm{I}_\Upsilon)$. So I_Υ is the least fixed point of Υ. Note however that this set theoretical argument is questionable from a constructive point of view: while defining a new subset of the natural numbers, that is to say, while defining the new subset I_Υ of natural numbers, we already assume knowledge of the full powerset $\mathfrak{P}(\mathbb{N})$, that is, we require knowledge of *all* subsets of the natural numbers.

But inductive definitions can also be understood as sets "constructed from below", once a sufficiently large segment of ordinals is given. By transfinite induction on α we define the sets $\mathrm{I}_\Upsilon^\alpha := \Upsilon(\bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi)$. Obviously for $\alpha \le \beta$ we have $\mathrm{I}_\Upsilon^\alpha \subset I_\Upsilon^\beta$, so we are constructing an increasing sequence of sets of natural numbers. This sequence of sets is entirely below the least fixed point I_Υ, as can be shown by induction on α. By induction hypothesis $\bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi \subset \mathrm{I}_\Upsilon$, hence by monotonicity of Υ we have $\mathrm{I}_\Upsilon^\alpha = \Upsilon(\bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi) \subset \Upsilon(\mathrm{I}_\Upsilon) = \mathrm{I}_\Upsilon$. Note that here we only used the fixed point property of I_Υ, not its definition.

Since set theory asserts that $\mathfrak{P}(\mathbb{N})$ is a *set*, the mapping $\mathfrak{On} \to \mathfrak{P}(\mathbb{N}), \alpha \mapsto \mathrm{I}_\Upsilon^\alpha$ from the class of ordinals into the powerset cannot be injective. So this process has to have stationary points, that is, $\mathrm{I}_\Upsilon^\alpha = \bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi$ for some α. At this α the fixed point is reached, for $\bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi = \mathrm{I}_\Upsilon^\alpha = \Upsilon(\bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi)$, hence by our first observation $\mathrm{I}_\Upsilon \subset \bigcup_{\xi<\alpha} \mathrm{I}_\Upsilon^\xi$. Since the sequence is increasing and cannot exceed I_Υ it remains stationary from that very moment on.

The theories ID_n^c formalize the n-fold iteration of this process of defining a set of natural numbers as the least fixed point of a monotone operator on sets of natural numbers. So one might argue that these are constructively justified theories. To emphasize the constructive character of these theories the operators considered are not arbitrary, but we restrict ourselves to those *definable* from the current state of knowledge. So we start by operators expressible by arithmetical formulae. After introducing their least fixed points we have understood these sets as well and may use them to define more complicated operators (and hence more complicated fixed points) and so on. This stepwise increase in knowledge of more complicated sets is reflected in a hierarchy $\mathcal{L}_0, \mathcal{L}_1, \ldots$ of richer and richer languages. To be able to define these languages syntactically, the semantical criterion of monotonicity is replaced by the syntactical criterion of positivity. Our embedding into (fragments of) second order arithmetic will show that positivity actually is only used as a sufficient, but syntactically expressible, criterion for monotonicity. This will be discussed in Remark 1.4.8.

The language of arithmetic, denoted by $\mathcal{L}_0$, consists of a single relation symbol $=$ for equality and function symbols for all primitive recursive functions. These are built from function symbols for the constant function zero of every arity (where we denote the nullary zero by 0), the unary function symbol $\mathcal{S}$ for the successor function and n-ary function symbols for the i'th projection for every $0 \leq i \leq n-1$, by arity-respecting composition and primitive recursion. The intended meaning of these function symbols is formalized in their axiomatization in Definition 1.1.2.

For systems based on a language extending that of arithmetic, we use the usual notation $\underline{n}$ for the n'th numeral, that is, if n is a natural number then $\underline{n}$ is short for $\underbrace{\mathcal{S}(\mathcal{S}(\ldots(\mathcal{S}}_{n}\,0)))$.

As logical connectives we use those of first order logic: conjunction, implication, universal quantification and absurdity. Moreover we use the other connectives as abreviations in the usual way, so $\neg A \equiv A \to \bot$, $A \vee B \equiv \neg(\neg A \wedge \neg B)$, $\exists x A \equiv \neg \forall x \neg A$.

Starting from the language $\mathcal{L}_0$ of arithmetic we define languages $\mathcal{L}_i$ and sets $\mathrm{Pos}_i, \mathrm{Neg}_i \subset \mathcal{L}_i[\mathfrak{X}]$ of positive and negative operator forms by induction on i. To do so, we set $\mathcal{L}_{i+1} = \mathcal{L}_i \cup \{\mathcal{P}^{\mathfrak{A}}_{i+1} \mid \mathfrak{A}(\mathfrak{X}, x) \in \mathrm{Pos}_i, \mathrm{FV}(\mathfrak{A}) \subset \{x, \mathfrak{X}\}\}$, where the $\mathcal{P}^{\mathfrak{A}}_{i+1}$ are new predicate symbols (to be understood as the least fixed point of the operator $\mathfrak{A}$) and Pos_i and Neg_i are simultaneously inductively defined as follows.

- If $A \in \mathcal{L}_i$ then $A \in \mathrm{Pos}_i \cap \mathrm{Neg}_i$.
- $\mathfrak{X}t \in \mathrm{Pos}_i$.
- If $A, B \in \mathrm{Pos}_i$ then $A \wedge B, \forall x A \in \mathrm{Pos}_i$.
- If $A, B \in \mathrm{Neg}_i$ then $A \wedge B, \forall x A \in \mathrm{Neg}_i$.
- If $A \in \mathrm{Neg}_i, B \in \mathrm{Pos}_i$ then $A \to B \in \mathrm{Pos}_i$.
- If $A \in \mathrm{Pos}_i, B \in \mathrm{Neg}_i$ then $A \to B \in \mathrm{Neg}_i$.

Immediately from the definition we get

Proposition 1.1.1. *If $A, B \in \mathrm{Pos}_i$ then $A \vee B, \exists x A \in \mathrm{Pos}_i$ and if $A, B \in \mathrm{Neg}_i$ then $A \vee B, \exists x A \in \mathrm{Neg}_i$.*

The system ID^c_n is based on *classical* predicate logic. Its axioms are the defining axioms for the primitive recursive function symbols, the axioms of Peano Arithmetic (in particular *full induction* on the natural numbers for *arbitrary $\mathcal{L}_n$-formulae*). Moreover, for *arbitrary formulae $\mathcal{F}$* of the language (i.e. $\mathcal{L}_n$) and $0 < i \leq n$ the following axioms are added, which define $\mathcal{P}^{\mathfrak{A}}_i$ as the least fixed point of the operator $\mathfrak{A}$.

- $\forall x(\mathfrak{A}(\mathcal{P}^{\mathfrak{A}}_i, x) \to \mathcal{P}^{\mathfrak{A}}_i x)$

- $\forall x(\mathfrak{A}(\mathcal{F},x) \to \mathcal{F}(x)) \to \forall x(\mathcal{P}_i^{\mathfrak{A}}x \to \mathcal{F}(x))$

For ease of reading we will use the abbreviation $\mathcal{A} \subset \mathcal{B} \equiv \forall x(\mathcal{A}(x) \to \mathcal{B}(x))$. With this abbreviation the defining axioms for the fixed points read

- $\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, \cdot) \subset \mathcal{P}_i^{\mathfrak{A}}$
- $\mathfrak{A}(\mathcal{F}, \cdot) \subset \mathcal{F} \to \mathcal{P}_i^{\mathfrak{A}} \subset \mathcal{F}$

Definition 1.1.2. The non-induction axioms are the following.

- $t = t$
- $s = t \to t = s$
- $s = t \to t = r \to s = r$
- $\vec{t} = \vec{s} \to \mathfrak{f}\vec{t} = \mathfrak{f}\vec{s}$, for arbitrary function symbols $\mathfrak{f}$
- $\mathcal{S}t = 0 \to \bot$
- $\mathfrak{f}0\vec{t} = \mathfrak{g}\vec{t}$ and $\mathfrak{f}(\mathcal{S}x)\vec{t} = \mathfrak{h}x(\mathfrak{f}x\vec{t})\vec{t}$ where $\mathfrak{f}$ is the function symbol for the function built by primitive recursion from $\mathfrak{g}$ and $\mathfrak{h}$
- $\mathfrak{f}\vec{t} = 0$ if $\mathfrak{f}$ is the function symbol for the n-ary zero
- $\mathfrak{f}\vec{t} = t_{i+1}$ if $\mathfrak{f}$ is the function symbol for the i'th projection
- $\mathfrak{f}\vec{s} = \mathfrak{g}(\mathfrak{h}_1\vec{s})\ldots(\mathfrak{h}_n\vec{s})$ if $\mathfrak{f}$ is the function symbol for the composition of $\mathfrak{g}$ and $\vec{\mathfrak{h}}$

Remark 1.1.3 (Injectivity of the successor). Let $\mathfrak{h}$ be the function symbol for binary first projection and Pred the function symbol for the function built by primitive recursion from $\mathfrak{h}$ and 0. We have the axioms $\mathrm{Pred}(\mathcal{S}x) = \mathfrak{h}x(\mathrm{Pred}x)$ and $\mathfrak{h}x(\mathrm{Pred}x) = x$. Hence, by transitivity of equality we have $\mathrm{Pred}(\mathcal{S}x) = x$. In particular the assumption $\mathcal{S}t = \mathcal{S}s$ implies $\mathrm{Pred}(\mathcal{S}t) = \mathrm{Pred}(\mathcal{S}s)$ which, implies $t = s$. So injectivity of the successor is derivable. Note that all these proofs are based on minimal logic.

The system ID_n^- is literally the same as ID_n^c, but based on *minimal* logic; in particular not even ex-falso-quodlibet is available. It is called the "negative fragment", because all connectives (conjunction, implication, universal quantification and absurdity) are negative. Recall that absurdity has no special meaning in minimal logic, but is just an unspecified nullary junctor. We still use $\exists$ and $\vee$ as abbreviations, but keep in mind that they do *not* behave as positive connectives.

Proposition 1.1.4.
$\mathfrak{A} \in \mathrm{Pos}_i \quad \Rightarrow \quad \mathrm{ID}_n^- \vdash \mathcal{A} \subset \mathcal{B} \to \mathfrak{A}(\mathcal{A}, \cdot) \subset \mathfrak{A}(\mathcal{B}, \cdot)$.
$\mathfrak{A} \in \mathrm{Neg}_i \quad \Rightarrow \quad \mathrm{ID}_n^- \vdash \mathcal{A} \subset \mathcal{B} \to \mathfrak{A}(\mathcal{B}, \cdot) \subset \mathfrak{A}(\mathcal{A}, \cdot)$.

Proof. Induction on $\mathfrak{A}$. For example, the second statement in the case of "$\mathfrak{A} \to \mathfrak{B}$": $\mathfrak{A} \in \mathrm{Pos}_i$, $\mathfrak{B} \in \mathrm{Neg}_i$, hence (under the assumption $\mathcal{A} \subset \mathcal{B}$) by the induction hypotheses for arbitrary x it holds that $\mathfrak{A}(\mathcal{A},x) \to \mathfrak{A}(\mathcal{B},x)$ and $\mathfrak{B}(\mathcal{B},x) \to \mathfrak{B}(\mathcal{A},x)$. Assume $\mathfrak{A}(\mathcal{B},x) \to \mathfrak{B}(\mathcal{B},x)$ and $\mathfrak{A}(\mathcal{A},x)$; hence $\mathfrak{B}(\mathcal{A},x)$ by three applications of modus ponens.

It should be noted that the proof does not use any axioms. Hence the statement is even valid in minimal predicate logic. This will be used later (in Remark 1.4.6). ◻

Proposition 1.1.5. $\mathrm{ID}_n^- \vdash \forall x(\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, x) \leftrightarrow \mathcal{P}_i^{\mathfrak{A}} x)$.

Proof. We argue within ID_n^- and have to show $\forall x(\mathcal{P}_i^{\mathfrak{A}} x \to \mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, x))$. To this end, we use the second fixed point axiom; hence we have to show $\mathfrak{A}(\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, \cdot), x) \to \mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, x)$ for arbitrary x. Since $\mathfrak{A}(\mathcal{P}_i^{A}, \cdot) \subset \mathcal{P}_i^{\mathfrak{A}}$ by the first fixed point axiom, we are done by monotonicity of $\mathfrak{A}$ (Proposition 1.1.4). ◻

Remark 1.1.6. As obviously $\mathrm{ID}_n^- \subset \mathrm{ID}_n^c$, Propositions 1.1.4 and 1.1.5 also hold for ID_n^c.

Finally we define the constructively most justified system ID_n^* to be the fragment of the system ID_n^- where all the operators $\mathfrak{A}$ of the fixed points $\mathcal{P}_i^{\mathfrak{A}}$ are strictly positive in the second order variable. That is, for $\mathfrak{A}(\mathfrak{X}, x)$ to be allowed to form a fixed point $\mathcal{P}_i^{\mathfrak{A}}$ we require that $\mathfrak{X}$ in $\mathfrak{A}$ never occurs to the left of an arrow. In other words, $\mathfrak{X}$ in $\mathfrak{A}$ is required to occur only in the conclusions, but not the premises, of implications.

1.2 Double Negation Translation of ID_n^c into the Minimal Logic Variant ID_n^-

Although classical logic has more logical rules, it is well-known that classical and minimal logic do not differ on negated formulae, as minimal logic proves stability for those, that is, $\neg\neg\neg A \to \neg A$ for arbitrary A. This observation motivates the idea [42, §10] to eliminate classical logic by doubly negating all atomic formulae. This technique, originally due to Gödel and Gentzen, can be used to embed the systems of iterated inductive definitions based on classical logic into systems of iterated inductive definitions based on minimal logic. We follow Buchholz' exposition [11] of this argument.

Definition 1.2.1. By induction on $A \in \mathcal{L}_i[\mathfrak{X}]$ we define its double negation translation A' and simultaneously showing $\mathfrak{A} \in \mathrm{Pos}_i$ (or $\mathfrak{A} \in \mathrm{Neg}_i$) implies $\mathfrak{A}' \in \mathrm{Pos}_i$ (or $\mathfrak{A}' \in \mathrm{Neg}_i$ respectively).

- $(\mathfrak{X}t)' \equiv \neg\neg\mathfrak{X}t$
- $(\mathcal{P}_i^{\mathfrak{A}} t)' \equiv \neg\neg\mathcal{P}_i^{\mathfrak{A}'} t$
- $\bot' \equiv \bot$
- $A' \equiv \neg\neg A$, if $A \in \mathcal{L}_0$, A atomic, different from $\bot$
- $(A \wedge B)' \equiv A' \wedge B'$
- $(A \to B)' \equiv A' \to B'$
- $(\forall x A)' \equiv \forall x A'$

Immediately by unfolding the abbreviations we get

Proposition 1.2.2. $(A \vee B)' \equiv \neg(\neg A' \wedge \neg B')$ *and* $(\exists x A)' \equiv \neg\forall x \neg A'$.

By a simple induction on $\mathfrak{A}$ we show

Proposition 1.2.3. $(\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, t))' \equiv \mathfrak{A}'(\mathcal{P}_i^{\mathfrak{A}'}, t)$.

Since (doubly) negated formulae are stable, a simple induction on A shows

Lemma 1.2.4. $\mathrm{ID}_n^- \vdash \neg\neg A' \leftrightarrow A'$.

In the target of our embedding the only inductive predicates are of the form $\mathcal{P}_i^{\mathfrak{A}'}$, that is, only fixed points of properly negated forms are built. We can now show that these predicates are stable.

Lemma 1.2.5. $\mathrm{ID}_i^- \vdash \neg\neg\mathcal{P}_i^{\mathfrak{A}'} t \leftrightarrow \mathcal{P}_i^{\mathfrak{A}'} t$.

Proof. We have $\mathrm{ID}_i^- \vdash \mathfrak{A}'(\mathcal{P}_i^{\mathfrak{A}'}, t) \leftrightarrow \mathcal{P}_i^{\mathfrak{A}'} t$ and $\mathrm{ID}_i^- \vdash \mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, t) \leftrightarrow \mathcal{P}_i^{\mathfrak{A}} t$ by Proposition 1.1.5. Moreover we have $\mathrm{ID}_i^- \vdash \neg\neg\mathfrak{A}'(\mathcal{P}_i^{\mathfrak{A}'}, t) \leftrightarrow \mathfrak{A}'(\mathcal{P}_i^{\mathfrak{A}'}, t)$ by Lemma 1.2.4 which shows that $\mathfrak{A}'$ is stable, independently of the stability of $\mathfrak{X}$. So the claim follows from the syntactical equality shown in Proposition 1.2.3. $\square$

In general, it is *not* the case, that $(\mathfrak{A}(\mathcal{F}, x))' \equiv \mathfrak{A}'(\mathcal{F}', x)$. Consider for example $\mathfrak{A}(\mathfrak{X},x) \equiv \mathfrak{X}x$ and $\mathcal{F}(z) \equiv \forall x.x = x$. Then $(\mathfrak{A}(\mathcal{F}, x))' \equiv \mathcal{F}'(x) \equiv (\forall x.x = x)' \equiv \forall x.\neg\neg x = x$ and $\mathfrak{A}'(\mathcal{F}', x) \equiv \neg\neg\mathcal{F}'(x) \equiv \neg\neg\forall x.\neg\neg x = x$. But at least the formulae are provably equivalent.

Proposition 1.2.6. $\mathrm{ID}_n^- \vdash (\mathfrak{A}(\mathcal{F}, x))' \leftrightarrow \mathfrak{A}'(\mathcal{F}', x)$.

Proof. Induction on $\mathfrak{A}$, using Lemma 1.2.4 for the only non-trivial case $\mathfrak{A} \equiv \mathfrak{X}t$. $\square$

To complete our embedding we have to show that the translation of the axioms are provable.

Lemma 1.2.7.
$\mathrm{ID}_n^- \vdash (\forall x(\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, x) \to \mathcal{P}_i^{\mathfrak{A}} x))'$.
$\mathrm{ID}_n^- \vdash (\forall x(\mathfrak{A}(\mathcal{F}, x) \to \mathcal{F}(x)) \to \forall x(\mathcal{P}_i^{\mathfrak{A}} x \to \mathcal{F}(x)))'$.

Proof. Using the syntactical equality shown in Proposition 1.2.3 we calculate $(\forall x(\mathfrak{A}(\mathcal{P}_i^{\mathfrak{A}}, x) \to \mathcal{P}_i^{\mathfrak{A}} x))' \equiv \forall x(\mathfrak{A}'(\mathcal{P}_i^{\mathfrak{A}'}, x) \to \neg\neg\mathcal{P}_i^{\mathfrak{A}'} x)$. This formula is provable from the corresponding induction axiom, since trivially $\mathcal{P}_i^{\mathfrak{A}'} t \to \neg\neg\mathcal{P}_i^{\mathfrak{A}'} t$.

Again we calculate $(\forall x(\mathfrak{A}(\mathcal{F}, x) \to \mathcal{F}(x)) \to \forall x(\mathcal{P}_i^{\mathfrak{A}} x \to \mathcal{F}(x)))' \equiv$
$\forall x((\mathfrak{A}(\mathcal{F}, x))' \to \mathcal{F}'(x)) \to \forall x(\neg\neg\mathcal{P}_i^{\mathfrak{A}'} x \to \mathcal{F}'(x))$. By Lemmata 1.2.5 and 1.2.6 this formula is provably equivalent to an instance of the corresponding induction axiom. $\square$

So we have completed the embedding. Summarizing we obtain

Corollary 1.2.8. *If* $\mathrm{ID}_n^c \vdash A$ *then* $\mathrm{ID}_n^- \vdash A'$.

Proof. By Lemma 1.2.7 the translations of the induction axioms are provable. The translations of the non-induction axioms (Definition 1.1.2) follow immediately from the corresponding axiom (essentially noting that $A \to B$ implies $\neg\neg A \to \neg\neg B$).

Moreover, classical logic is admissible since the translations of all formulae are stable by Lemma 1.2.4. $\square$

Since minimal logic proves $\neg\neg\neg R(x,y) \leftrightarrow \neg R(x,y)$ we get conservativity for Π^0_2-statements.

Corollary 1.2.9. *For atomic $R(x,y) \in \mathcal{L}_0$ it holds that, if $\mathrm{ID}^c_n \vdash \forall x \exists y R(x,y)$ then $\mathrm{ID}^-_n \vdash \forall x \neg \forall y \neg R(x,y)$.*

1.3 Definition of Second Order Arithmetic HA^2 and the Fragments HA^2_n and $\widetilde{\mathrm{HA}}^2_n$

Second order arithmetic formalizes the idea of complete understanding of the powerset of the natural numbers. This means that we may not only quantify over individuals (thought of as natural numbers), but also over their properties (or "sets"). As sets of natural numbers (up to appropriate coding) are real numbers this system is also called "analysis" and second order quantifiers are also called "analytic quantifiers".

Second order arithmetic is a highly impredicative system. Our formulae, properties of natural numbers, may refer to the totality of all possible properties of natural numbers. For this reason it is desirable to reduce fragments of second order arithmetic to systems that can be understood more constructively.

The system HA^2 is second order minimal predicate logic. Its language is that of arithmetic, extended by set variables and universal quantification over them. The axioms are the *non*-induction axioms of arithmetic, that is, those of Definition 1.1.2. Apart from the rules of minimal first order logic we have introduction and elimination of second order universal quantification in the following form

$$\frac{\Gamma \vdash A(\mathfrak{X}) \quad \mathfrak{X} \text{ not free in } \Gamma}{\Gamma \vdash \forall \mathfrak{X} A(\mathfrak{X})} \qquad\qquad \frac{\Gamma \vdash \forall \mathfrak{X} A(\mathfrak{X})}{\Gamma \vdash A(\mathcal{A})}$$

Since no induction is available, we should think of our universe as also containing objects which are not natural numbers. So we will often need the property of being a natural number, that is, the property of being an object for which the principle of induction holds. This predicate $\mathbb{N}x$ is defined in the usual way as $\mathbb{N}x \equiv \forall \mathfrak{X}.\forall y(\mathfrak{X}y \to \mathfrak{X}(\mathcal{S}(y))) \to \mathfrak{X}0 \to \mathfrak{X}x$.

When working in fragments of second order arithmetic we use the abbreviation $\mathcal{A} \subset \mathcal{B}$ for $\forall x.\mathbb{N}x \to \mathcal{A}(x) \to \mathcal{B}(x)$. This differs from the use of this abbreviation in the systems of iterated inductive definitions. But as these are different systems, which even have a different language, there is no danger of confusion.

We will (in Definitions 1.3.1 and 1.3.4) define two families of subsystems of HA^2. The restriction will be a restriction of the language. In HA^2_n we will allow n nested, but not "interleaved" (in the sense of Matthes [32]), second order quantifiers, so that no second order variable is allowed to occur free in the scope of another second order quantifier.

In $\widetilde{\mathrm{HA}}^2_n$ we will further restrict the use of the "lowest" second order quantification: $\mathfrak{X}_0$ may only be used to form the predicate $\mathbb{N}$. In particular, in the scope of a $\forall \mathfrak{X}_0$-quantification only atoms of the form $\mathfrak{X}_0 t$ occur.

We can, locally in ID^*_n, prove (in Corollary 1.6.13) cut-elimination for HA^2_{n+1}. Nevertheless it will turn out (in Corollary 1.4.11) that the embedding of ID^-_n into HA^2_{n+1} will use *only* the fragment $\widetilde{\mathrm{HA}}^2_{n+1}$. This will be crucial (see Proposition 3.4.3) when characterizing the functions provably recursive in ID^c_n.

Later we will also study (in Subsection 2.2) extensions of these fragments by induction on arithmetical formulae. This is important to make these systems more flexible and in particular to have a decent coding available. However, it is emphasized that the embeddings studied in Subsection 1.4 use only the system $\widetilde{\mathrm{HA}}^2_{n+1}$, as described here.

For technical reasons, it does not suffice to count the nesting depths of second order variables but we have to use second order variables of different levels (see especially the proof of Lemma 1.5.14). However, universal formulae $\forall\mathfrak{X}_k.A(\mathfrak{X}_k)$ may be instantiated with *arbitrary* formulae of the language, *as long as a legal formula is obtained* (confer Remark 1.3.8). Allowing instantiation to the *full* language is essential, as in a "predicative version" of this system only the functions of Grzegorczyk's [27] class $\mathfrak{E}_4$ would be obtained as provably recursive functions [30].

From now on, we assume a fixed assignment of levels $0, 1, \dots$ to all second-order variables except for a distinguished one, $\widehat{\mathfrak{X}}$. Let $\mathfrak{X}_i$ range over second order variables of level i. We assume that there are infinitely many second order variables of every level. The predicate $\mathbb{N}x$ uses a second order variable of level 0, that is, we have $\mathbb{N}x \equiv \forall\mathfrak{X}_0.\forall x(\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}(x))) \to \mathfrak{X}_0 0 \to \mathfrak{X}_0 x$.

Definition 1.3.1 ($\mathcal{I}^2_n$). By induction on n we define sets $\mathcal{I}^2_n[\widehat{\mathfrak{X}}]$ of formulae of second order arithmetic as follows.

- $\mathcal{I}^2_0[\widehat{\mathfrak{X}}]$ is the set of all first order $\mathcal{L}_0[\widehat{\mathfrak{X}}]$-formulae.
- $\mathcal{I}^2_{n+1}[\widehat{\mathfrak{X}}]$ is the first order closure (conjunction, implication, first order universal quantification) of $\mathcal{L}_0[\widehat{\mathfrak{X}}] \cup \{\forall\mathfrak{X}_n A[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n] \mid A \in \mathcal{I}^2_n[\widehat{\mathfrak{X}}]\}$.

We write $\mathcal{I}^2_n$ for the set of all $\mathcal{I}^2_n[\widehat{\mathfrak{X}}]$-formulae *without* free second order variables and $\mathcal{I}^2_n[\mathfrak{X}_n]$ for $\{A[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n] \mid A \in \mathcal{I}^2_n[\widehat{\mathfrak{X}}]\}$. Moreover, we define the sets $\mathcal{I}^*_n = \bigcup\limits_{k\leq n} \mathcal{I}^2_k[\mathfrak{X}_k]$ with the reading that each of the $\mathfrak{X}_k$ in the union should range over *all* the second order variables of level k.

Finally, HA^2_n is defined to be the fragment of HA^2, where all occurring formulae are in $\mathcal{I}^*_n$.

Remark 1.3.2. Note that $\mathfrak{X}_0 0 \wedge \mathfrak{X}_1 0$ is *not* a legal formula; nor is $\mathfrak{X}_0 0 \wedge \forall\mathfrak{X}_0.(\mathfrak{X}_0 0 \to \mathfrak{X}_0 0)$.

Remark 1.3.3. Usually fragments of analysis are given in the full language by restricting the way $\forall\mathfrak{X}$-quantifications are eliminated. This representation is particularly suited for proof theoretical investigations, as the elimination of $\forall\mathfrak{X}$ is the only elimination rule where the conclusion has a *higher* logical complexity than the premise.

A particular such presentation would allow elimination of $\forall\mathfrak{X}$-quantification only for variables and predicate symbols and add appropriate constants, axiomatized to represent the allowed comprehension predicates.

Consider such a presentation, where we have constants axiomatizing comprehension for all $\mathcal{I}^*_n$-formulae. Then by partial cut elimination we can achieve that all occurring formulae are subformulae of the conclusion or one of the axioms. So, if the conclusion is a $\mathcal{I}^*_n$-formula, then such a proof can be transformed into a HA_n^2 proof. In other words our systems HA_n^2 can be thought of as canonical proofs for $\mathcal{I}_n^*$-comprehension.

It may however be interesting to note that when showing the main Lemma 1.5.15 towards the admissibility of $\forall\mathfrak{X}$-elimination we actually use that we have a good overview of normal (semi-formal) proofs of $\forall\mathfrak{X}A(\mathfrak{X})$.

We now define a similar family $\widetilde{\mathcal{I}}_n^2$ of languages with the additional restriction that that level 0 quantification may only be used to define the predicate $\mathbb{N}$.

Definition 1.3.4 ($\widetilde{\mathcal{I}}_n^2$). By induction on $n \geq 1$ we define sets $\widetilde{\mathcal{I}}_n^2[\widehat{\mathfrak{X}}]$.

- $\widetilde{\mathcal{I}}_1^2[\widehat{\mathfrak{X}}]$ is the set of all first-order $\mathcal{L}_0[\widehat{\mathfrak{X}}, \mathbb{N}]$ formulae
- $\widetilde{\mathcal{I}}_{n+1}^2[\widehat{\mathfrak{X}}]$ is the first order closure of $\mathcal{L}_0[\widehat{\mathfrak{X}}] \cup \{\forall\mathfrak{X}_n A[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n] \mid A \in \widetilde{\mathcal{I}}_n^2[\widehat{\mathfrak{X}}]\}$ for all $n \geq 1$.

To make the presentation smoother, we write $\widetilde{\mathcal{I}}_0^2[\widehat{\mathfrak{X}}]$ for the set of all formulae built from atoms of the form $\mathfrak{X}_0 t$ by conjunction, implication and (first order) universal quantification.

We write $\widetilde{\mathcal{I}}_n^2$ for the set of all $\widetilde{\mathcal{I}}_n^2[\widehat{\mathfrak{X}}]$-formulae *without* free second order variables and $\widetilde{\mathcal{I}}_n^2[\mathfrak{X}_n]$ for $\{A[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n] \mid A \in \widetilde{\mathcal{I}}_n^2[\widehat{\mathfrak{X}}]\}$. Moreover, we define the sets $\widetilde{\mathcal{I}}_n^* = \bigcup_{k \leq n} \widetilde{\mathcal{I}}_k^2[\mathfrak{X}_k]$ again with the reading that each of the $\mathfrak{X}_k$ in the union should range over *all* the second order variables of level k.

Finally, $\widetilde{\mathrm{HA}}_n^2$ is defined to be the fragment where all occurring formulae are in $\widetilde{\mathcal{I}}_n^*$.

Remark 1.3.5. $\mathcal{I}_n^* \subset \mathcal{I}_{n+1}^*$ and $\widetilde{\mathcal{I}}_n^* \subset \mathcal{I}_n^*$. In particular, $\widetilde{\mathrm{HA}}_n^2$ is a subsystem of HA_n^2.

Remark 1.3.6. We note that, up to (level-*ignoring*) α-equality all formulae in $\mathcal{I}_n^2$ are built from a single second order variable. This can be proved by induction on the Definition 1.3.1 of $\mathcal{I}_n^2$. In Section 4 we will further analyze this property.

Our sets $\mathcal{I}_n^2[\mathfrak{X}_n]$ are closed under substitution. More precisely, by simple induction on A one shows

Lemma 1.3.7. *If* $A(\mathfrak{X}_n), \mathcal{A} \in \mathcal{I}_n^2[\mathfrak{X}_n]$ *then* $A(\mathcal{A}) \in \mathcal{I}_n^2[\mathfrak{X}_n]$.

Remark 1.3.8. Note that in Lemma 1.3.7 it was crucial that the free second order variable of $\mathcal{A}$ is of level n. Substitution in free variables of too low level might lead to illegal formulae. Consider for example $A(\mathfrak{X}_1) \equiv \mathfrak{X}_1 0 \wedge \forall\mathfrak{X}_0.\mathfrak{X}_0 0 \to \mathfrak{X}_0 0$ and $\mathcal{A} \equiv \mathfrak{X}_0$. Then $A(\mathcal{A})$ is the non well-formed expression $\mathfrak{X}_0 0 \wedge \forall\mathfrak{X}_0.\mathfrak{X}_0 0 \to \mathfrak{X}_0 0$.

This will not be a problem when embedding ID_n^- into HA_n^2 as all formulae we deal with in this embedding will contain no free second order variables. However, we have to be very careful about this when performing the analysis of HA_n^2.

Lemma 1.3.9. $\widetilde{\mathrm{HA}}_1^2 \vdash \mathbb{N}0$.

Proof. Trivial argument within $\widetilde{\mathrm{HA}}_1^2$. □

Lemma 1.3.10. $\widetilde{\mathrm{HA}}^2_1 \vdash \forall x(\mathbb{N}x \to \mathbb{N}(\mathcal{S}(x)))$.

Proof. Argue within $\widetilde{\mathrm{HA}}^2_1$. Assume $\mathbb{N}x$. Let $\mathfrak{X}$ be arbitrary and assume $\forall y.\mathfrak{X}y \to \mathfrak{X}(\mathcal{S}(y))$ and $\mathfrak{X}0$. Hence $\mathfrak{X}x$ by $\mathbb{N}x$ and $\mathfrak{X}(\mathcal{S}(x))$ by the assumption. ∎

1.4 Embedding of ID_n^- into $\widetilde{\mathrm{HA}}^2_{n+1}$

As second order arithmetic is an impredicative system, the set theoretical definition $\mathrm{I}_\Upsilon = \bigcap\{X \in \mathfrak{P}(\mathbb{N}) \mid \Upsilon(X) \subset X\}$ of the least fixed point can be formalized easily. The argument (given in the introduction to Subsection 1.1) showing that I_Υ indeed *is* the least fixed point can be formalized as well. Hence, systems of iterated inductive definitions can easily be embedded into second order arithmetic. In this subsection we study the embedding and focus on the language of the embedded formulae. More precisely we will show (in Corollary 1.4.12) that ID_n^- can be embedded into $\widetilde{\mathrm{HA}}^2_{n+1}$.

Definition 1.4.1. For $A \in \mathcal{L}_n[\mathfrak{X}_n]$ we define a formula A^* of second order arithmetic inductively as follows.

- $(\mathcal{P}_i^{\mathfrak{A}}t)^* \equiv \forall \mathfrak{X}_i.(\mathfrak{A}(\mathfrak{X}_i, \cdot))^* \subset \mathfrak{X}_i \to \mathfrak{X}_i t$
- $(\mathfrak{X}_i t)^* \equiv \mathfrak{X}_i t$
- $A^* \equiv A$ for atomic $A \in \mathcal{L}_0$
- $(A \wedge B)^* \equiv A^* \wedge B^*$
- $(\forall x A)^* \equiv \forall x.\mathbb{N}x \to A^*$
- $(A \to B)^* \equiv A^* \to B^*$

A simple induction shows that the range of the embedding is a very restricted fragment of the language of second order arithmetic. More precisely we have

Lemma 1.4.2. *If $A \in \mathcal{L}_n$ then $A^* \in \widetilde{\mathcal{I}}^2_{n+1}$. In particular A^* has* no *free second order variables.*

Lemma 1.4.3. $\mathfrak{A}^*((\mathcal{P}_i^{\mathfrak{A}})^*, x) \in \widetilde{\mathcal{I}}^2_{i+1}$.

Proof. Immediate from Lemma 1.4.2, noting that for $\mathfrak{A} \in \mathrm{Pos}_i$ the predicate $\mathfrak{X}$ does not occur under any second-order quantifier (since Pos_i only contains first-order formulae in $\mathfrak{X}$). ∎

Remark 1.4.4. When introducing second order arithmetic we also fixed a different use of the symbol $\subset$. However, via the *-embedding these two notions are coherent. More precisely $(\mathcal{A} \subset \mathcal{B})^* \equiv \mathcal{A}^* \subset \mathcal{B}^*$.

Since all arithmetical quantifiers are relativized to $\mathbb{N}$, the principle of induction on the natural numbers holds for the translated formulae. More precisely we have

Lemma 1.4.5. *For every formula $\mathcal{A}$ of* ID_n^-,

$$\widetilde{\mathrm{HA}}^2_{n+1} \vdash (\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \forall x\mathcal{A}(x))^*.$$

Proof. Unfolding the definition yields $(\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \forall x\mathcal{A}(x))^*$ $\equiv (\forall x.\mathbb{N}x \to \mathcal{A}^*(x) \to \mathcal{A}^*(\mathcal{S}x)) \to \mathcal{A}^*(0) \to \forall x.\mathbb{N}x \to \mathcal{A}^*(x)$.

Arguing informally within $\widetilde{\mathrm{HA}}^2_{n+1}$ we assume $\forall x.\mathbb{N}x \to \mathcal{A}^*(x) \to \mathcal{A}^*(\mathcal{S}x)$ and $\mathcal{A}^*(0)$. Moreover let x be arbitrary and assume $\mathbb{N}x$. We have to show $\mathcal{A}^*(x)$.

Instantiate $\mathbb{N}x$ to $\mathcal{A}^*(\cdot) \wedge \mathbb{N}(\cdot)$ to obtain $\forall y((\mathcal{A}^*(y) \wedge \mathbb{N}y) \to (\mathcal{A}^*(\mathcal{S}(y)) \wedge \mathbb{N}(\mathcal{S}(y)))) \to (\mathcal{A}^*(0) \wedge \mathbb{N}0) \to (\mathcal{A}^*(x) \wedge \mathbb{N}x)$. Since we have $\mathcal{A}^*(0)$ and (by Lemma 1.3.9) also $\mathbb{N}0$, it suffices to show $\forall y.(\mathcal{A}^*(y) \wedge \mathbb{N}y) \to (\mathcal{A}^*(\mathcal{S}(y)) \wedge \mathbb{N}(\mathcal{S}(y)))$. So, for arbitrary y assume $\mathcal{A}^*(y) \wedge \mathbb{N}y$. From our first assumption we get $\mathcal{A}^*(\mathcal{S}y)$ and from Lemma 1.3.10 we get $\mathbb{N}(\mathcal{S}y)$. ∎

Remark 1.4.6. As in the proof of proposition 1.1.4 one shows that the translations of positive formulae are monotone.

The following lemma formalizes the informal arguments given at the beginning of Subsection 1.1.

Lemma 1.4.7.
$\widetilde{\mathrm{HA}}^2_{i+1} \vdash (\mathfrak{A}(\mathcal{P}^{\mathfrak{A}}_i, \cdot) \subset \mathcal{P}^{\mathfrak{A}}_i)^*$.
$\widetilde{\mathrm{HA}}^2_{n+1} \vdash (\mathfrak{A}(\mathcal{F}, \cdot) \subset \mathcal{F} \to \mathcal{P}^{\mathfrak{A}}_i \subset \mathcal{F})^*$ *for* $\mathcal{F} \in \mathcal{L}_n$ *and* $i \leq n$.

Proof. By definition we have $(\mathfrak{A}(\mathcal{P}^{\mathfrak{A}}_i, \cdot) \subset \mathcal{P}^{\mathfrak{A}}_i)^* \equiv \forall x.\mathbb{N}x \to \mathfrak{A}^*((\mathcal{P}^{\mathfrak{A}}_i)^*, x) \to \forall \mathfrak{X}_i.\mathfrak{A}^*(\mathfrak{X}_i, \cdot) \subset \mathfrak{X}_i \to \mathfrak{X}_i x$, where $(\mathcal{P}^{\mathfrak{A}}_i(y))^* \equiv \forall \mathfrak{X}_i.(\mathfrak{A}(\mathfrak{X}_i, \cdot))^* \subset \mathfrak{X}_i \to \mathfrak{X}_i y$.

Arguing informally in $\widetilde{\mathrm{HA}}^2_{n+1}$, the translation can be proved as follows. Let x be arbitrary and assume $\mathbb{N}x$ and $\mathfrak{A}^*(\mathcal{P}^{\mathfrak{A}^*}_i, x)$. Let $\mathfrak{X}_i$ be arbitrary and assume $\mathfrak{A}^*(\mathfrak{X}_i, \cdot) \subset \mathfrak{X}_i$. From $(\mathfrak{A}(\mathfrak{X}_i, \cdot))^* \subset \mathfrak{X}_i$ we get $(\mathcal{P}^{\mathfrak{A}}_i)^* \subset \mathfrak{X}_i$. Hence, by monotonicity (remark 1.4.6) of $\mathfrak{A}^*$ we get $\mathfrak{A}^*((\mathcal{P}^{\mathfrak{A}}_i)^*) \subset \mathfrak{A}^*(\mathfrak{X}_i, \cdot) \subset \mathfrak{X}_i$, so we can conclude $\mathfrak{X}_i x$, since $\mathfrak{A}^*((\mathcal{P}^{\mathfrak{A}}_i)^*, x)$ and $\mathbb{N}x$.

By definition we have $(\mathfrak{A}(\mathcal{F}, \cdot) \subset \mathcal{F} \to \mathcal{P}^{\mathfrak{A}}_i \subset \mathcal{F})^* \equiv \mathfrak{A}^*(\mathcal{F}^*, \cdot) \subset \mathcal{F}^* \to \forall x.\mathbb{N}x \to (\forall \mathfrak{X}_i.\mathfrak{A}^*(\mathfrak{X}_i, \cdot) \subset \mathfrak{X}_i \to \mathfrak{X}_i x) \to \mathcal{F}^*(x)$. Arguing informally in $\widetilde{\mathrm{HA}}^2_{n+1}$ this can be proved as follows. Assume $\mathfrak{A}^*(\mathcal{F}^*, \cdot) \subset \mathcal{F}^*$. Let x be arbitrary and moreover assume $\mathbb{N}x$ and $\forall \mathfrak{X}_i.\mathfrak{A}^*(\mathfrak{X}_i, \cdot) \subset \mathfrak{X}_i \to \mathfrak{X}_i x$. Specializing the last assumption to $\mathcal{F}^*$ we get $\mathcal{F}^*(x)$ from $\mathfrak{A}^*(\mathcal{F}^*, \cdot) \subset \mathcal{F}^*$ which is our first assumption. ∎

Remark 1.4.8 (Monotone inductive definitions). Note that in the proof of Lemma 1.4.7 only monotonicity of $\mathfrak{A}^*$ was used. The only point in restricting $\mathcal{P}^{\mathfrak{A}}_i$ to positive $\mathfrak{A}$ is to have a canonical proof that $\mathfrak{A}^*$ is monotone. So one might be tempted to allow $\mathcal{P}^{\mathfrak{A}}_i$ for arbitrary $\mathfrak{A}$ and relativize the first axiom to $\mathfrak{A}$ being monotone. However, monotonicity is a second order concept and does not fit well into the arithmetical framework of systems of inductive definitions. Note that even in HA^2_{n+1}, due to the restriction to one second-order variable, some mild coding is necessary to express monotonicity: $\forall \mathfrak{X}.\mathfrak{X}\langle 0, \cdot\rangle \subset \mathfrak{X}\langle 1, \cdot\rangle \to \ldots$. The admissibility of the necessary induction to prove the needed properties of pairing and projection will be discussed in Subsection 2.2.

On the other hand, this thought shows that the following "obvious formulation" of a system of *monotone* iterated inductive definitions can be embedded into $\widetilde{\mathrm{HA}}^2_{n+1}$, and therefore (as will be shown in Corollary 1.6.19) can be reduced to ID^*_n.

The predicate symbol $\mathcal{P}^{\mathfrak{A}}_{i+1}$ can be formed for *arbitrary* formulae $\mathfrak{A} \in \mathcal{L}_i[\mathfrak{X}]$. The induction axiom $\mathfrak{A}(\mathcal{F}, \cdot) \subset \mathcal{F} \to \mathcal{P}^{\mathfrak{A}}_i \subset \mathcal{F}$ is always present and we have a rule that, from $\mathfrak{X} \subset \mathfrak{Y} \to \mathfrak{A}(\mathfrak{X}, \cdot) \subset \mathfrak{A}(\mathfrak{Y}, \cdot)$ for *new predicate variables* $\mathfrak{X}$ *and* $\mathfrak{Y}$ we may conclude $\mathfrak{A}(\mathcal{P}^{\mathfrak{A}}_i, \cdot) \subset \mathcal{P}^{\mathfrak{A}}_i$.

Proof. Careful inspection of the proof of lemma 1.4.7. The translation of the proof of $\mathfrak{X} \subset \mathfrak{Y} \to \mathfrak{A}(\mathfrak{X}, \cdot) \subset \mathfrak{A}(\mathfrak{Y}, \cdot)$ yields proofs of the monotonicity statements $\mathcal{A} \subset \mathcal{B} \to \mathfrak{A}(\mathcal{A}, \cdot) \subset \mathfrak{A}(\mathcal{B}, \cdot)$ from the same assumptions, since $\mathfrak{X}$ and $\mathfrak{Y}$ are new. $\square$

Lemma 1.4.9. *For every term t we have a proof of t being total. More precisely, if the free variables of t are among $\vec{x}$ then $\widetilde{\mathrm{HA}}^2_1 \vdash \mathbb{N}\vec{x} \to \mathbb{N}t$.*

Proof. By induction on the buildup of the term t and the buildup of symbols for primitive recursive functions we construct a proof in $\widetilde{\mathrm{HA}}^2_1$. For the base cases 0 and $\mathcal{S}$ use Lemmata 1.3.9 and 1.3.10. For the case of a recursion we use induction on the natural numbers, as provided by Lemma 1.4.5. $\square$

Remark 1.4.10. Note that the proofs in Lemma 1.4.9 have type (in a sense to be made precise in Subsection 3.4) $\vec{\mathcal{N}} \to \mathcal{N}$ and might contain lots of inductions (in the sense of $\forall\mathfrak{X}_0$-elimination for forumlae of the from $\mathbb{N}s$). This is precisely the point where primitive recursive function symbols enter the second order theory.

It should be noted that in ID^c_0 all primitive recursive functions can be proved to be total by specifying the corresponding primitive recursive function symbol (no proof obligations remain). However when "extracting programs" (in the sense of Proposition 3.4.8; compare also Definition 3.4.4) the first order structure is ignored (only the second order structure and logic remain). This should also be compared with the use of this lemma in Corollary 1.4.11.

Summarizing the results we obtain

Corollary 1.4.11. *If $\mathrm{ID}^-_n \vdash A$ then $\widetilde{\mathrm{HA}}^2_{n+1} \vdash A^*$.*

Proof. From Lemmata 1.4.7 and 1.4.5 and Lemma 1.4.9 for the admissibility of the $\forall x$-elimination rule (which is now relativized to $\mathbb{N}$). $\square$

In particular, for Π^0_2-statements this corollary reads as

Corollary 1.4.12. *For atomic $R(x,y) \in \mathcal{L}_0$, if $\mathrm{ID}^-_n \vdash \forall x \neg\forall y \neg R(x,y)$ then $\widetilde{\mathrm{HA}}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$.*

1.5 Semi-Formal Proofs for the Fragment $\mathrm{HA}^2_{n^*}$

From a proof theoretic point of view, the most difficult rule of second order arithmetic is the elimination of a second order quantifier, as this rule significantly lacks the subformula property. In fact, when stepping from $\forall\mathfrak{X}A(\mathfrak{X})$ to $A(\mathcal{A})$ the formula might get arbitrary complex.

In this subsection we shall develop a semi-formal notion of "normal proofs", that is, of proofs with the subformula property, in such a way that $\forall\mathfrak{X}$-elimination becomes admissible. When trying to do so one should first understand that $\forall\mathfrak{X}$-elimination is *not* admissible in the usual notion of normal proof for second order arithmetic. This might surprise, as morally, in a normal proof a universal statement should be proved by the corresponding introduction rule — and substitution of $\mathcal{A}$ for $\mathfrak{X}$ in the premise of an $\forall\mathfrak{X}$-introduction rule does indeed yield a proof of the conclusion of the elimination. However, proofs might contain assumptions and in a proof of $\forall\mathfrak{X}.A(\mathfrak{X})$ *by assumption* there is nothing to substitute.

The Ω-rule, introduced by Buchholz [11] and used in the context of second order arithmetic by Buchholz and Schütte [13], solves this problem by anticipating cut elimination. If finally we consider a proof with no open assumptions, then in every subproof the open assumptions will sooner or later be replaced by proofs. So it suffices to have a proof of $A(\mathcal{A})$ for every possible proof of $\forall\mathfrak{X}_n.A(\mathfrak{X}_n)$ that does not have assumptions containing $\forall\mathfrak{X}_n$-formulae positively. But such proofs are proofs *by introduction* and hence the substitution argument works.

Before we can define semi-formal proofs we first need to make precise which sets of possible assumptions cannot be used to prove $\forall\mathfrak{X}_nA(\mathfrak{X}_n)$ without giving a recipe on how to prove $A(\mathcal{A})$ for all possible $\mathcal{A}$. These are the sets Neg^2_n containing second order universal quantification of level n only negatively.

Definition 1.5.1. We inductively define sets Neg^2_n and Pos^2_n of second order formulae as follows.

- If $A \in \mathcal{I}^2_k[\mathfrak{X}_k]$ for some $k \leq n$ then $A \in \mathrm{Neg}^2_n \cap \mathrm{Pos}^2_n$.
- If $A \in \mathcal{I}^2_n[\mathfrak{X}_n]$ then $\forall\mathfrak{X}_nA \in \mathrm{Pos}^2_n$.
- If $A \in \mathrm{Neg}^2_n$ then $\forall xA \in \mathrm{Neg}^2_n$.
- If $A \in \mathrm{Pos}^2_n$ then $\forall xA \in \mathrm{Pos}^2_n$.
- If $A \in \mathrm{Neg}^2_n$ and $B \in \mathrm{Pos}^2_n$ then $A \to B \in \mathrm{Pos}^2_n$.
- If $A \in \mathrm{Pos}^2_n$ and $B \in \mathrm{Neg}^2_n$ then $A \to B \in \mathrm{Neg}^2_n$.
- If $A, B \in \mathrm{Neg}^2_n$ then $A \wedge B \in \mathrm{Neg}^2_n$.
- If $A, B \in \mathrm{Pos}^2_n$ then $A \wedge B \in \mathrm{Pos}^2_n$.

Remark 1.5.2. $\mathrm{Neg}^2_n \subset \mathrm{Neg}^2_{n+1}$ and $\mathrm{Pos}^2_n \subset \mathrm{Pos}^2_{n+1}$.

Remark 1.5.3. $\mathbb{N}x \in \mathrm{Pos}^2_0$.

Informally, normal proofs are those where no elimination follows an introduction. Technically this can be expressed by allowing elimination rules only if the major premise is an elimination, assumption or axiom. In our formalization in Definition 1.5.4 this will read as $(d)_0 \leq 4$.

We also have to consider the complexity of the definition, that is to say, the number of inductive definitions needed. The (ω)-rule is technically convenient, but is an additional inductive definition. Hence, to get a notion that can be arithmetically defined we use "universal generalization" as a proof principle. For technical reasons we also have to use a form of (ω)-rule, which however does not make full induction admissible. As opposed to the usual (ω)-rule, used for example by Schütte [38], in our rule a typical term would be a new variable, rather than a numeral. A more careful discussion can be found in Remark 1.5.12.

Let $n^* \geq 1$ be a natural number. In this subsection we shall develop a notion of semiformal proofs definable and properties thereof provable in $\mathrm{ID}^*_{n^*-1}$. Nevertheless we will continue to state all our theorems as theorems of our meta theory PRA.

As we will later (for example in Lemma 1.6.18) need that certain transformations of proofs are recursive we will assign a natural number as proof object, rather than just defining a notion of provability.

Definition 1.5.4 (Normal derivations $d \vdash_n \Gamma \sqsupset A$). By iterated inductive definitions we define the relations $d \vdash_n \Gamma \sqsupset A$ for $n < n^*$ and d a natural number, $A \in \mathcal{I}^*_{n+1}$ a formula and $\Gamma \subset \mathcal{I}^*_{n+1}$ a finite set of formulae.

- (Axiom) $\langle 0 \rangle \vdash_n \Gamma \sqsupset A$ if $A \equiv B\sigma$, where B is an axiom in the sense of Definition 1.1.2. (In particular set of axioms allowed here is closed under substitution.) A discussion on how far the set of axioms can be extended will follow in Subsection 2.2.
- (Assumption) $\langle 1 \rangle \vdash_n \Gamma \sqsupset A$ if $A \in \Gamma$.
- ($\to$-elim) If $d \vdash_n \Gamma \sqsupset A \to B$ with $(d)_0 \leq 4$ and $e \vdash_n \Gamma \sqsupset A$ then $\left\langle 2, d, e, \ulcorner A \urcorner \right\rangle \vdash_n \Gamma \sqsupset B$.
- ($\wedge$-elim) If $d \vdash_n \Gamma \sqsupset A_0 \wedge A_1$ and $(d)_0 \leq 4$ then $\left\langle 3, d, i, \ulcorner A_{1-i} \urcorner \right\rangle \vdash_n \Gamma \sqsupset A_i$.
- ($\forall x$-elim) If $d \vdash_n \Gamma \sqsupset \forall x A(x)$ with $(d)_0 \leq 4$ then $\left\langle 4, d, \ulcorner \forall x A(x) \urcorner \right\rangle \vdash_n \Gamma \sqsupset A(t)$.
- ($\to$-intro) If $d \vdash_n \Gamma, A \sqsupset B$ then $\langle 5, d \rangle \vdash_n \Gamma \sqsupset A \to B$.
- ($\forall x$-intro) If $d \vdash_0 \Gamma' \sqsupset A(x)$ then $\left\langle 6, d, \ulcorner x \urcorner \right\rangle \vdash_0 \Gamma \sqsupset \forall z A(z)$ where Γ' is the subset of those formulae of Γ not containing x free.
- (ω-rule) If $n \geq 1$ and for all terms t it holds that $\{e\} \left(\ulcorner t \urcorner\right) \vdash_n \Gamma \sqsupset A(t)$ then $\langle 6, e \rangle \vdash_n \Gamma \sqsupset \forall x A(x)$.
- ($\wedge$-intro) If $d_0 \vdash_n \Gamma \sqsupset A_0$ and $d_1 \vdash_n \Gamma \sqsupset A_1$ then $\langle 7, d_0, d_1 \rangle \vdash_n \Gamma \sqsupset A_0 \wedge A_1$.
- ($\forall \mathfrak{X}$-intro) If $d \vdash_n \Gamma' \sqsupset A(\mathfrak{X})$ then $\left\langle 8, d, \ulcorner \mathfrak{X} \urcorner \right\rangle \vdash_n \Gamma \sqsupset \forall \mathfrak{X} A(\mathfrak{X})$ where Γ' is the subset of those formulae of Γ not containing $\mathfrak{X}$ free.
- (Ω_{k+1}) If $k < n$ and $d \vdash_n \Gamma \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$ and (for all $\Delta \subset \mathrm{Neg}^2_k$ it holds that if $d' \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$ then $\{e\} \left(\left\langle d', \ulcorner \Delta \urcorner \right\rangle\right) \vdash_n \Delta, \Gamma \sqsupset B$) then $\left\langle 9 + k, \ulcorner \forall \mathfrak{X}_k A(\mathfrak{X}_k) \urcorner, d, e \right\rangle \vdash_n \Gamma \sqsupset B$.

Remark 1.5.5. The reader is invited to verify that the above Definition 1.5.4 of the relations $\cdot \vdash_n \cdot \sqsupset \cdot$ for $n = 0, \ldots, n^* - 1$ can be formalized in $\mathrm{ID}^*_{n^*-1}$. In this formalization we would use appropriate Gödel codes instead of Γ and A. Recalling that we identify α-equal formulae also in our meta theory, the definition fits with our convention that α-equal formulae have identical Gödel codes.

So officially, for $1 \leq n < n^*$ we have a formula $\mathfrak{A}_n \in \mathrm{Pos}_{n-1}$, formalizing the underlying operator and a predicate $\mathcal{P}_n^{\mathfrak{A}_n}$ for the fixed point. We write $d \vdash_n \Gamma \sqsupset A$ for the official $\mathcal{P}_n^{\mathfrak{A}_n} \left\langle d, \ulcorner \Gamma \urcorner, \ulcorner A \urcorner \right\rangle$.

The relation $\cdot \vdash_0 \cdot \sqsupset \cdot$ is in fact a primitive recursive relation, which is formalized as such.

We have "saved" an inductive definition by inspecting only a particular branch at the first "ω-branching" level and so could obtain an arithmetical relation. This technique has a strong similarity to the one used in Arai's "slow growing analogue to Buchholz' proof" [3]. There "pointwise transfinite induction" for the ordinal notation system for the ordinal of ID^c_p was shown within ID^c_{p-1}. The crucial observation was that the lowest inductive definition becomes arithmetical when for an ordinal term of type ω only the n'th element of the fundamental sequence is considered, for a fixed but arbitrary n given from the outside. In our case, to show a property for all terms we are happy with a proof just for a single variable — provided it is sufficiently new.

A rationale why we still include an infinitely branching (ω)-rule can be found in Remark 1.5.12.

Remark 1.5.6. Note that in Definition 1.5.4 the "witness" d is constructed in such a way, that for given Γ and A such that $d \vdash_n \Gamma \sqsupset A$ holds, one can uniquely reconstruct all the formulae Γ', A' and relations $d' \vdash_k \Gamma' \sqsupset A'$ that led to the relation $d \vdash_n \Gamma \sqsupset A$ in the inductive definition. This will be used silently in the sequel.

These derivations are called normal since they have, with exception of the (Ω)-rule, the subformula property (and even for the (Ω)-rule we have the subformula property for the "important" premises). In particular we have

Lemma 1.5.7. *If Γ consists only of formulae ending in $\perp$, and for some atomic formula $R(\vec{t})$, different from $\perp$, we have $d \vdash_0 \Gamma \sqsupset R(\vec{t})$ then $R(\vec{t})$ holds.*

Proof. Inspection of the proof rules shows that $\vdash_0 \ \sqsupset$ -proofs of atomic formulae can only be eliminations (that is $(d)_0 \leq 4$). Since *no* axioms (compare Definition 1.1.2) have $\perp$ in a premise, and no assumptions end in an atom different from $\perp$, in fact *no* assumption was used at all. Hence all occurring formulae are in fact quantifier-free. So using Δ^0_0-truth predicates, induction on the derivation shows that it only proves true formulae. ∎

Remark 1.5.8. As a theorem of ID^*_n, if $d \vdash_k \Gamma \sqsupset A$ and $k \leq n$ then $d \vdash_n \Gamma \sqsupset A$.

Remark 1.5.9 (Weakening). We have chosen ($\forall x$-intro) and ($\forall \mathfrak{X}$-intro) in such a way that, provably in ID^*_k, it is the case that $d \vdash_k \Gamma \sqsupset A$ and $\Gamma \subset \Gamma'$ imply $d \vdash_k \Gamma' \sqsupset A$.

First order substitutions do not fit well with the (Ω)-rule. As proofs of level 0 do not contain (Ω)-rules, first order substitutions are admissible. More precisely we have

Lemma 1.5.10 (First order substitution). *There is a primitive recursive $\mathfrak{f}$ such that if $d \vdash_0 \Gamma \sqsupset A$ then $\mathfrak{f}(\ulcorner A \urcorner, \ulcorner \Gamma \urcorner, d, \ulcorner \sigma \urcorner) \vdash_0 \Gamma\sigma \sqsupset A\sigma$. If $(d)_0 \leq 4$ then $(\mathfrak{f}(\dots, d, \dots))_0 \leq 4$. Moreover, the number of inferences is* not *changed.*

Proof. Induction on $d \vdash_0 \Gamma \sqsupset A$. Note that (in Definition 1.5.4) we have required our set of axioms to be closed under substitution. In the case of ($\forall x$-intro) we have to do case distinctions on the variable that is introduced. Moreover, it might be necessary to conclude from $A(x')$ for a different x'.

Note that the rules (Ω) and (ω) cannot have been used.
All the remaining cases are immediate by induction hypothesis. □

Lemma 1.5.11 ($\forall x$-elimination). *There is a partially recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_n \Gamma \sqsupset \forall x A(x)$ then $\mathfrak{f}(d, \ulcorner\Gamma\urcorner, \ulcorner\forall x A\urcorner, \ulcorner t\urcorner) \vdash_n \Gamma \sqsupset A(t)$.*

Proof. Induction on $d \vdash_n \Gamma \sqsupset \forall x A(x)$. If $(d)_0 \leq 4$ then $\langle 4, d, \ulcorner\forall x A\urcorner\rangle$ is the desired derivation; if $d = \langle 6, d', \ulcorner x\urcorner\rangle$ apply Lemma 1.5.10 to d'; if $d = \langle 6, e\rangle$ then $\{e\}\left(\ulcorner t\urcorner\right)$ is the desired derivation and in the case of (Ω_{k+1}) apply the induction hypothesis. □

Remark 1.5.12. Considering the proofs of Lemma 1.5.10 and 1.5.11 we can explain, why we had to use the technical trick of using both, a ($\forall x$-intro) and an (ω)-rule. The first is necessary to have $\cdot \vdash_0 \cdot \sqsupset \cdot$ an arithmetical relation. However, later we will (in Definition 1.6.1) define (partial) truth predicates and have to show (in Lemma 1.6.9) that everything derivable in $\cdot \vdash_n \cdot \sqsupset \cdot$ is true. In particular, we have to show that from a proof of a $\forall x A$ statement we get the truth of $A(t)$ for *all* terms t. In the case of a universal statement introduced by ($\forall x$-intro), soundness is ensured by admissibility of first order substitution. But such a proof fails at an (Ω)-rule: to conclude $\ldots \vdash_n \Gamma\sigma \sqsupset B\sigma$ by an (Ω)-rule again we would have to provide (in a uniform way) a proof of $\ldots \vdash_n \Delta, \Gamma\sigma \sqsupset B\sigma$. However, the induction hypothesis would only give a proof of $\ldots \vdash_n \Delta\sigma, \Gamma\sigma \sqsupset B\sigma$. This should also be compared to the careful formulation of Lemma 1.5.14.

Lemma 1.5.13. *As a theorem of ID^*_n, if $d \vdash_n \Gamma \sqsupset A$ and $(d)_0 \leq 4$ then A is subformula of an axiom or subformula of an element of Γ.*

Proof. Induction on d. □

Lemma 1.5.14 (Substitution for second order variables). *There is a primitive recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_n \Gamma \sqsupset A$ then $\mathfrak{f}(\ulcorner\Gamma\urcorner, \ulcorner A\urcorner, \ulcorner\Sigma\urcorner, \ulcorner\Sigma'\urcorner, d) \vdash_n \Gamma\Sigma'\Sigma \sqsupset A\Sigma'\Sigma$ where Σ is a substitution of second order variables with only variables of level n in its domain such that $\Gamma\Sigma$ and $A\Sigma$ are well-formed, and Σ' is a permutation of second order variables of level less than n.*

Proof. Induction along the inductive Definition 1.5.4 of $d \vdash_n \Gamma \sqsupset A$.

The case (Axiom) is trivial, since axioms do not contain second order variables. The case (Assumption) is trivial, too. The other cases with $(d)_0 \leq 4$ are immediate by induction hypothesis. (Here we simultaneously prove that $(\mathfrak{f}(\ldots, d))_0 \leq 4$ if $(d)_0 \leq 4$ and use Lemma 1.5.13.) So are the cases with $(d)_0 \in \{5, 6, 7\}$, using Lemma 1.5.10 for ($\forall x$-intro).

For ($\forall\mathfrak{X}$-intro) with the abstracted variable $\mathfrak{X}$ of level less than n let $\mathfrak{X}'$ be a new variable of the same level as $\mathfrak{X}$. Let $\Sigma'' = [\mathfrak{X} \mapsto \mathfrak{X}', \mathfrak{X}' \mapsto \mathfrak{X}]\Sigma'$ be the composition of permutations. We note that $\mathfrak{X}$ is not in the free variables of Γ. Hence the induction hypothesis yields $\mathfrak{f}(\ulcorner\Gamma\urcorner, \ulcorner A(\mathfrak{X})\urcorner, \ulcorner\Sigma\urcorner, \ulcorner\Sigma''\urcorner, (d)_1) \vdash_n \Gamma\Sigma'\Sigma \sqsupset A(\mathfrak{X}')$. So ($\forall\mathfrak{X}'$-intro) can be applied again.

For ($\forall\mathfrak{X}$-intro) with the abstracted variable $\mathfrak{X}$ of level n let $\mathfrak{X}'$ be a new second order variable of level n. Apply the induction hypothesis for Σ extended by $\mathfrak{X} \mapsto \mathfrak{X}'$ and conclude by ($\forall\mathfrak{X}$-intro) again, with $\mathfrak{X}'$ as abstracted variable.

For the case (Ω_{k+1}) we have by induction hypothesis $\ldots \vdash_n \Gamma\Sigma'\Sigma \sqsupset \forall\mathfrak{X}_k A$. For all $\Delta \subset \mathrm{Neg}^2_k$ and $d' \vdash_k \Delta \sqsupset \forall\mathfrak{X}_k A$ we have by induction hypothesis $\ldots \vdash_k \Delta\bar{\Sigma}' \sqsupset \forall\mathfrak{X}_k A$ where $\bar{\Sigma}'$ is the inverse of Σ'. Hence, by the premise of the (Ω_{k+1})-rule we have $\ldots \vdash_n \Delta\bar{\Sigma}', \Gamma \sqsupset B$. Again by induction hypothesis we obtain $\ldots \vdash_n \Delta, \Gamma\Sigma'\Sigma \sqsupset B\Sigma'\Sigma$, noting that $\Delta\Sigma = \Delta$. Therefore an application of the (Ω_{k+1})-rule again completes the derivation.

All the remaining cases are immediate by induction hypothesis, where in the case of ($\forall x$-intro) we need a renaming argument, provided by Lemma 1.5.10. □

The following lemma explains the usefulness of the (Ω)-rule: derivations of $\forall\mathfrak{X}_n A(\mathfrak{X}_n)$ contain enough information to construct a derivation of $A(\mathcal{A})$. Hence, the (Ω)-rule can be used (as will be done in Lemma 1.6.10) to construct derivations corresponding to $\forall\mathfrak{X}$-elimination.

Note that the proof of the following lemma does *not* introduce new (Ω)-inferences. The inferences are instead only reconstructed where they occur in the given semi-formal derivation. However, the lemma only considers quantification at the *topmost* level and this is important, for otherwise the restriction $\Gamma \subset \mathrm{Neg}^2_n$ would not suffice. Nevertheless, in applications of the (Ω)-rule the contexts are of sufficiently small level. So we will use the lemma for various values n strictly smaller than n^*.

Lemma 1.5.15 (Admissibility of $\forall\mathfrak{X}_n$-elimination). *There is a primitive recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_n \Gamma \sqsupset \forall\mathfrak{X}_n A(\mathfrak{X}_n)$ with $\Gamma \subset \mathrm{Neg}^2_n$ and $\mathcal{A} \in \mathcal{I}^*_n$ such that $A(\mathcal{A})$ is a well-formed formula, then $\mathfrak{f}(\ulcorner\forall\mathfrak{X}_n A(\mathfrak{X}_n)\urcorner, \ulcorner\mathcal{A}\urcorner, \ulcorner\Gamma\urcorner, d) \vdash_n \Gamma \sqsupset A(\mathcal{A})$.*

Proof. Induction along the inductive Definition 1.5.4 of $d \vdash_n \Gamma \sqsupset \forall\mathfrak{X}_n A(\mathfrak{X}_n)$.

$(d)_0 \leq 4$ would require an axiom or assumption containing $\forall\mathfrak{X}_n A(\mathfrak{X}_n)$ strictly positive, which cannot be as our axioms are first order and $\Gamma \subset \mathrm{Neg}^2_n$.

If the last rule was ($\forall\mathfrak{X}_n$-intro) then Lemma 1.5.14 applies. By the form of the conclusion the only other inference the derivation could end in is (Ω_{k+1}). So assume that $\ldots \vdash_n \Gamma \sqsupset \forall\mathfrak{X}_n A(\mathfrak{X}_n)$ was concluded from: $d \vdash_n \Gamma \sqsupset \forall\mathfrak{X}_k B(\mathfrak{X}_k)$ for some $k < n$ and from the fact that for all $\Delta \subset \mathrm{Neg}^2_k$, if $d' \vdash_k \Delta \sqsupset \forall\mathfrak{X}_k B(\mathfrak{X}_k)$ then $\{e\}(d') \vdash_n \Delta, \Gamma \sqsupset \forall\mathfrak{X}_n A(\mathfrak{X}_n)$. We have for every $\Delta \subset \mathrm{Neg}^2_k \subset \mathrm{Neg}^2_n$ and d' such that $d' \vdash_k \Delta \sqsupset \forall\mathfrak{X}_k B$ by induction hypothesis $\mathfrak{f}(\ldots, \{e\}(\langle d', \ulcorner\Delta\urcorner\rangle)) \vdash_n \Delta, \Gamma \sqsupset A(\mathcal{A})$. Hence an application of (Ω_{k+1}) yields the desired derivation. □

Lemma 1.5.16 (Collapsing). *There is a partially recursive function $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_n \Gamma \sqsupset A$ with $\Gamma \subset \mathrm{Neg}^2_\ell$ and $A \in \mathrm{Pos}^2_\ell$ for some $\ell < n$ then $\mathfrak{f}(\ell, d, \ulcorner\Gamma\urcorner, \ulcorner A\urcorner) \vdash_\ell \Gamma \sqsupset A$.*

Proof. Induction along the inductive Definition 1.5.4 of $d \vdash_n \Gamma \sqsupset A$, showing the claim simultaneously for all $\ell < n$.

If $(d)_0 \leq 4$ then also $d \vdash_\ell \Gamma \sqsupset A$.

The cases ($\to$-intro), ($\forall x$-intro), ($\wedge$-intro) and ($\forall\mathfrak{X}$-intro) are immediate by the induction hypotheses, where in the case ($\to$-intro) we use that $A \in \mathrm{Pos}^2_\ell$.

If $\ell = 0$ and $d = \langle 6, e\rangle$, that is, in the case of an (ω)-rule that has to be collapsed to a proof without, we apply the induction hypotheses to $\{e\}(\ulcorner y\urcorner)$ for some new y and conclude by ($\forall x$-intro). The case of the (ω)-rule with $\ell > 0$ is immediate by induction hypothesis.

In the case of the (Ω_{k+1})-rule with $\ell \leq k$ we have $\mathfrak{f}(k, d, \ldots) \vdash_k \Gamma \sqsupset \forall\mathfrak{X}_k A$ by induction hypothesis. Since $\ell \leq k$ and hence $\Gamma \subset \mathrm{Neg}^2_\ell \subset \mathrm{Neg}^2_k$ by the premise of the rule we get $\{e\}\left(\langle \mathfrak{f}(k,d,\dots), \ulcorner\Gamma\urcorner\rangle\right) \vdash_n \Gamma \sqsupset B$. By application of the induction hypothesis(!) we get $\mathfrak{f}(\ell, \{e\}(\dots), \ulcorner\Gamma\urcorner) \vdash_\ell \Gamma \sqsupset B$.

In the case of the (Ω_{k+1})-rule with $\ell > k$ we have by induction hypothesis $\mathfrak{f}(\ell, d, \dots) \vdash_\ell \Gamma \sqsupset \forall\mathfrak{X}_k A$ and $\mathfrak{f}(\ell, \{e\}\left(\langle d', \ulcorner\Delta\urcorner\rangle\right), \dots) \vdash_\ell \Delta, \Gamma \sqsupset B$ if $\Delta \subset \mathrm{Neg}^2_k$ and $d' \vdash_k \Delta \sqsupset \forall\mathfrak{X}_k A(\mathfrak{X}_k)$. Hence an application of the (Ω_{k+1})-rule yields the desired derivation. ∎

1.6 Forcing Relation for $\mathrm{HA}^2_{n^*}$ and Cut Elimination

With our semi-formal notion of normal proofs we have an appropriate semantics for $\forall\mathfrak{X}$-statements. Based on this semantics we define (partial) truth predicates in the usual way, similar to Tait's computability predicates [39].

To be proof theoretically optimal, these predicates are defined on the meta-level, by a family of formulae. Of course, an additional inductive definition would suffice to define these predicates internally, requiring, however, a stronger system. This increase by an additional inductive definition catches up at the level $\mathrm{ID}^c_{<\omega}$. So a reader only interested in $\mathrm{ID}^c_{<\omega}$, for example mainly in the results discussed in Section 4, can think of the forcing relation (introduced in Definition 1.6.1) as defined by an inductive definition and hence read the following lemmata as internal to $\mathrm{ID}^*_{n^*}$, rather than as the observation that there are families of proofs of $\mathrm{ID}^*_{n^*-1}$.

Again, let $n^* \geq 1$ be a natural number. In this subsection we will develop a proof of cut elimination which can be formalized locally in $\mathrm{ID}^*_{n^*-1}$. We continue the habit of stating our theorems as theorems of PRA.

Definition 1.6.1 ($f \Vdash \Gamma \sqsupset C$). By induction on $\mathfrak{k}$ we define a family $f \Vdash_{\mathfrak{k}} \Gamma \sqsupset C$ of formulae in the language of $\mathrm{ID}^*_{n^*-1}$ such that for all $C \in \mathcal{I}^*_{n^*}$ with less than $\mathfrak{k}$ logical symbols, all $f \in \mathbb{N}$ and $\Gamma \subset \mathrm{Neg}^2_{n^*-1}$ the following properties hold.

- $f \Vdash_{\mathfrak{k}} \Gamma \sqsupset C \quad\Leftrightarrow\quad f \vdash_{n^*-1} \Gamma \sqsupset C$
  if C atomic or of the form $\forall\mathfrak{X} A$
- $f \Vdash_{\mathfrak{k}} \Gamma \sqsupset A \to B \quad\Leftrightarrow$
  $\forall g \in \mathbb{N}, \Delta \subset \mathrm{Neg}^2_{n^*-1}(g \Vdash_{\mathfrak{k}-1} \Delta \sqsupset A \Rightarrow \{f\}\left(\langle g, \ulcorner\Delta\urcorner\rangle\right) \Vdash_{\mathfrak{k}-1} \Gamma, \Delta \sqsupset B)$
- $f \Vdash_{\mathfrak{k}} \Gamma \sqsupset \forall x A(x) \quad\Leftrightarrow\quad \forall t(\{f\}\left(\ulcorner t\urcorner\right) \Vdash_{\mathfrak{k}-1} \Gamma \sqsupset A(t))$
- $f \Vdash_{\mathfrak{k}-1} \Gamma \sqsupset A \wedge B \quad\Leftrightarrow\quad (f)_0 \Vdash_{\mathfrak{k}} \Gamma \sqsupset A$ and $(f)_1 \Vdash_{\mathfrak{k}-1} \Gamma \sqsupset B$

Remark 1.6.2 (On the formalization of the relation $\cdot \Vdash \cdot \sqsupset \cdot$). Definition 1.6.1 could be defined by induction on C. However, we need slightly more uniformity. So, officially, we define a family $A_{\mathfrak{k}}(x,y,z)$ of formulae indexed by $\mathfrak{k}$, with free number variables for f, $\ulcorner\Gamma\urcorner$ and $\ulcorner C\urcorner$ in such a way that whenever the number of logical symbols in C is at most $\mathfrak{k}$ then $A_{\mathfrak{k}}(f, \ulcorner\Gamma\urcorner, \ulcorner C\urcorner)$ is equivalent to $f \Vdash \Gamma \sqsupset C$ as defined (say) by induction on C. This can easily be achieved by setting (with the $\mathfrak{A}_{n^*-1}$ of Remark 1.5.5)

$$
\begin{aligned}
A_{\mathfrak{k}}(x,y,z) :\equiv & (\text{“}z \text{ codes a formula which is atomic or of the form } \forall\mathfrak{X}B\text{”} \\
& \quad \wedge \quad \mathcal{P}^{\mathfrak{A}_{n^*-1}}_{n^*-1}\langle x,y,z\rangle) \\
& \vee (\text{“}x \text{ codes a formula of the form } B \to C\text{”} \\
& \quad \wedge \quad \forall g \forall \Delta \subset \mathrm{Neg}_{n^*-1}.\ A_{\mathfrak{k}-1}(g, \ulcorner\Delta\urcorner, \ulcorner B\urcorner) \\
& \qquad\qquad\qquad\qquad \to A_{\mathfrak{k}-1}(\{f\}\left(\langle g, \ulcorner\Delta\urcorner\rangle\right), \ulcorner\Gamma,\Delta\urcorner, \ulcorner C\urcorner)) \\
& \vee \ldots
\end{aligned}
$$

It should also be noted that for $\mathfrak{k}' \geq \mathfrak{k}$ it is provable in $\mathrm{ID}^*_{n^*-1}$, that if z codes a formula with at most $\mathfrak{k}$ logical symbols then $A_{\mathfrak{k}'}(x,y,z)$ and $A_{\mathfrak{k}}(x,y,z))$ are equivalent. This can be shown by induction on $\mathfrak{k}$. From now on we will write $f \Vdash \Gamma \sqsupset C$ for $A_{\mathfrak{k}}(f, \ulcorner\Gamma\urcorner, \ulcorner C\urcorner)$ with appropriately chosen $\mathfrak{k}$, silently assuming $\mathfrak{k}$ to be big enough. The reader should verify that this reading yields a primitive recursive family of proofs for the following theorems.

Remark 1.6.3. Note that in the case of an implication our forcing relation only has to deal with proofs from negative assumptions. This will be crucial when we will show the admissibility of the (Ω)-rule in Lemma 1.6.8.

Remark 1.6.4. As weakening is admissible for derivations (Remark 1.5.9), an easy induction shows that it is also admissible for the forcing relation. More precisely, provably in $\mathrm{ID}^*_{n^*-1}$, if $f \Vdash \Gamma \sqsupset A$, $\Gamma \subset \Gamma'$ then $f \Vdash \Gamma' \sqsupset A$.

Lemma 1.6.5 (Renaming of second order variables). *For every C and level-preserving permutation Σ of second order variables there are partially recursive functions $\mathfrak{f}^C_\Sigma$ such that the dependency of $\mathfrak{f}^C_\Sigma$ from $\ulcorner C\urcorner$ and $\ulcorner\Sigma\urcorner$ is primitive recursive. Moreover, for every natural number $\mathfrak{k}$ the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. If C has at most $\mathfrak{k}$ logical symbols and $f \Vdash \Gamma \sqsupset C$ then $\mathfrak{f}^C_\Sigma(f) \Vdash \Gamma\Sigma \sqsupset C\Sigma$.*

Proof. By induction on the logical complexity of C. Let $\bar{\Sigma}$ be the inverse of Σ.

If C is atomic or of the form $\forall\mathfrak{X}A$ then Lemma 1.5.14 applies.

If $C = A \to B$ and $g \Vdash \Delta \sqsupset A\Sigma$ then, by induction hypothesis $\mathfrak{f}^B_{\bar{\Sigma}}(g) \Vdash \Delta\bar{\Sigma} \sqsupset A$. Hence, $\{f\}\left(\mathfrak{f}^B_{\bar{\Sigma}}(g)\right) \Vdash \Gamma, \Delta\bar{\Sigma} \sqsupset B$ and by induction hypothesis, $\mathfrak{f}^A_\Sigma(\{f\}\left(\mathfrak{f}^B_{\bar{\Sigma}}(g)\right)) \Vdash \Gamma\Sigma, \Delta \sqsupset B\Sigma$.

The case $A \wedge B$ is immediate by the induction hypotheses and so is the case $\forall x A$. ∎

Lemma 1.6.6. *For every formula C there are partially recursive functions $\mathfrak{f}^C_\Gamma$ and $\mathfrak{g}^C_\Gamma$ such that the dependency of $\mathfrak{f}^C_\Gamma$ and $\mathfrak{g}^C_\Gamma$ from $\ulcorner C\urcorner$ is primitive recursive. Moreover, for every natural number $\mathfrak{k}$ the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. If C has at most $\mathfrak{k}$ logical symbols, then*

- $e \vdash_{n^*-1} \Gamma \sqsupset C$, $(e)_0 \leq 4$, $C \in \mathrm{Neg}^2_{n^*-1} \quad \Rightarrow \quad \mathfrak{f}^C_\Gamma(e) \Vdash \Gamma \sqsupset C$,
- $f \Vdash \Gamma \sqsupset C$, $C \in \mathrm{Pos}^2_{n^*-1} \quad \Rightarrow \quad \mathfrak{g}^C_\Gamma(f) \vdash_{n^*-1} \Gamma \sqsupset C$.

Proof. Induction on the logical complexity of C. Trivial, if C is atomic or of the form $\forall\mathfrak{X}A$.

If C is of the form $A \to B$ we argue as follows.

- We have $A \in \mathrm{Pos}^2_{n^*-1}$ and $B \in \mathrm{Neg}^2_{n^*-1}$. Assume $g \Vdash \Delta \sqsupset A$. By induction hypothesis $\mathfrak{g}^A_\Delta(g) \vdash_{n^*-1} \Delta \sqsupset A$. Since $(e)_0 \leq 4$ by assumption, we get $\langle 2, e, \mathfrak{g}^A_\Delta(g), \ulcorner A \urcorner \rangle \vdash_{n^*-1} \Gamma, \Delta \sqsupset B$. Hence by induction hypothesis $\mathfrak{f}^B_{\Gamma,\Delta}(\langle 2, e, \mathfrak{g}^A_\Delta(g), \ulcorner A \urcorner \rangle) \Vdash \Gamma, \Delta \sqsupset B$.

- We have $A \in \mathrm{Neg}^2_{n^*-1}$ and $B \in \mathrm{Pos}^2_{n^*-1}$. By definition we have $\langle 1 \rangle \vdash_{n^*-1} A \sqsupset A$. Hence by induction hypothesis $\mathfrak{f}^A_A(\langle 1 \rangle) \Vdash A \sqsupset A$. Thus $\{f\}\left(\langle \mathfrak{f}^A_A(\langle 1 \rangle), \ulcorner A \urcorner \rangle\right) \Vdash \Gamma, A \sqsupset B$. So $\mathfrak{g}^B_{A,\Gamma}(\{f\}(\ldots)) \vdash_{n^*-1} \Gamma, A \sqsupset B$ by induction hypothesis. We conclude by ($\to$-intro).

If C is of the form $\forall x A(x)$ we argue in the following way.

- Since $(d)_0 \leq 4$, we have $\langle 4, e, \ulcorner \forall x A(x) \urcorner \rangle \vdash_{n^*-1} \Gamma \sqsupset A(t)$, hence $\mathfrak{f}^{A(t)}_\Gamma(\langle 4, e, \ulcorner \forall x A(x) \urcorner \rangle) \Vdash \Gamma \sqsupset A(t)$.

- We have for every term t that $\{f\}\left(\ulcorner t \urcorner\right) \Vdash \Gamma \sqsupset A(t)$, hence by induction hypothesis $\mathfrak{g}^{A(t)}_\Gamma(\{f\}\left(\ulcorner t \urcorner\right)) \vdash_n \Gamma \sqsupset A(t)$. The ($\omega$)-rule applies. (If $n^* = 1$ then $\mathfrak{g}^{A(y)}_\Gamma(\{f\}\left(\ulcorner y \urcorner\right)) \vdash_0 \Gamma \sqsupset A(y)$ for a new variable y, and we apply ($\forall x$)-intro.)

If C is of the form $A_0 \wedge A_1$ we are done immediately by the induction hypotheses. ∎

Remark 1.6.7. The $\mathfrak{f}$ and $\mathfrak{g}$ in the above lemma correspond to the $\uparrow$ and $\downarrow$ of "Normalization by Evaluation" [8].

Lemma 1.6.8 (Admissibility of (Ω_{k+1}) for $k \leq n^*-1$). *For every formula C and set Γ of formulae there is a partially recursive $\mathfrak{f}^C_\Gamma$ such that the dependency of $\mathfrak{f}^C_\Gamma$ from $\ulcorner C \urcorner$ and $\ulcorner \Gamma \urcorner$ is primitive recursive. Moreover, for every bound on the logical complexity of C the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. Assume for $\Gamma \subset \mathrm{Neg}^2_{n^*-1}$ that $d \Vdash \Gamma \sqsupset \forall \mathfrak{X}_k A$ and $k \leq n^* - 1$. Also assume that for all $\Delta \subset \mathrm{Neg}^2_k$ if $d' \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A$ then $\{e\}\left(\langle d', \ulcorner \Delta \urcorner \rangle\right) \Vdash \Delta, \Gamma \sqsupset C$. Then $\mathfrak{f}^C_\Gamma(\langle d, e \rangle) \Vdash \Gamma \sqsupset C$.*

Proof. Induction on the logical complexity of C.

Case C is atomic or of the form $\forall \mathfrak{X} B$ and so we have to show $\ldots \vdash_{n^*-1} \Gamma \sqsupset C$. If $k < n^* - 1$ we may apply and (Ω_{k+1})-inference, using $d \vdash_{n^*-1} \Gamma \sqsupset \forall \mathfrak{X}_k A$ and $\{e\}\left(\langle d', \ulcorner \Delta \urcorner \rangle\right) \vdash_{n^*-1} \Delta, \Gamma \sqsupset C$. If $k = n^* - 1$ then $\{e\}\left(\langle d, \ulcorner \Gamma \urcorner \rangle\right) \Vdash \Gamma \sqsupset C$ by assumption.

Case $C = B \to B'$. Let $\Gamma' \subset \mathrm{Neg}^2_{n^*-1}$ and assume $g \Vdash \Gamma' \sqsupset B$. We have $d \Vdash \Gamma, \Gamma' \sqsupset \forall \mathfrak{X}_k A$. Moreover, for any $\Delta \subset \mathrm{Neg}^2_k$, $d' \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A$ we have $\{e\}\left(\langle d', \ulcorner \Delta \urcorner \rangle\right) \Vdash \Gamma, \Delta \sqsupset B \to B'$ by assumption and hence $\left\{\{e\}\left(\langle d', \ulcorner \Delta \urcorner \rangle\right)\right\}\left(\langle g, \ulcorner \Gamma' \urcorner \rangle\right) \Vdash \Gamma, \Delta, \Gamma' \sqsupset B'$. So by induction hypothesis (using $\Gamma, \Gamma' \subset \mathrm{Neg}^2_{n^*-1}$) we get $\ldots \Vdash \Gamma, \Gamma' \sqsupset B'$.

Case $C = \forall x B(x)$. The claim follows by induction hypothesis since $\{\{e\}(d')\}\left(\ulcorner t \urcorner\right) \Vdash \Delta, \Gamma \sqsupset B(t)$ for $\Delta \subset \mathrm{Neg}^2_{n^*-1}$ and $d' \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A$.

Case $C = B_0 \wedge B_1$. The claim follows by induction hypothesis since $(\{e\}(d'))_i \Vdash \Delta, \Gamma \sqsupset B_i$ for $\Delta \subset \mathrm{Neg}^2_{n^*-1}$ and $d' \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A$. ∎

Lemma 1.6.9 (Forcing under substitution). *There is a partially recursive $\mathfrak{f}$ such that for every natural number k the following is a theorem of $\mathrm{ID}^*_{n^*-1}$.*

Assume that each of the formulae $A, \vec{A}$ *has at most* k *symbols. Assume moreover that* $d \vdash_{n^*-1} \Delta, \vec{A} \sqsupset A$ *and* $f_i \Vdash \Gamma \sqsupset A_i$ *for* $i = 1, \ldots, \ell$ *and some* $\Delta, \Gamma \subset \mathrm{Neg}^2_{n^*-1}$. *Then* $\mathfrak{f}(\langle d, \langle f_1, \ulcorner A_1 \urcorner \rangle, \ldots, \langle f_n, \ulcorner A_n \urcorner \rangle\rangle) \Vdash \Delta, \Gamma \sqsupset A$.

Proof. Within $\mathrm{ID}^*_{n^*-1}$ we argue by the induction principle provided by the inductive Definition 1.5.4 of $\cdot \vdash_n \cdot \sqsupset \cdot$, where in the case of $\cdot \vdash_0 \cdot \sqsupset \cdot$ we actually use induction on the number of inferences.

Case (Axiom). By Lemma 1.6.6, noting that axioms are in $\mathcal{I}^2_0 \subset \mathrm{Neg}^2_0 \subset \mathrm{Neg}^2_{n^*-1}$.

Case (Assumption). By the assumption $f_i \Vdash \Gamma \sqsupset A$ of this lemma.

Case (→-elim). By induction hypothesis $e \Vdash \Delta, \Gamma \sqsupset A \to B$ and $e' \Vdash \Delta, \Gamma \sqsupset A$ for appropriate e and e'. Hence $\{e'\}(e) \Vdash \Delta, \Gamma \sqsupset B$ follows from the definition of the forcing relation.

Cases (∀x-elim), (∧-elim). Similar.

Case (→-intro). By induction hypothesis we know that $f \Vdash \Gamma' \sqsupset A$ implies $\mathfrak{f}(\langle d, \langle f_1, \ulcorner A_1 \urcorner \rangle, \ldots, \langle f_n, \ulcorner A_n \urcorner \rangle, \langle f, \ulcorner A \urcorner \rangle\rangle) \Vdash \Delta, \Gamma, \Gamma' \sqsupset B$. Hence $e \Vdash \Delta, \Gamma \sqsupset A \to B$ by Definition 1.6.1 of the forcing relation, where e is an index of the recursive function $f \mapsto \mathfrak{f}\left(\langle d, \langle f_1, \ulcorner A_1 \urcorner \rangle, \ldots, \langle f_n, \ulcorner A_n \urcorner \rangle, \langle f, \ulcorner A \urcorner \rangle\rangle\right)$.

Case (∀x-intro). By Lemma 1.5.11 we find an index e such that for every term t we have $\{e\}\left(\ulcorner t \urcorner\right) \vdash_0 \Delta, \vec{A} \sqsupset A(t)$ with a derivation with no more inferences. Hence by induction hypothesis we get the desired forcings in a recursive way for every t.

Cases (ω), (∧-intro). Immediate by the induction hypotheses.

Case (∀$\mathfrak{X}$-intro). We have $d \vdash_n \Delta, \vec{A} \sqsupset A(\mathfrak{X})$ and $\mathfrak{X}$ not free in $\Delta, \vec{A}$. Let $\mathfrak{Y}$ be a new second order variable and $\Sigma = [\mathfrak{X} \mapsto \mathfrak{Y}, \mathfrak{Y} \mapsto \mathfrak{X}]$. Noting that $\mathfrak{X}$ does not occur in A_i, Lemma 1.6.5 yields $\mathfrak{f}^{A_i}_{\Sigma}(f_i) \Vdash \Gamma\Sigma \sqsupset A_i$. Hence, by induction hypothesis $\ldots \Vdash \Gamma\Sigma, \Delta \sqsupset A(\mathfrak{X})$, and by Lemma 1.6.6 we get $\ldots \vdash_{n^*-1} \Gamma\Sigma, \Delta \sqsupset A(\mathfrak{X})$. By (∀$\mathfrak{X}$-intro) we obtain $\ldots \vdash_{n^*-1} \Gamma\Sigma, \Delta \sqsupset \forall \mathfrak{X} A(\mathfrak{X})$ and hence by Lemma 1.5.14, since $\mathfrak{X}$ is not free in Δ and $\mathfrak{Y}$ is new, $\ldots \vdash_{n^*-1} \Gamma, \Delta \sqsupset \forall \mathfrak{X} A(\mathfrak{X})$. This is what we had to show.

Case (Ω_{k+1}). We have $k < n^* - 1$ and $d \vdash_{n^*-1} \Delta, \vec{A} \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$. Moreover, for all $\Delta' \subset \mathrm{Neg}^2_k$ it holds that if $d' \vdash_k \Delta' \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$ then $\{e\}\left(\langle d', \ulcorner \Delta' \urcorner \rangle\right) \vdash_{n^*-1} \Delta', \Delta, \vec{A} \sqsupset B$.

By induction hypothesis we have $\ldots \Vdash \Delta, \Gamma \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$. Moreover, also by induction hypothesis we know for every $\Delta' \subset \mathrm{Neg}^2_k$ that if $d \vdash_k \Delta' \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$ then $\ldots \Vdash \Delta', \Delta, \Gamma \sqsupset B$. Using the assumption that $\Delta, \Gamma \subset \mathrm{Neg}^2_{n^*-1}$ we can apply Lemma 1.6.8 to obtain $\ldots \Vdash \Delta, \Gamma \sqsupset B$. □

Lemma 1.6.10 (Admissibility of (∀$\mathfrak{X}$-elimination) in the forcing relation). *There is a partially recursive* $\mathfrak{f}$ *such that for all* A *and for every* $\mathcal{A} \in \mathcal{I}^*_{n^*}$, *such that* $A(\mathcal{A})$ *is well formed, the following is a theorem of* $\mathrm{ID}^*_{n^*-1}$. *If* $\Gamma \subset \mathrm{Neg}^2_{n^*-1}$ *and* $f \Vdash \Gamma \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$ *then* $\mathfrak{f}(f, \ulcorner \Gamma \urcorner, \ulcorner A \urcorner, \ulcorner \mathfrak{X}_k \urcorner, \ulcorner \mathcal{A} \urcorner) \Vdash \Gamma \sqsupset A(\mathcal{A})$.

Proof. If $k = n^* - 1$ then we argue as follows: $f \vdash_{n^*-1} \Gamma \sqsupset \forall \mathfrak{X}_{n^*-1} A(\mathfrak{X}_{n^*-1})$ by assumption, hence $\ldots \vdash_{n^*-1} \Gamma \sqsupset A(\mathcal{A})$ by Lemma 1.5.15, and $\ldots \Vdash \Gamma \sqsupset A(\mathcal{A})$ by Lemma 1.6.9.

If $k < n^* - 1$ we use the (Ω_{k+1})-rule to construct a derivation of $A(\mathcal{A})$. From this derivation we get a forcing of $A(\mathcal{A})$ by Lemma 1.6.9. We have $f \vdash_{n^*-1} \Gamma \sqsupset \forall \mathfrak{X}_k A$ by assumption. So assume $\Delta \subset \mathrm{Neg}^2_k$ and $d \vdash_k \Delta \sqsupset \forall \mathfrak{X}_k A$. Then by Lemma 1.5.15 we have $\ldots \vdash_k \Delta \sqsupset A(\mathcal{A})$, and therefore also $\ldots \vdash_{n^*-1} \Delta, \Gamma \sqsupset A(\mathcal{A})$. □

Lemma 1.6.11 (Admissibility of ($\forall\mathfrak{X}$-introduction) for the forcing relation). *There is a partially recursive $\mathfrak{f}$ such that for every formula A the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. If $f \Vdash \Gamma \sqsupset A(\mathfrak{X}_k)$ for some $k \le n^*-1$ and $\Gamma \subset \mathrm{Neg}^2_{n^*-1}$, such that $\mathfrak{X}_k$ is not free in Γ, then $\mathfrak{f}(f, \ulcorner \mathfrak{X}_k \urcorner, \ulcorner A \urcorner, \ulcorner \Gamma \urcorner) \Vdash \Gamma \sqsupset \forall \mathfrak{X}_k A(\mathfrak{X}_k)$.*

Proof. We have $A(\mathfrak{X}_k) \in \mathrm{Pos}^2_{n^*-1}$. Hence by Lemma 1.6.6 we have $\ldots \vdash_{n^*-1} \Gamma \sqsupset A(\mathfrak{X}_k)$, so by an application of ($\forall\mathfrak{X}$-intro) we get a derivation of $\forall \mathfrak{X}_k A(\mathfrak{X}_k)$, hence a forcing of that formula. ☐

Theorem 1.6.12 (Cut elimination). *For every proof in $\mathrm{HA}^2_{n^*}$ of A from assumptions $\vec{A}$ there is a partially recursive $\mathfrak{f}$ such that the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. For every σ, if $f_i \Vdash \Gamma\sigma \sqsupset A_i\sigma$ for all i then $\mathfrak{f}(\ulcorner \sigma \urcorner, \vec{f}) \Vdash \Gamma\sigma \sqsupset A\sigma$.*

Proof. Induction on the $\mathrm{HA}^2_{n^*}$ derivation.

Case (Assumption). By the assumption $f_i \Vdash \Gamma \sqsupset A_i$.

Case (Axiom). By Lemma 1.6.6, using that our axioms are closed under substitution.

Cases of introduction or elimination of propositional connectives, ($\forall x$-elimination). By logic, induction hypothesis and the Definition 1.6.1 of the forcing relation.

Case ($\forall x$-intro). By induction hypothesis.

Cases of introduction and elimination of second order quantifiers. By Lemmata 1.6.11 and 1.6.10. ☐

Corollary 1.6.13 (Relative consistency of $\mathrm{HA}^2_{n^*}$). *If $\mathrm{HA}^2_{n^*}$ proves $0 = 1$ then so does $\mathrm{ID}^*_{n^*-1}$.*

Proof. Assume that there is a proof of $0 = 1$ in $\mathrm{HA}^2_{n^*}$. Then by Theorem 1.6.12 we have $\mathrm{ID}^*_{n^*-1} \vdash \exists f (f \Vdash \emptyset \sqsupset 0 = 1)$. Hence $\mathrm{ID}^*_{n^*-1} \vdash \exists n (n \vdash_{n^*-1} \emptyset \sqsupset 0 = 1)$ by Lemma 1.6.6. Thus $\mathrm{ID}^*_{n^*-1} \vdash \exists n' (n' \vdash_0 \emptyset \sqsupset 0 = 1)$ by Lemma 1.5.16. So we obtain $\mathrm{ID}^*_{n^*-1} \vdash 0 = 1$ by Lemma 1.5.7. ☐

Remark 1.6.14 (Equiconsistency of $\mathrm{HA}^2_{n^*}$ and $\mathrm{ID}^*_{n^*-1}$). Together with Corollary 1.4.11 the above Corollary 1.6.13 reads as follows. $\mathrm{HA}^2_{n^*}$ proves $0 = 1$ if and only if $\mathrm{ID}^*_{n^*-1}$ does.

Corollary 1.6.15 (Conservativity for first order formulae). *If $\mathrm{HA}^2_{n^*} \vdash A$ for some first order formula A then $\mathrm{ID}^*_{n^*-1}$ shows $\mathrm{HA}^2_0 \vdash A$.*

Proof. From the proof of A we have by Theorem 1.6.12 that $\mathrm{ID}^*_{n^*-1} \vdash \exists f (f \Vdash \emptyset \sqsupset A)$. Lemma 1.6.6 yields $\mathrm{ID}^*_{n^*-1} \vdash \exists g (g \vdash_{n^*-1} \emptyset \sqsupset A)$, hence by the Collapsing Lemma 1.5.16 we get $\mathrm{ID}^*_{n^*-1} \vdash \exists g' (g' \vdash_0 \emptyset \sqsupset A)$. Inspection of Definition 1.5.4 shows that all the rules of $\cdot \vdash_0 \cdot \sqsupset \cdot$ are also proof rules of HA^2_0, hence $\mathrm{ID}^*_{n^*-1}$ can read g' as a proof in HA^2_0. ☐

Remark 1.6.16. If one believes in reflection of $\mathrm{ID}^*_{n^*-1}$, Corollary 1.6.15 might look disturbing. Although the systems HA^2_n increase in consistency strength over PRA, they should have the same arithmetical consequences. But what about the consistency statement for HA^2_n, which is provable in HA^2_{n+1}? The answer is, that in the absence of induction the correct notion of the consistency statement is no longer the Π^0_1-statement "For all proofs it is not the case…", but the Σ^1_1-statement "For all *natural numbers* it is not the case that they code a proof..." which has the shape $\forall x.\mathbb{N}x \to \ldots$ and is no longer an arithmetical formula.

Lemma 1.6.17. *There is a primitive recursive* $\mathfrak{f}$ *with* $\mathrm{ID}_0^* \vdash \forall n.\mathfrak{f}n \Vdash \emptyset \sqsupset \mathbb{N}\underline{n}$.

Proof. Argue informally in ID_0^*. For every n we have to construct a derivation (in the sense of $\cdot \vdash_0 \cdot \sqsupset \cdot$) of $\forall \mathfrak{X}_0.\mathfrak{X}_0 0 \to \forall x(\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}(x))) \to \mathfrak{X}_0\underline{n}$. To do so, it suffices to construct a derivation of $\mathfrak{X}_0\underline{n}$ of $\mathfrak{X}_0 0$ and $\forall x(\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}(x)))$. We do this in the obvious way and verify correctness of this construction by induction(!) on n. ∎

Lemma 1.6.18 (Provably recursive functions of $\mathrm{HA}^2_{n^*}$). *If* $\mathrm{HA}^2_{n^*} \vdash \forall x.\mathbb{N}x \to \neg\forall y \neg R(x,y)$ *then there is a partially recursive* $\mathfrak{f}$ *such that* $\mathrm{ID}^*_{n^*-1} \vdash \forall x.R(x,\mathfrak{f}x)$.

Proof. Assume $\mathrm{HA}^2_{n^*} \vdash \forall x.\mathbb{N}x \to \neg\forall y\neg R(x,y)$. By Theorem 1.6.12 we have a partially recursive $\mathfrak{f}$ with $\mathrm{ID}^*_{n^*-1} \vdash (\mathfrak{f}(\ulcorner\sigma\urcorner) \Vdash \emptyset \sqsupset \forall x.\mathbb{N}x \to \neg\forall y\neg R(x,y))$, where σ is the identical substitution. Note that the argument of $\mathfrak{f}$ is fixed, but in our meta theory we cannot prove $\mathfrak{f}$ to be defined at this argument.

By Lemma 1.6.17 we obtain a natural number f such that $\mathrm{ID}^*_{n^*-1} \vdash \forall n(\{\{f\}(\ulcorner\dot{\underline{n}}\urcorner)\}(\mathfrak{f}(n)) \Vdash \emptyset \sqsupset \neg\forall y\neg R(\underline{n},y))$.

Now argue informally in $\mathrm{ID}^*_{n^*-1}$. By Lemma 1.6.6 we have a partially recursive $\mathfrak{g}$ such that $\forall n(\mathfrak{g}(n) \vdash_{n^*-1} \emptyset \sqsupset \neg\forall y\neg R(\underline{n},y))$. Hence by Lemma 1.5.16 we get another partially recursive $\mathfrak{g}'$ such that $\forall n(\mathfrak{g}'(n) \vdash_0 \emptyset \sqsupset \neg\forall y\neg R(\underline{n},y))$.

This derivation must consist of a subderivation of $\bot$ from the assumption $\forall y\neg R(\underline{n},y)$. Because the only axiom ending in $\bot$ is $\mathcal{S}x = 0 \to \bot$ and by Lemma 1.5.7 no derivation of $\mathcal{S}x = 0$ from negated assumptions exists, this derivation must have the shape that the assumption $\forall y\neg R(\underline{n},y)$ is instantiated to a proof of $R(\underline{n},y)$ and, by Lemma 1.5.7 this formula holds true. We can extract the y in a recursive way. ∎

Summing up, we get the following

Corollary 1.6.19. *Let n be a natural number and $R(\cdot,\cdot)$ a primitive recursive atom. Then the following are equivalent.*

(α) $\mathrm{ID}^c_n \vdash \forall x\exists y R(x,y)$

(β) $\mathrm{ID}^-_n \vdash \forall x\neg\forall y\neg R(x,y)$

(γ) $\widetilde{\mathrm{HA}}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$

(δ) $\mathrm{HA}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$

(ε) $\mathrm{HA}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y\neg R(x,y)$

(φ) *For some natural number e it is the case that* $\mathrm{ID}^*_n \vdash \forall x R(x,\{\underline{e}\}(x))$ *and this statement is to be read that in particular* $\mathrm{ID}^*_n \vdash \forall x\exists y.\{\underline{e}\}(x) = y$.

Proof. $(\alpha)\Rightarrow(\beta)$ is Corollary 1.2.9; $(\beta)\Rightarrow(\gamma)$ is Corollary 1.4.12; $(\gamma)\Rightarrow(\delta)$ holds since $\widetilde{\mathrm{HA}}^2_{n+1}$ is a subsystem of HA^2_{n+1}; $(\delta)\Rightarrow(\varepsilon)$ is trivial; $(\varepsilon)\Rightarrow(\varphi)$ is Lemma 1.6.18 and $(\varphi)\Rightarrow(\alpha)$ holds since ID^*_n is a subsystem of ID^c_n. ∎

2 Adding Induction on the Natural Numbers

Our fragments of second order arithmetic do not contain induction on the natural numbers. Now we will investigate in which way the strength of these systems is influenced if induction is added and obtain two results. On the one hand, adding full induction results in an increase of the proof theoretical strength by an additional inductive definition. On the other hand induction on arithmetical formulae only affects the arithmetical consequences of the systems.

2.1 $\mathbb{N}$ Is the Best Zeroth Inductive Definition

It should be noted that Corollary 1.6.19 in particular shows a reduction from $\mathrm{HA}^2_{n^*+1}$ to $\widetilde{\mathrm{HA}}^2_{n^*+1}$. We will now give some meaning to this reduction by showing that $\widetilde{\mathrm{HA}}^2_{n^*+1}$ *is* the same as "$\mathrm{HA}^2_{n^*}$ + Induction". So, the best thing to do with your 0'th second order quantification is to define the natural numbers, if you don't have them already. This subsection can be understood as an analogue in the framework of second order arithmetic to Arai's observation [3] that "ID^c_p without Induction" essentially *is* ID^c_{p-1}.

Within this subsection we assume a canonical choice $\mathfrak{X}_0, \mathfrak{X}_1, \ldots$ of second order variables of levels $0, 1, \ldots$ and may assume that, up to level-preserving α-equivalence all occurring formulae are built from these canonical variables only.

The system "HA^2_n + Full Induction" (or "HA^2_n + Induction" for short) is formally defined to be the system HA^2_n enriched with the axiom scheme

$$\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \mathcal{A}(t)$$

for *arbitrary* formulae $\mathcal{A}$ and terms t of the language.

In systems where we have the full principle of induction available, all objects *are* natural numbers, hence the predicate $\mathbb{N}$ trivializes. In $\widetilde{\mathrm{HA}}^2_{n+1}$ however, $\mathbb{N}$ is the *only* predicate that can be formed with $\mathfrak{X}_0$. This observation can be used to save one level in our quantification hierarchy.

Definition 2.1.1. For $A \in \widetilde{\mathcal{I}}^*_{n+1}$ define $A^\flat \in \mathcal{I}^*_n$ by induction on A as follows.

- $\mathbb{N}t^\flat \equiv t = t$
- $R\vec{t}^{\,\flat} \equiv R\vec{t}$, for $R\vec{t}$ an atomic formula in $\mathcal{L}_0$ and $\bot^\flat \equiv \bot$
- $\mathfrak{X}_{n+1}t^\flat \equiv \mathfrak{X}_n t$
- $(\forall \mathfrak{X}_{n+1} A)^\flat \equiv \forall \mathfrak{X}_n A^\flat$
- For all other connectives $\cdot^\flat$ is defined in a homomorphic way, that is, $(\forall x A)^\flat \equiv \forall x A^\flat$, $(A \to B)^\flat \equiv A^\flat \to B^\flat$ and $(A \wedge B)^\flat \equiv A^\flat \wedge B^\flat$

Lemma 2.1.2. *If* $\widetilde{\mathrm{HA}}^2_{n+1}$ *proves the sequent* $\Gamma \vdash A$, *that is, if* $\widetilde{\mathrm{HA}}^2_{n+1}$ *proves* A *under the assumptions* Γ, *then "*HA^2_n + Induction*" proves the sequent* $\Gamma^\flat \vdash A^\flat$.

Proof. Induction on the derivation. Note that all the logical rules and axioms translate identically, except for introduction and elimination of $\forall \mathfrak{X}_0$. Introduction of $\forall \mathfrak{X}_0$ can only yield a derivation of $\mathbb{N}t$ under some assumptions; since

$\mathbb{N}t^\flat \equiv t = t$ we can derive the whole statement directly, as it is an axiom. To translate $\forall \mathfrak{X}_0$ elimination we need a way to get an instance of the induction scheme out of a derivation of $t = t$. Fortunately, in "HA^2_n + Induction" we have induction as an axiom scheme and hence do not need the (useless) derivation of $t = t$. ∎

Induction can be eliminated by relativizing all first order quantifiers to natural numbers. We already used this technique to embed ID^-_n into the system HA^2_{n+1} which does not have induction.

Definition 2.1.3. For $A \in \mathcal{I}^2_n[\mathfrak{X}_n]$ we define $A^\sharp \in \widetilde{\mathcal{I}}^2_{n+1}[\mathfrak{X}_{n+1}]$ by induction on the formula A as follows.

- $(\forall x A)^\sharp \equiv \forall x.\mathbb{N}x \to A^\sharp$
- $(\mathfrak{X}_n t)^\sharp \equiv \mathfrak{X}_{n+1} t$
- $(R\vec{t})^\sharp \equiv R\vec{t}$ and $\bot^\sharp \equiv \bot$
- $(\forall \mathfrak{X}_n A)^\sharp \equiv \forall \mathfrak{X}_{n+1} A^\sharp$
- $(A \wedge B)^\sharp \equiv A^\sharp \wedge B^\sharp$ and $(A \to B)^\sharp = A^\sharp \to B^\sharp$

Lemma 2.1.4. *For every formula $\mathcal{A} \in \mathcal{I}^*_n$*

$$\widetilde{\mathrm{HA}}^2_{n+1} \vdash (\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \forall x \mathcal{A}(x))^\sharp.$$

Proof. This is a simple argument and in fact literally the same as the proof of Lemma 1.4.5. ∎

Lemma 2.1.5. *If "HA^2_n + Induction" proves the sequent $\Gamma \vdash A$ then $\widetilde{\mathrm{HA}}^2_{n+1}$ proves the sequent $\mathbb{N}\vec{x}, \Gamma^\sharp \vdash A^\sharp$ where $\vec{x}$ are the free variables of Γ, A.*

Proof. Induction on the derivation. All proof rules translate identical, except for the induction principle, where we use Lemma 2.1.4, and $\forall x$-elimination, where we use Lemma 1.4.9. It should be noted that at the only position where the free variables of Γ, A decrease, (when an $\forall x$-introduction is used) we can also get rid of the assumption $\mathbb{N}x$ via $\to$-introduction. ∎

To get a characterization of "HA^2_n + Induction" we finally have to note that $\cdot^\sharp$ and $\cdot^\flat$ are sufficiently inverse to each other. More precisely we have

Lemma 2.1.6. *For $A \in \mathcal{I}^*_n$ the formulae A and $A^{\sharp\flat}$ are equivalent over minimal predicate logic and the equality axiom $t = t$.*

Proof. By induction on A one easily verifies that the only place where A and $A^{\sharp\flat}$ actually differ is in the case of a first order universal quantification. In this case we have $(\forall x A)^{\sharp\flat} \equiv \forall x.x = x \to A^{\sharp\flat}$ and the equivalence can easily be shown. ∎

Summing up we obtain the following

Corollary 2.1.7. *For $A \in \mathcal{I}^*_n$ with free variables among $\vec{x}$ the following are equivalent*

- "HA^2_n + Induction" $\vdash A$,
- $\widetilde{\mathrm{HA}}^2_{n+1} \vdash \mathbb{N}\vec{x} \to A^\sharp$.

Proof. By Lemma 2.1.5 the assumption "HA^2_n + Induction" $\vdash A$ implies $\widetilde{\mathrm{HA}}^2_{n+1} \vdash \mathbb{N}\vec{x} \to A^\sharp$. By Lemma 2.1.2 this in turn implies "HA^2_n + Induction" $\vdash \vec{x} = \vec{x} \to A^{\sharp^\flat}$. The equivalence is completed by Lemma 2.1.6. ∎

For Π^0_2-statements we obtain the following

Corollary 2.1.8. *Let n be a natural number and $R(\cdot,\cdot)$ a primitive recursive function symbol. Then the following are equivalent (and equivalent to the other statements of Corollary 1.6.19).*

- "HA^2_n *+ Induction*" $\vdash \forall x \neg\forall y \neg R(x,y)$
- $\mathrm{HA}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$

Proof. By Corollary 2.1.7 the first statement is equivalent to $\widetilde{\mathrm{HA}}^2_{n+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$ which by Corollary 1.6.19 is equivalent to the second statement. ∎

Finally we note that the reduction from HA^2_{n+1} via $\widetilde{\mathrm{HA}}^2_{n+1}$ to "HA^2_n + Induction" is rather indirect, as the first of these two steps involved the formalization of a cut elimination. It would be desirable to find a more direct embedding of these two fragments of analysis.

2.2 Compatibility with Arithmetical Induction

From a foundational point of view the systems HA^2_n form a nice and clean hierarchy and adding induction brings us up one level in this hierarchy. However, from a practical point of view the complete lack of induction in the systems HA^2_n makes them hard to use, as in particular no reasonable coding is available. So it seems desirable to add induction. Since for coding only arithmetical induction is used (and usually even only Π^0_1-induction) we can add enough induction for coding without spoiling the proof theoretic strength.

In this subsection we will investigate the effect of adding Π^0_k-induction to the systems HA^2_n. We will show (in Remark 2.2.18) that only the arithmetical consequences are affected, while the consistency strength and the provably recursive functions in the sense of Corollary 1.6.19 remain unchanged (Lemma 2.2.20). So in a formal sense, arithmetical induction and second order quantification are orthogonal concepts.

Definition 2.2.1. The system $\mathrm{HA}^2_{\mathrm{n},(\mathrm{k})}$ is defined to be HA^2_n extended by the *axiom* scheme

$$\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \forall x \mathcal{A}(x)$$

for every Π^0_k-formula $\mathcal{A}$.

Remark 2.2.2. In particular, $\mathrm{HA}^2_{\mathrm{n},(0)}$ is HA^2_n with induction for all Δ^0_0-formulae. This contains all the equations of Primitive Recursive Arithmetic and in particular coding is available. Hence arguments can be formalized in the usual way, even if coding is needed. This will be used in Subsection 3.1.

Remark 2.2.3. Obviously "$\mathrm{HA}^2_{n^*}$ + Arithmetical Induction" is just the union of the systems $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k})}$ for all k. So by analyzing these we also get results on the compatibility of arithmetical induction with our fragments of second order arithmetic.

Let $n^* \geq 1$ and $k^* \geq 0$ be fixed natural numbers. We will describe how to adapt the proofs in Subsections 1.5 and 1.6 to obtain an analysis of $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$ rather than only of $\mathrm{HA}^2_{n^*}$. We will only discuss the changes that have to be made in the proofs of the previous section. In that way the essential points of the given analysis are highlighted. This might help a reader who got lost in the technical details to recall the main ideas.

Definition 2.2.4 (Normal derivations for $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$). By iterated inductive definitions we define the relation $d \vdash_{n,(k^*)} \Gamma \sqsupset A$ for $n < n^*$ and d a natural number, $A \in \mathcal{I}^*_{n+1}$ a formula and $\Gamma \subset \mathcal{I}^*_{n+1}$ a finite set of formulae in the same way as we defined (in Definition 1.5.4) the relation $d \vdash_n \Gamma \sqsupset A$ with the difference that the case (Axiom) now reads as follows.

- (Axiom) $\langle 0 \rangle \vdash_{n,(k^*)} \Gamma \sqsupset A$ if A is $B\sigma$, where B is either one of the axioms of Definition 1.1.2 or of the form $\forall x(\mathcal{A}(x) \to \mathcal{A}(\mathcal{S}x)) \to \mathcal{A}(0) \to \forall x \mathcal{A}(x)$ for some $\Pi^0_{k^*}$-formula $\mathcal{A}$.

Concerning formalization and notation the same remarks as for the relation $d \vdash_n \Gamma \sqsupset A$ apply (see Remark 1.5.5).

Remark 2.2.5. It is trivially noted that all axioms have bounded logical complexity, that the relation $\cdot \vdash_0 \cdot \sqsupset \cdot$ is primitive recursive, and one can still uniquely reconstruct from a given witness d all the formulae Γ', A' and relations $d' \vdash_{k,(k^*)} \Gamma' \sqsupset A'$ that led to the relation $d \vdash_{n,(k^*)} \Gamma \sqsupset A$ in the inductive definition.

Remark 2.2.6. It should be noted, that Lemma 1.5.7 does not hold any longer for the relation $\cdot \vdash_{0,(k^*)} \cdot \sqsupset \cdot$, as we now do have axioms with $\perp$ as a premise. This is due to the fact that we have induction for arbitrary $\Pi^0_{k^*}$-formulae, including those with $\perp$.

However, we do not need that lemma in such a sharp formulation and can come up with a replacement for every of its uses, that is, for Corollary 1.6.13 and Lemma 1.6.18.

Obviously weakening still holds for the new notion of semi-formal proof.

Proposition 2.2.7 ($\forall x$-elimination). *There is a partially recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_{n,(k^*)} \Gamma \sqsupset \forall x A(x)$ then $\mathfrak{f}(d, \ulcorner\Gamma\urcorner, \ulcorner\forall x A\urcorner, \ulcorner t\urcorner) \vdash_{n,(k^*)} \Gamma \sqsupset A(t)$.*

Proof. We first note that, with identical proof, first order substitution (Lemma 1.5.10) also holds for our new notion of semi-formal derivations. The reason for this is that the only property of our axioms actually used is that they are closed under substitution. Thus we can use the old proof of $\forall x$-elimination (Lemma 1.5.11) unchanged. □

Proposition 2.2.8 (Substitution for second order variables). *There is a primitive recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_{n,(k^*)} \Gamma \sqsupset$*

A then $\mathfrak{f}(\ulcorner\Gamma\urcorner, \ulcorner A\urcorner, \ulcorner\Sigma\urcorner, \ulcorner\Sigma'\urcorner, d) \vdash_{n,(k^*)} \Gamma\Sigma'\Sigma \sqsupset A\Sigma'\Sigma$, *where Σ is a substitution of second order variables with only variables of level n in its domain such that $\Gamma\Sigma$ and $A\Sigma$ are well-formed, and Σ' is a permutation of second order variables of level less than n.*

Proof. All our axioms are still *arithmetical* formulae and as such not affected by second order substitutions. Therefore we can use the proof of Lemma 1.5.14 unchanged. ◻

Proposition 2.2.9 (Admissibility of $\forall\mathfrak{X}_n$-elimination). *There is a primitive recursive $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_{n,(k^*)} \Gamma \sqsupset \forall\mathfrak{X}_n A(\mathfrak{X}_n)$ with $\Gamma \subset \mathrm{Neg}^2_n$ and $\mathcal{A} \in \mathcal{I}^*_n$ such that $A(\mathcal{A})$ is a well-formed formula, then $\mathfrak{f}(\ulcorner\forall\mathfrak{X}_n A(\mathfrak{X}_n)\urcorner, \ulcorner\mathcal{A}\urcorner, \ulcorner\Gamma\urcorner, d) \vdash_{n,(k^*)} \Gamma \sqsupset A(\mathcal{A})$.*

Proof. Obviously the subformula property (Lemma 1.5.13) also holds for our new notion of derivation. Moreover, in Proposition 2.2.8 we have just seen that the substitution property for second order variables still holds. But this was essentially the only property used to show admissibility of $\forall\mathfrak{X}_n$-elimination (Lemma 1.5.15). ◻

Proposition 2.2.10 (Collapsing). *There is a partially recursive function $\mathfrak{f}$ such that the following is a theorem of ID^*_n. If $d \vdash_{n,(k^*)} \Gamma \sqsupset A$ with $\Gamma \subset \mathrm{Neg}^2_\ell$ and $A \in \mathrm{Pos}^2_\ell$ for some $\ell < n$ then $\mathfrak{f}(\ell, d, \ulcorner\Gamma\urcorner, \ulcorner A\urcorner) \vdash_{\ell,(k^*)} \Gamma \sqsupset A$.*

Proof. The old proof (Lemma 1.5.16) only depended on the axioms in that they had to be independent of the level n of the inductively defined notion of semi-formal derivation. But also for our new notion they are. ◻

Definition 2.2.11. The relation $f \Vdash_{(k^*)} \Gamma \sqsupset A$ is defined as $f \Vdash \Gamma \sqsupset A$, but based on $\vdash_{n^*-1,(k^*)}\ \sqsupset$ rather than on $\vdash_{n^*-1}\ \sqsupset$. That is we have

$$f \Vdash \Gamma \sqsupset C \quad \Leftrightarrow \quad f \vdash_{n^*-1,(k^*)} \Gamma \sqsupset C$$
if C atomic or of the form $\forall\mathfrak{X}A$

The same remarks on formalization and notational conventions as for the relation $\Vdash\ \sqsupset$ apply.

As our new notion of derivations still enjoys weakening, so does our new notion of the forcing relation.

Lemma 2.2.12 (Renaming of second order variables). *For every C and level preserving permutation Σ of second order variables there are partially recursive functions $\mathfrak{f}^C_\Sigma$ such that the dependency of $\mathfrak{f}^C_\Sigma$ from $\ulcorner C\urcorner$ and $\ulcorner\Sigma\urcorner$ is primitive recursive. Moreover, for every natural number $\mathfrak{k}$ the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. If C has at most $\mathfrak{k}$ logical symbols and $f \Vdash_{(k^*)} \Gamma \sqsupset C$ then $\mathfrak{f}^C_\Sigma(f) \Vdash_{(k^*)} \Gamma\Sigma \sqsupset C\Sigma$.*

Proof. The only property of the underlying notion of derivation used in the proof of Lemma 1.6.5 was that renaming of second order variables is possible. This is provided by Proposition 2.2.8. ◻

Lemma 2.2.13. *For every formula C there are partially recursive functions $\mathfrak{f}^C_\Gamma$ and $\mathfrak{g}^C_\Gamma$ such that the dependency of $\mathfrak{f}^C_\Gamma$ and $\mathfrak{g}^C_\Gamma$ from $\ulcorner C\urcorner$ is primitive recursive. Moreover, for every natural number k the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. Assume that C has at most k logical symbols. Then*

- $e \vdash_{n^*-1,(k^*)} \Gamma \sqsupset C,\ (e)_0 \leq 4,\ C \in \mathrm{Neg}^2_{n^*-1} \quad\Rightarrow\quad \mathfrak{f}^C_\Gamma(e) \Vdash_{(k^*)} \Gamma \sqsupset C$
- $f \Vdash_{(k^*)} \Gamma \sqsupset C,\ C \in \mathrm{Pos}^2_{n^*-1} \quad\Rightarrow\quad \mathfrak{g}^C_\Gamma(f) \vdash_{n^*-1,(k^*)} \Gamma \sqsupset C$

Proof. We can use the proof of Lemma 1.6.6 unchanged. This works as the main ingredient of this proof was the definition of the forcing relation as a "logical relation". In fact, only the following properties of the underlying notion of derivation have been used. For derivations which are "not by introduction" ($(e)_0 \leq 4$), an elimination rule may be added, and there are canonical proofs of $A \vdash A$, not by introduction. ∎

Lemma 2.2.14 (Forcing under substitution). *There is a partially recursive $\mathfrak{f}$ such that for every natural number k the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. Assume that each of the formulae $A, \vec{A}$ has at most k symbols. Assume moreover that $d \vdash_{n^*-1,(k^*)} \Delta, \vec{A} \sqsupset A$ and $f_i \Vdash_{(k^*)} \Gamma \sqsupset A_i$ for $i = 1, \ldots, \ell$ and some $\Delta, \Gamma \subset \mathrm{Neg}^2_{n^*-1}$. Then $\mathfrak{f}(\langle d, \langle f_1, A_1\rangle, \ldots, \langle f_n, A_n\rangle\rangle) \Vdash_{(k^*)} \Delta, \Gamma \sqsupset A$.*

Proof. First we note, that the Ω-rule is still admissible for our new forcing relation. The reason is that the proof of Lemma 1.6.8 only depended on the forcing relation as a logical relation and the fact that an Ω-rule was present in the underlying notion of derivability.

Next we note that our new set of axioms still does not contain second order quantification. So all our axioms are still in $\mathrm{Neg}^2_{n^*-1}$ and so Proposition 2.2.13 can be applied.

So everything immediately depending on the special form of our derivations is present. Moreover, with Propositions 2.2.7, 2.2.12, 2.2.13 and 2.2.8 we have all the other ingredients needed in the proof of Lemma 1.6.9 which therefore can be used unchanged. ∎

Lemma 2.2.15 (Admissibility of ($\forall\mathfrak{X}$-elimination) for the forcing relation). *There is a partially recursive $\mathfrak{f}$, such that for all A and for every $\mathcal{A} \in \mathcal{I}^*_{n^*}$, such that $A(\mathcal{A})$ is well-formed the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. If $\Gamma \subset \mathrm{Neg}^2_{n^*-1}$ and $f \Vdash_{(k^*)} \Gamma \sqsupset \forall\mathfrak{X}_k A(\mathfrak{X}_k)$ then $\mathfrak{f}(f, \ulcorner\Gamma\urcorner, \ulcorner A\urcorner, \ulcorner\mathfrak{X}_k\urcorner, \ulcorner\mathcal{A}\urcorner) \Vdash_{(k^*)} \Gamma \sqsupset A(\mathcal{A})$.*

Proof. We can use the proof of Lemma 1.6.10 unchanged. This proof only depended on the Ω-rule being present in the underlying notion of derivations, the admissibility of $\forall\mathfrak{X}$-elimination in the derivations (Proposition 2.2.9) and the property that derivations can be forced (Proposition 2.2.14). ∎

Lemma 2.2.16 (Cut elimination). *For every proof in $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$ of A from assumptions $\vec{A}$ there is a partially recursive $\mathfrak{f}$ such that the following is a theorem of $\mathrm{ID}^*_{n^*-1}$. For every σ, if $f_i \Vdash_{(k^*)} \Gamma\sigma \sqsupset A_i\sigma$ for all i then $\mathfrak{f}(\ulcorner\sigma\urcorner, \vec{f}) \Vdash_{(k^*)} \Gamma\sigma \sqsupset A\sigma$.*

Proof. Putting together Propositions 2.2.13, the fact that our axioms are still a subset of $\mathrm{Neg}^2_{n^*-1}$ and Lemma 2.2.15 in the same way as in the proof of Theorem 1.6.12. ∎

Corollary 2.2.17 (Conservativity for first order formulae). *If $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)} \vdash A$ for some first order formula A then $\mathrm{ID}^*_{n^*-1} \vdash (\mathrm{HA}^2_{0,(\mathrm{k}^*)} \vdash A)$.*

As $\mathrm{HA}^2_{0,(\mathrm{k}^*)}$ is a subsystem of Π_{k^*}-IA the last statement in particular implies $\mathrm{ID}^*_{n^*-1} \vdash (\Pi_{k^*}\text{-IA} \vdash A)$.

Proof. From the proof of A we obtain by Lemma 2.2.16 that $\mathrm{ID}^*_{n^*-1} \vdash \exists f(f \Vdash_{(k^*)} \emptyset \sqsupset A)$. By Proposition 2.2.13 we get $\mathrm{ID}^*_{n^*-1} \vdash \exists g(g \vdash_{n^*-1,(k^*)} \emptyset \sqsupset A)$ and by Proposition 2.2.10 we obtain $\mathrm{ID}^*_{n^*-1} \vdash \exists g'(g' \vdash_{0,(k^*)} \emptyset \sqsupset A)$. Inspection of the Definition 2.2.4 shows that all the rules of $\cdot \vdash_{0,(k^*)} \cdot \sqsupset \cdot$ are also proof rules of $\mathrm{HA}^2_{0,(\mathrm{k}^*)}$, hence $\mathrm{ID}^*_{n^*-1}$ can read g' as a proof in $\mathrm{HA}^2_{0,(\mathrm{k}^*)}$. ◻

Remark 2.2.18. As $\mathrm{ID}^*_{n^*}$ reflects $\mathrm{ID}^*_{n^*-1}$ we obtain as an immediate consequence of Corollary 2.2.17 that the following is a theorem of $\mathrm{ID}^*_{n^*}$. *All arithmetical consequences of* $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$ *are already consequences of* Π_{k^*}-IA.

Remark 2.2.19. Arguing within set theory, it is possible to obtain the conservativity result for arithmetical formulae (Remark 2.2.18) by the following model theoretic argument.

Assume $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)} \vdash A$ for some first order formula A. We want to show $\mathrm{HA}^2_{0,(\mathrm{k}^*)} \vdash A$. By soundness we have $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)} \models A$ and by completeness it suffices to show $\mathrm{HA}^2_{0,(\mathrm{k}^*)} \models A$. So assume $\mathfrak{M} \models \mathrm{HA}^2_{0,(\mathrm{k}^*)}$ for some first order structure $\mathfrak{M}$, recalling that $\mathrm{HA}^2_{0,(\mathrm{k}^*)}$ is a first order theory. We have to show $\mathfrak{M} \models A$. Let $\tilde{\mathfrak{M}}$ be the second order extension of $\mathfrak{M}$ obtained by taking the full power set(!) as second order structure. Then $\tilde{\mathfrak{M}} \models \mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$. In fact, $\tilde{\mathfrak{M}}$ even is a model of full comprehension (but not necessarily of full induction). Hence $\tilde{\mathfrak{M}} \models A$. But for first order formulae A we have $\tilde{\mathfrak{M}} \models A$ if and only if $\mathfrak{M} \models A$.

However, it should be noted that this simple argument does *not* render Corollary 2.2.17 useless, as it provides also the precise proof theoretical strength of this conservativity statement. Note that the conservativity statement in particular implies the consistency of $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)}$ over Peano Arithmetic. So Corollary 2.2.17 is optimal, as $\mathrm{HA}^2_{n^*}$ and $\mathrm{ID}^*_{n^*-1}$ are equiconsistent by Remark 1.6.14.

Lemma 2.2.20. *If* $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)} \vdash \forall x.\mathbb{N}x \to \neg\forall y \neg R(x,y)$ *then there is a partially recursive* $\mathfrak{f}$ *such that* $\mathrm{ID}^*_{n^*-1} \vdash \forall x.R(x, \mathfrak{f}x)$.

Proof. Assume $\mathrm{HA}^2_{\mathrm{n}^*,(\mathrm{k}^*)} \vdash \forall x.\mathbb{N}x \to \neg\forall y \neg R(x,y)$. By Lemma 2.2.16 we obtain $\mathrm{ID}^*_{n^*-1} \vdash \exists f(f \Vdash_{(k^*)} \emptyset \sqsupset \forall x.\mathbb{N}x \to \neg\forall y\neg R(x,y))$. By Lemma 1.6.17 and the fact that $\Vdash_{(k^*)} \sqsupset$ is more liberal than $\Vdash \sqsupset$ for formulae of the from $\forall \mathfrak{X}$ we get $\mathrm{ID}^*_{n^*-1} \vdash \forall n(\{\{f\}(\ulcorner \underline{\dot{n}} \urcorner)\}(\mathfrak{f}(n)) \Vdash_{(k^*)} \emptyset \sqsupset \neg\forall y \neg R(\underline{n}, y))$.

Now argue informally in $\mathrm{ID}^*_{n^*-1}$. By Proposition 2.2.13 we have a partially recursive $\mathfrak{g}$ such that $\forall n(\mathfrak{g}(n) \vdash_{n^*-1,(k^*)} \emptyset \sqsupset \neg\forall y\neg R(\underline{n}, y))$. Hence by Proposition 2.2.10 we get another partially recursive $\mathfrak{g}'$ such that $\forall n(\mathfrak{g}'(n) \vdash_{0,(k^*)} \emptyset \sqsupset \neg\forall y\neg R(\underline{n}, y))$. This is a derivation in the system Π_{k^*}-IA. *Since every formula provably in* Π_{k^*}-IA *is true,* there really is a y such that $R(n,y)$ holds. Since R is decidable, this y can be found in a recursive way. ◻

Remark 2.2.21. In the proof of Lemma 2.2.20 we used the well-known property that arithmetic reflects the system Π_{k^*}-IA which reads as follows.

$$\mathrm{PA} \vdash \forall n.(\Pi_{k^*}\text{-IA} \vdash \exists y R(\underline{n}, y)) \to \exists y R(n,y)$$

However, we only needed a simpler form thereof, where we only have to deal with proofs that are already partially cut-eliminated.

Corollary 2.2.22. *The statements in Corollary 1.6.19 are also equivalent to*

(∗) $\mathrm{HA}^2_{\mathrm{n+1,(k)}} \vdash \forall x.\mathbb{N}x \to \neg\forall y \neg R(x, y)$.

Proof. (δ)⇒(∗) trivial. (∗)⇒(φ) is Lemma 2.2.20. □

3 Term Systems for the Provably Recursive Functions

In Corollary 1.6.19 we have seen to what extent our fragments of analysis and systems of iterated inductive definitions have the same provably recursive functions. We now aim at describing these functions in a more algorithmic way. To do so we will define term systems that capture the provably recursive functions of our systems in the same way as Gödel's system T captures the functions provably recursive in arithmetic [25].

3.1 Weak Normalization for $\mathrm{F}^{\times}_{n^*+1}$

The polymorphic lambda calculus, also called "system F" has been introduced independently by Girard [22] and Reynolds [37]. The idea is that algorithms are very often uniform in the underlying data type. For example, the identity combinator $\lambda x.x$ can be understood as a function of type $\rho \to \rho$, regardless of the nature of type ρ. To adequately express this uniformity, universal types $\forall\alpha.\sigma(\alpha)$ are introduced. The identity, for example, is assigned the type $\forall\alpha.\alpha \to \alpha$.

In a similar way as arbitrary formulae may be inserted for universally quantified predicate variables in HA^2, arbitrary types may be inserted for universally quantified type variables. This is the basis for the well-known correspondence between system F and second order arithmetic. We will show that this correspondence also holds for fragments. More precisely, we will introduce fragments of systems F by introducing the same restrictions for type quantification as we did for set quantification in our fragments of analysis. For technical reasons we do not use pure system F, but system F extended by primitive (as opposed to coded) pairs.

As we have already analyzed fragments of second order arithmetic by more constructively justified means we are now in the comfortable situation that we can use quite impredicative arguments and hence formalize normalization proofs in a natural way.

We assume a fixed enumeration $\alpha_0, \alpha_1, \ldots$ of type variables. Quantification over the variable α_i will correspond to quantification of a second order variable of level i in HA^2_i.

Definition 3.1.1 (Types of $\mathrm{F}^{\times}_{n+1}$). The types ρ_n of level n are given by the grammar

$$\rho_n, \sigma_n ::= \alpha_n \mid \rho_n \to \sigma_n \mid \rho_n \times \sigma_n \mid \forall\alpha_\ell.\rho_\ell$$

with the proviso that in the last clause we have $\ell < n$. The types of $\mathrm{F}^{\times}_{n+1}$ are the types with level at most $n+1$. Note that the only free type variable in types of level n is α_n. In particular all universal types are closed.

Raw terms are those of the untyped λ-calculus [4] with an additional constructor for the ordered pair, written $\langle s; t\rangle$ and projections, written $t\mathtt{L}$ and $t\mathtt{R}$. When speaking about β-reductions this should include both, the β-reduction of the λ-calculus, that is the reduction $(\lambda x t)s \mapsto t[x{:=}s]$ and the reductions for the ordered pair, that is the reductions $\langle t; s\rangle\mathtt{L} \mapsto t$ and $\langle t; s\rangle\mathtt{R} \mapsto s$.

Remark 3.1.2. As usual, natural numbers are encoded as Church numerals in the λ-calculus. The n'th Church numeral is denoted by c_n and defined to be

$$c_n = \lambda xy.x^n y = \lambda xy.\underbrace{x(x(\dots(x}_{n}\,y))).$$

Remark 3.1.3 (Axiomatization of the β-equality relation). Let $\cdot \sim \cdot$ be an abbreviation for a formalization of β-equality (as Σ^0_1-formula). We list all the properties of $\sim$ that we will need. So an alternative approach would be the addition of a new relation symbol equipped with these properties as axioms. Note all the following properties are provably in $\mathrm{HA}^2_{0,(1)}$.

- $x \sim y \to y \sim x$ and $x \sim y \to y \sim z \to x \sim z$
- for every λ-term $t(x,\vec{y})$ with free variables among $x,\vec{y}$ we have the relation $\ulcorner(\lambda x t(x,\dot{\vec{y}})\dot{x}\urcorner \sim \ulcorner t(\dot{x},\dot{\vec{y}})\urcorner$

  Moreover, we have $\ulcorner c_{\dot{n}}\dot{y}\urcorner \sim \ulcorner \lambda x.\dot{y}^{\dot{n}}x\urcorner$. Note that this is *not* a consequence of the above axiom, as we have more uniformity. This will be needed in the proof of Lemma 3.1.11 and Proposition 3.5.11.
- $\ulcorner\langle\dot{x};\dot{y}\rangle\mathtt{L}\urcorner \sim x$ and $\ulcorner\langle\dot{x};\dot{y}\rangle\mathtt{R}\urcorner \sim y$
- $x \sim y \to \ulcorner\dot{x}\dot{z}\urcorner \sim \ulcorner\dot{y}\dot{z}\urcorner$ and $x \sim y \to \ulcorner\dot{z}\dot{x}\urcorner \sim \ulcorner\dot{z}\dot{y}\urcorner$
- $x \sim y \to \ulcorner\dot{x}\mathtt{L}\urcorner \sim \ulcorner\dot{y}\mathtt{L}\urcorner$ and $x \sim y \to \ulcorner\dot{x}\mathtt{R}\urcorner \sim \ulcorner\dot{y}\mathtt{R}\urcorner$

Definition 3.1.4 ($\Gamma \vdash t\colon \rho$). For Γ a type context, that is, a finite set of pairs $x\colon \sigma$ of variables and types where each variable occurs at most once, t a term and ρ a type we define the arithmetical relation $\Gamma \vdash t\colon \rho$ inductively as follows.

- $\Gamma, x\colon \rho \vdash x\colon \rho$.
- If $\Gamma, x\colon \rho \vdash t\colon \sigma$ then $\Gamma \vdash \lambda x t\colon \rho \to \sigma$.
- If $\Gamma \vdash t\colon \rho \to \sigma$ and $\Gamma \vdash s\colon \rho$ then $\Gamma \vdash ts\colon \sigma$.
- If $\Gamma \vdash t\colon \rho$ and $\Gamma \vdash s\colon \sigma$ then $\Gamma \vdash \langle t; s\rangle\colon \rho \times \sigma$.
- If $\Gamma \vdash t\colon \rho \times \sigma$ then $\Gamma \vdash t\mathtt{L}\colon \rho$ and $\Gamma \vdash t\mathtt{R}\colon \sigma$.
- If $\Gamma \vdash t\colon \rho_\ell$ with α_ℓ not free in Γ then $\Gamma \vdash t\colon \forall\alpha_\ell\rho_\ell$.
- If $\Gamma \vdash t\colon \forall\alpha_\ell\rho_\ell$ then $\Gamma \vdash t\colon \rho_\ell[\alpha_\ell{:=}\sigma_{\ell'}]$ for *arbitrary* ℓ'.

We call $\Gamma \vdash t\colon \rho$ an $\mathrm{F}^\times_{n+1}$-derivation, if and only if all occurring types in the whole underlying generation tree of that relation are $\mathrm{F}^\times_{n+1}$-types.

Definition 3.1.5. For a fixed choice of second order variables $\mathfrak{X}_0, \mathfrak{X}_1, \dots$ of levels $0, 1, \dots$ we define for each $\mathrm{F}^\times_{n+1}$-type ρ a formula $\Phi^\rho(x) \in \mathcal{I}^+_{n+1}$, called the computability predicate of type ρ, by induction on ρ as follows.

$$\begin{array}{lcl}
\Phi^{\alpha_\ell}(x) & = & \exists y. y \sim x \wedge \mathfrak{X}_\ell y \\
\Phi^{\rho\times\sigma}(x) & = & \Phi^\rho(\ulcorner\dot{x}\mathtt{L}\urcorner) \wedge \Phi^\sigma(\ulcorner\dot{x}\mathtt{R}\urcorner) \\
\Phi^{\rho\to\sigma}(x) & = & \forall y. \Phi^\rho(y) \to \Phi^\sigma(\ulcorner\dot{x}\dot{y}\urcorner) \\
\Phi^{\forall\alpha_\ell.\rho}(x) & = & \forall\mathfrak{X}_\ell \Phi^\rho(x)
\end{array}$$

Remark 3.1.6. Recalling that $\exists$ is an abbreviation for $\neg\forall\neg$ a simple induction on the type shows that all formulae Φ^ρ enjoy stability provably in HA^2_{n+1}.

As we have built in β-equality in the base case of our computability predicates, they are invariant under β-equality. More precisely we have

Lemma 3.1.7. *For every* $\mathrm{F}^{\times}_{n+1}$*-type* ρ

$$\mathrm{HA}^2_{\mathrm{n+1,(1)}} \vdash x \sim y \to \Phi^\rho(x) \to \Phi^\rho(y).$$

Proof. Induction on ρ using the properties of $\sim$, as summarized in Remark 3.1.3. ∎

Due to the technical trick of building in β-equality in the base case of our computability predicates they do not preserve substitution. But at least, the formulae are provably equivalent.

Lemma 3.1.8. *For* $\mathrm{F}^{\times}_{n+1}$*-types* ρ *and* σ

$$\mathrm{HA}^2_{\mathrm{n+1,(1)}} \vdash \Phi^\rho(t)[\mathfrak{X}_\ell{:=}\Phi^\sigma] \leftrightarrow \Phi^{\rho[\alpha_\ell:=\sigma]}(t).$$

Proof. Induction on ρ. The only case that is not immediate from the induction hypotheses is the base case. Here we have to show within $\mathrm{HA}^2_{\mathrm{n+1,(1)}}$ that $\exists y . y \sim t \wedge \Phi^\sigma(y)$ and $\Phi^\sigma(t)$ are equivalent. This follows from Lemma 3.1.7 and Remark 3.1.6. ∎

Lemma 3.1.9. *If* $\vec{x}\colon \vec{\rho} \vdash t\colon \rho$ *is a* $\mathrm{F}^{\times}_{n^*+1}$*-derivation then*

$$\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} . \Phi^{\rho_1}(x_1) \to \ldots \to \Phi^{\rho_n}(x_n) \to \Phi^\rho(\ulcorner t(\dot{\vec{x}}) \urcorner).$$

Proof. By induction on $\vec{x}\colon \vec{\rho} \vdash t\colon \rho$ we construct a derivation in $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}}$.

The variable rule is trivial.

If $\ldots \vdash \lambda x t\colon \rho \to \sigma$ was concluded from $\ldots, x\colon \rho \vdash t\colon \sigma$ we have $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} \ldots \to \Phi^\rho(x) \to \Phi^\sigma(\ulcorner t(\dot{x}, \dot{\vec{x}}) \urcorner)$ by induction hypothesis. We have to show $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} \ldots \to (\forall x . \Phi^\rho(x) \to \Phi^\sigma(\ulcorner (\lambda x t(x, \dot{\vec{x}}))\dot{x} \urcorner))$. This can be derived by first order logic and Lemma 3.1.7 since $\ulcorner (\lambda x t(x, \dot{\vec{x}}))\dot{x} \urcorner \sim \ulcorner t(\dot{x}, \dot{\vec{x}}) \urcorner$ holds.

The cut rule is proved by a logical cut and $\forall$-introduction by the introduction rule for second order quantifiers. The $\times$-elimination rules follow by logic (via $\wedge$-elimination).

If $\ldots \vdash \langle t; s\rangle\colon \rho \times \sigma$ was concluded from $\ldots \vdash t\colon \rho$ and $\ldots \vdash s\colon \sigma$ we have by the induction hypotheses that $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} . \ldots \to \Phi^\rho(\ulcorner t(\dot{\vec{x}}) \urcorner)$ and $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} . \ldots \to \Phi^\sigma(\ulcorner s(\dot{\vec{x}}) \urcorner)$. We have to show $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} . \ldots \to \Phi^\rho(\ulcorner \left\langle t(\dot{\vec{x}}); s(\dot{\vec{x}}) \right\rangle \mathtt{L} \urcorner) \wedge \Phi^\sigma(\ulcorner \left\langle t(\dot{\vec{x}}); s(\dot{\vec{x}}) \right\rangle \mathtt{R} \urcorner)$. Since $\ulcorner \left\langle t(\dot{\vec{x}}); s(\dot{\vec{x}}) \right\rangle \mathtt{L} \urcorner \sim \ulcorner t(\dot{\vec{x}}) \urcorner$ and $\ulcorner \left\langle t(\dot{\vec{x}}); s(\dot{\vec{x}}) \right\rangle \mathtt{R} \urcorner \sim \ulcorner s(\dot{\vec{x}}) \urcorner$ the claim follows from Lemma 3.1.7.

If $\ldots \vdash t\colon \rho[\alpha{:=}\sigma]$ was concluded from $\ldots \vdash t\colon \forall \alpha \rho$ we have by induction hypothesis $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} \ldots \to \forall \mathfrak{X}_0 . \Phi^\rho(\ulcorner t(\dot{\vec{x}}) \urcorner))$. We have to show $\mathrm{HA}^2_{\mathrm{n^*+1,(1)}} \vdash \forall \vec{x} \ldots \to \Phi^{\rho[\alpha:=\sigma]}(\ulcorner t(\dot{\vec{x}}) \urcorner)$. This follows from the elimination rule for second order quantifiers and the equivalence shown in Lemma 3.1.8. ∎

Lemma 3.1.10. *Provably in* $\mathrm{HA}^2_{1,(1)}$*, the formula*

$$\Phi^{\mathcal{N}}(\ulcorner t \urcorner) \equiv \forall \mathfrak{X}_0.\forall y.(\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner \dot{y}\dot{z} \urcorner)) \to (\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner t\dot{y}\dot{z} \urcorner))$$

implies $\neg\forall n \neg(\ulcorner t\mathtt{f}\mathtt{s} \urcorner \sim \ulcorner \mathtt{f}^{\dot{n}}\mathtt{s} \urcorner)$ *where* $\mathtt{f}$ *and* $\mathtt{s}$ *are fresh (syntactic) variables.*

Proof. The syntactical equivalence is immediate from the definition. For the implication we argue informally within $\mathrm{HA}^2_{1,(1)}$. We specialize $\mathfrak{X}_0 x$ to $\neg\forall n \neg(x \sim \ulcorner \mathtt{f}^{\dot{n}}\mathtt{s} \urcorner)$ where $\mathtt{f}$ and $\mathtt{s}$ are fresh syntactical variables. We write Φ^* for the result of the substitution on Φ^{α_0} and specialize y to $\ulcorner \mathtt{f} \urcorner$. We note that $\forall z.\Phi^*(z) \to \Phi^*(\ulcorner \mathtt{f}\dot{z} \urcorner)$, so we get $\forall z.\Phi^*(z) \to \Phi^*(\ulcorner t\mathtt{f}\dot{z} \urcorner)$. Now we instantiate z to $\ulcorner \mathtt{s} \urcorner$. Furthermore we note $\Phi^*(\ulcorner \mathtt{s} \urcorner)$. Hence we have $\Phi^*(\ulcorner t\mathtt{f}\mathtt{s} \urcorner)$, that is,

$$\neg\forall y \neg(y \sim \ulcorner t\mathtt{f}\mathtt{s} \urcorner \wedge \neg\forall n \neg(y \sim \ulcorner \mathtt{f}^{\dot{n}}\mathtt{s} \urcorner))$$

So the claim follows from the reflexivity and transitivity of $\sim$ by minimal logic. ∎

Next we show that, provably in $\mathrm{HA}^2_{1,(1)}$, all Church numerals are inhabitants of $\mathcal{N} = \forall \alpha_0.(\alpha_0 \to \alpha_0) \to \alpha_0 \to \alpha_0$. More precisely we have

Lemma 3.1.11. $\mathrm{HA}^2_{1,(1)} \vdash \forall n.\mathbb{N}n \to \Phi^{\mathcal{N}}(\ulcorner c_{\dot{n}} \urcorner)$.

Proof. We argue informally in $\mathrm{HA}^2_{1,(2)}$. Let n be given and use the induction principle provided by the assumption $\mathbb{N}n$. We have to show $\Phi^{\mathcal{N}}(\ulcorner \lambda xy.y \urcorner)$ and $\Phi^{\mathcal{N}}(\ulcorner c_{\dot{n}} \urcorner) \to \Phi^{\mathcal{N}}(\ulcorner c_{\dot{n}+1} \urcorner)$.

Concerning the first claim, we have to show for arbitrary $\mathfrak{X}_0$ and y under the assumption $\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner \dot{y}\dot{z} \urcorner)$ that $\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner (\lambda xy.y)\dot{y}\dot{z} \urcorner)$. We have $\ulcorner (\lambda xy.y)\dot{y}\dot{z} \urcorner \sim \ulcorner \dot{z} \urcorner$, and hence by Lemma 3.1.7 also the claimed implication $\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner (\lambda xy.y)\dot{y}\dot{z} \urcorner)$.

Concerning the second claim assume $\forall \mathfrak{X}_0.\forall y.(\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner \dot{y}\dot{z} \urcorner)) \to (\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner c_{\dot{n}}\dot{y}\dot{z} \urcorner))$. We have to show for arbitrary $\mathfrak{X}_0$ and y under the assumption $\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner \dot{y}\dot{z} \urcorner)$ that $\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner c_{\dot{n}+1}\dot{y}\dot{z} \urcorner)$. So for an arbitrary z assume $\Phi^{\alpha_0}(z)$. We have to show $\Phi^{\alpha_0}(\ulcorner c_{\dot{n}+1}\dot{y}\dot{z} \urcorner)$. From the first assumption we get $\Phi^{\alpha_0}(\ulcorner c_{\dot{n}}\dot{y}\dot{z} \urcorner)$, hence by $(\forall z.\Phi^{\alpha_0}(z) \to \Phi^{\alpha_0}(\ulcorner \dot{y}\dot{z} \urcorner))$ we get $\Phi^{\alpha_0}(\ulcorner \dot{y}(c_{\dot{n}+1}\dot{y}\dot{z}) \urcorner)$. Since $\ulcorner \dot{y}(c_{\dot{n}}\dot{y}\dot{z}) \urcorner \sim \ulcorner c_{\dot{n}+1}\dot{y}\dot{z} \urcorner$, we obtain the claim using Lemma 3.1.7. ∎

Lemma 3.1.12. *If* $\emptyset \vdash t \colon \mathcal{N} \to \mathcal{N}$ *is an* $\mathrm{F}^{\times}_{n+1}$*-derivation then*

$$\mathrm{HA}^2_{n+1,(2)} \vdash \forall n.\mathbb{N}n \to \exists m.\ulcorner t c_{\dot{n}}\mathtt{f}\mathtt{s} \urcorner \sim \ulcorner \mathtt{f}^{\dot{m}}\mathtt{s} \urcorner.$$

Proof. Assume $\emptyset \vdash t \colon \mathcal{N} \to \mathcal{N}$. Then by Lemma 3.1.9 we obtain $\mathrm{HA}^2_{n+1} \vdash \forall z.\Phi^{\mathcal{N}}(z) \to \Phi^{\mathcal{N}}(\ulcorner t\dot{z} \urcorner)$ and by Lemma 3.1.11 we get $\mathrm{HA}^2_{n+1} \vdash \forall n.\mathbb{N}n \to \Phi^{\mathcal{N}}(\ulcorner t c_{\dot{n}} \urcorner)$. So Lemma 3.1.10 yields the claim. ∎

Corollary 3.1.13. *If* $\emptyset \vdash t \colon \mathcal{N} \to \mathcal{N}$ *is a* $\mathrm{F}^{\times}_{n^*+1}$*-derivation then* t *denotes a function on Church numerals provably total in* $\mathrm{ID}^*_{n^*}$*.*

Proof. Assume $\emptyset \vdash t \colon \mathcal{N} \to \mathcal{N}$. By Lemma 3.1.12 we have $\mathrm{HA}^2_{n^*+1} \vdash \forall n.\mathbb{N}n \to \exists m.\ulcorner t c_{\dot{n}}\mathtt{f}\mathtt{s} \urcorner \sim \ulcorner \mathtt{f}^{\dot{m}}\mathtt{s} \urcorner$. Hence the claim follows by Corollary 2.2.22. ∎

The special case $n^* = 0$ of the above corollary is the theorem of Altenkirch and Coquand [2].

3.2 Embedding of the tree classes $\mathcal{O}_n$ into $\mathrm{F}_n^\times$

We have seen (in Corollary 1.6.19) that the functions provably recursive in HA^2_{n+1} and ID^*_n are the same. As objects defined by strictly positive inductive definitions essentially are trees, a functional interpretation into a more restricted system is possible, namely in extensions $\mathcal{O}_n$ of Gödel's system T (which formalizes recursion on the natural numbers [25]) to higher number classes. As these are known results [44, 34] we just recall the term systems and show in which way they can be embedded into the systems $\mathrm{F}_n^\times$.

Later (in Subsections 3.4 and 3.5) we will exhibit a direct functional interpretation for the fragments $\widetilde{\mathrm{HA}}^2_n$ of Analysis, though into the full system $\mathrm{F}_n^\times$.

Intuitively, the zeroth tree class consists of the empty tree; the first tree class is the set of natural numbers; the second tree class is the set of trees branching over the natural numbers; and, in general, the $(n+1)$'st tree class is the set of trees branching over the n'th tree class.

The systems $\mathcal{O}_n$ formalize the $(n+1)$'st tree class. The types of system $\mathcal{O}_n$ are the simple types, built upon the base types $\mathcal{O}_0, \ldots, \mathcal{O}_n$ for the first upto the $(n+1)$'st tree class. The terms of system $\mathcal{O}_n$ are the simply typed λ-terms built upon the following typed constants.

Every tree class is equipped with a constant $\mathtt{0}^\ell$ of type $\mathcal{O}_\ell$ to be interpreted as the empty tree and a constant $\mathtt{S}_\ell$ allowing to construct non-empty trees. $\mathtt{S}_0$ is the usual successor on the natural numbers and has type $\mathcal{O}_0 \to \mathcal{O}_0$. The constants $\mathtt{S}_{\ell+1}$ have type $(\mathcal{O}_\ell \to \mathcal{O}_{\ell+1}) \to \mathcal{O}_{\ell+1}$. So, if $\mathfrak{f}$ is a term of type $\mathcal{O}_\ell \to \mathcal{O}_{\ell+1}$, then $\mathtt{S}_{\ell+1}\mathfrak{f}$ is a term of type $\mathcal{O}_{\ell+1}$. We should think of $\mathtt{S}_{\ell+1}\mathfrak{f}$ as a non-empty tree, where the immediate subtrees are enumerated by $\mathfrak{f}$.

Moreover, for every type ρ and every ℓ there is a constant $\mathtt{R}_{\ell,\rho}$ embodying the iteration principle on $\mathcal{O}_\ell$ for type ρ.

For the natural numbers we have constants of type $\rho \to (\rho \to \rho) \to \mathcal{O}_0 \to \rho$ and the following equations.

$$\begin{aligned} \mathtt{R}_{0,\rho} st \mathtt{0}^0 &= s \\ \mathtt{R}_{0,\rho} st(\mathtt{S}_0 r) &= t(\mathtt{R}_{0,\rho} str) \end{aligned}$$

For the higher tree classes, iteration does not only involve a single "previous value" but a whole family of such, more precisely one "previous value" for every subtree. This family is given as a function. So the iterators $\mathtt{R}_{\ell+1,\rho}$ have type $\rho \to ((\mathcal{O}_\ell \to \rho) \to \rho) \to \mathcal{O}_{\ell+1} \to \rho$ and enjoy the following equalities

$$\begin{aligned} \mathtt{R}_{\ell+1,\rho} st \mathtt{0}^{\ell+1} &= s \\ \mathtt{R}_{\ell+1,\rho} st(\mathtt{S}_{\ell+1} r) &= t((\mathtt{R}_{\ell+1,\rho} st) \circ r) \end{aligned}$$

where we used the abbreviation $s \circ t \equiv \lambda x. s(tx)$ with x a "new" variable.

These term systems can be implemented in our fragments of system F in the same way as the Church numerals [15] are the usual implementation of natural numbers in the lambda-calculus. The idea is that an object is implemented by its own iteration principle.

More formally, we define the $\mathrm{F}_n^\times$-type $\mathcal{O}_n$ of tree ordinals of the $n+1$ number class by induction on n as follows.

$$\begin{aligned} \mathcal{O}_0 &= \mathcal{N} = \forall\alpha.(\alpha \to \alpha) \to \alpha \to \alpha \\ \mathcal{O}_{n+1} &= \forall\alpha.((\mathcal{O}_n \to \alpha) \to \alpha) \to \alpha \to \alpha \end{aligned}$$

We define the zero constructors to be $0^\ell = \lambda xy.y$ and easily verify that they can have type $\mathcal{O}_\ell$. Moreover, we define the usual successor $\mathtt{S}_0 = \lambda \mathtt{n}xy.x(\mathtt{n}xy)$ which can have type $\mathcal{N} \to \mathcal{N}$ and the successor functions for the higher number classes $\mathtt{S}_{\ell+1} = \lambda \mathtt{f}xy.x(\lambda z.\mathtt{f}zxy)$ which can have type $(\mathcal{O}_\ell \to \mathcal{O}_{\ell+1}) \to \mathcal{O}_{\ell+1}$.

As our encodings of the tree classes are equipped with their own iteration principles we can define the iterators quite uniformly as $\mathtt{R}_{\ell,\rho} = \lambda yx\mathtt{n}.\mathtt{n}xy$. It should be noted that these iterators can have the respective types. With that definition the following β-equalities (denoted by $\sim$) hold.

$$\begin{aligned}
\mathtt{R}_{0,\rho}st0^0 &\sim 0^0ts \sim s \\
\mathtt{R}_{0,\rho}st(\mathtt{S}_0r) &\sim \mathtt{S}_0rts \sim t(rts) \sim t(\mathtt{R}_{0,\rho}str)
\end{aligned}$$

$$\begin{aligned}
\mathtt{R}_{\ell+1,\rho}st0^{\ell+1} &\sim 0^{\ell+1}ts \sim s \\
\mathtt{R}_{\ell+1,\rho}st(\mathtt{S}_{\ell+1}r) &\sim \mathtt{S}_{\ell+1}rts \sim t(\lambda z.rzts) \sim t(\lambda z.\mathtt{R}_{\ell+1,\rho}st(rz)) \\
&\equiv t((\mathtt{R}_{\ell+1,\rho}st) \circ r)
\end{aligned}$$

3.3 On Strong Normalization

In Lemma 3.1.12 we have shown that all the terms typable in $\mathrm{F}^\times_{n+1}$ by $\mathcal{N}$ are weakly normalizing. This proof is optimal in that it is shown locally in $\mathrm{HA}^2_{\mathrm{n+1,(1)}}$ and we will later (in Corollary 3.4.25) show that the assumption "$\mathrm{F}^\times_{n+1}$ weakly normalizing" allows us to reflect the provably recursive functions of HA^2_{n+1} and hence also (by Corollary 2.2.22) of $\mathrm{HA}^2_{\mathrm{n+1,(1)}}$. This reflection obviously implies consistency.

However, it is well-known that, locally as a theorem of HA^2, the full system F is even *strongly* normalizing, that is, every reduction sequence is finite. So all fragments of system F are strongly normalizing as well. But the precise strength of these statements for the various fragments is not obvious from that observation.

In this subsection we will formalize the standard text book proof [23, p.114] of strong normalization, adapted for the fragments $\mathrm{F}^\times_n$ of system F. We will be able to formalize this proof locally in "HA^2_n + Full Induction", a system we know (Corollary 2.1.8) to be equivalent to HA^2_{n+1}. So we have an upper bound on the strength of the statement "$\mathrm{F}^\times_n$ is strongly normalizing". An obvious lower bound on the strength of this statement is of course "$\mathrm{F}^\times_n$ is weakly normalizing" which we will show to be of the strength of HA^2_n. These two bounds do not match. But the author is aware neither of a proof of strong normalization for $\mathrm{F}^\times_n$ that can be formalized in a weaker system, nor of a proof exploiting the statement "$\mathrm{F}^\times_n$ strongly normalizing" to a stronger degree than showing the consistency of HA^2_n. So the precise strength remains open. However, in the case of $\mathrm{ID}^c_{<\omega}$, discussed in Section 4, the increase by one level in the hierarchy of the systems ID^c_n catches up and we get a precise characterization.

Besides providing an upper bound on the strength of strong normalization, the reader might find this subsection interesting, for it exposes where induction is used in the traditional proof based on the "candidate method" [22].

Let $n^* \geq 1$ be a natural number. In the following we will denote by $\mathrm{HA}^2_{\mathrm{n}^*,(+)}$ the system "$\mathrm{HA}^2_{n^*}$ + Full Induction", as introduced in Subsection 2.1.

With the predicate "not an abstraction" we denote that the term is neither of the form λxt nor of the form $\langle s; t\rangle$. It should be noted that this is a primitive recursive predicate. The same holds for the one-step reduction relation. We write $r \rightsquigarrow r'$ if precisely one subterm of the form $(\lambda x.t)s$, $\langle t; s\rangle\mathtt{L}$ or $\langle t; s\rangle\mathtt{R}$ is replaced by $t[x{:=}s]$, t or s respectively. The reflexive transitive closure of $\rightsquigarrow$ is denoted by $\rightsquigarrow^*$. It can be expressed by a Σ^0_1-formula in the obvious way.

A term t is said to be strongly normalizing if every reduction sequence starting with t is finite. Since every term has only finitely many reducts this is equivalent to the existence of a finite height of the reduction tree. This can be expressed by stating that there is a natural number such that the reduction tree unfolded up to that height has only normal terms at its leaves. So "t is strongly normalizing" is a Σ^0_1-formula over Primitive Recursive Arithmetic. We assume a formalization in such a way that "t is strongly normalizing" implies that t actually is (the code of) a proper term.

Definition 3.3.1. The formula "$\mathcal{A}$ is a candidate" is the conjunction of the following formulae.

- $\forall x.\mathcal{A}(x) \to$ "x is strongly normalizing"
- $\forall x, y.\mathcal{A}(x) \to (x \rightsquigarrow^* y) \to \mathcal{A}(y)$
- $\forall x.$ "x not an abstraction" $\to (\forall y.(x \rightsquigarrow y) \to \mathcal{A}(y)) \to \mathcal{A}(x)$

It should be noted that no second order quantification is involved in this definition. In particular $\mathcal{A} \in \mathcal{I}^2_k[\mathfrak{X}_k]$ implies that "$\mathcal{A}$ is a candidate" $\in \mathcal{I}^2_k[\mathfrak{X}_k]$.

Remark 3.3.2. The last condition implies that all terms which are not abstractions and normal are already members of every candidate (since there are *no* one-step reducts). In particular, candidates are never empty as they contain all the variables.

Proposition 3.3.3. $\mathrm{HA}^2_{0,(1)} \vdash$ *"" is strongly normalizing" is a candidate".*

Proof. The first property is a logical tautology in this case. For the second property it should be noted that every reduct of a strongly normalizing term has a reduction tree of smaller height. For the third property it suffices to note that the height of the reduction tree of a term is bounded by one plus the maximum of the heights of the reduction trees of its (finitely many!) reducts. Obviously all these arguments can be formalized. ∎

Let $\alpha_0, \alpha_1, \ldots$ be a canonical choice of type variables of levels $0, 1, \ldots$; we may, up to level preserving α-equivalence, assume that all types are built from these variables only. Moreover, let $\mathfrak{X}_0, \mathfrak{X}_1, \ldots$ be a canonical choice of second order variables of levels $0, 1, \ldots$.

Definition 3.3.4. By induction on the $\mathrm{F}^\times_{n^*}$-type ρ we define formulae $\Psi^\rho(x) \in \mathcal{I}^2_{n^*}[\mathfrak{X}_{n^*}]$, called computability predicates, as follows.

$$\begin{array}{lcl}
\Psi^{\alpha_k}(x) & = & \mathfrak{X}_k x \\
\Psi^{\rho\to\sigma}(x) & = & \forall y.\Psi^\rho(y) \to \Psi^\sigma(\ulcorner \dot{x}\dot{y} \urcorner) \\
\Psi^{\rho\times\sigma}(x) & = & \Psi^\rho(\ulcorner \dot{x}\mathtt{L} \urcorner) \wedge \Psi^\sigma(\ulcorner \dot{x}\mathtt{R} \urcorner) \\
\Psi^{\forall\alpha_k\rho}(x) & = & \forall\mathfrak{X}_k.\text{"}\mathfrak{X}_k \text{ a candidate"} \to \Psi^\rho(x)
\end{array}$$

Lemma 3.3.5. *If the free variables of ρ are among $\{\alpha_k\}$ then $\mathrm{HA}^2_{\mathrm{n}^*,(+)} \vdash$ "$\mathfrak{X}_k$ is a candidate" $\to$ "Ψ^ρ is a candidate".*

Proof. Arguing informally in $\mathrm{HA}^2_{\mathrm{n}^*,(+)}$ we have to show that the three properties of Definition 3.3.1 are fulfilled. We do so by induction on ρ.

In the case of a type variable ρ_k there is nothing to show since by assumption $\mathfrak{X}_k$ is a candidate.

In the case $\rho = \sigma \to \tau$ we argue as follows.

- To show strong normalization of x it suffices to show strong normalization of $\ulcorner \dot{x}t \urcorner$ for some arbitrary term t. But this is the case since Ψ^τ is a candidate and Ψ^σ contains at least the variables (as it is a candidate).

- For closure under reduction we argue als follows. Assume $\Psi^\rho(x)$ and $x \rightsquigarrow^* x'$. We have to show $\Psi^\rho(x')$. Assume $\Psi^\sigma(y)$. We have to show $\Psi^\tau(\ulcorner \dot{x}'\dot{y} \urcorner)$. But this holds since $\Psi^\tau(\ulcorner \dot{x}\dot{y} \urcorner)$ and $\ulcorner \dot{x}\dot{y} \urcorner \rightsquigarrow^* \ulcorner \dot{x}'\dot{y} \urcorner$ and as a candidate Ψ^τ is closed under reduction.

- Assume x is not an abstraction and $\forall y.(x \rightsquigarrow y) \to \Psi^\rho(y)$. We have to show $\Psi^\rho(x)$. Hence assume $\Psi^\sigma(y)$. We have to show $\Psi^\tau(\ulcorner \dot{x}\dot{y} \urcorner)$. We do so by showing that all reducts of $\ulcorner \dot{x}\dot{y} \urcorner$ are in Ψ^τ. Since x is not an abstraction all reducts are of the form $\ulcorner \dot{x}'\dot{y} \urcorner$ or $\ulcorner \dot{x}\dot{y}' \urcorner$ with $x \rightsquigarrow^* x'$ or $y \rightsquigarrow^* y'$ respectively. The first case is handled by our assumption, the second case by a side induction on the height of the reduction tree of y.

  The side induction has to be formalized as follows: we show for all natural numbers ℓ that "for all y with height of the reduction tree at most ℓ if $\Psi^\sigma(y)$ then $\Psi^\tau(\ulcorner \dot{x}'\dot{y} \urcorner)$". This requires us to use induction on arbitrary complicated formulae. Note that Ψ^τ can not *be expressed uniformly in τ in our theory.*

The case $\sigma \times \tau$ is immediate from the induction hypotheses. We argue as in the case of $\sigma \to \tau$, but simpler and in particular do *not* need a side induction.

If $\rho = \forall \alpha_k \sigma$ we argue as follows.

- Assume $\Psi^{\forall \alpha_k \sigma}(x)$. We have to show that x is strongly normalizing. This follows from the induction hypothesis for our assumption instantiated to the candidate (Proposition 3.3.3) of strongly normalizing terms.

- Assume $\Psi^{\forall \alpha_k \sigma}$ and $x \rightsquigarrow^* y$. We have to show $\Psi^{\forall \alpha_k \sigma}(y)$. So assume an arbitrary candidate $\mathfrak{X}_k$. We need to show $\Psi^\sigma(y)$. But this is immediate from our induction hypothesis.

- Assume an x which is not an abstraction and assume that for all y with $x \rightsquigarrow y$ we have $\Psi^{\forall \alpha_k \sigma}(y)$. We have to show $\Psi^{\forall \alpha_k \sigma}(y)$. So assume an arbitrary candidate $\mathfrak{X}_k$. Then $\Psi^\sigma(y)$ is immediate from our induction hypothesis. ∎

Remark 3.3.6. As already emphasized, the proof of Lemma 3.3.5 uses induction on the "height of reduction tree" to show $\Psi^\tau(\ldots)$. As these formulae cannot be expressed uniformly in τ in our theory, induction on the full language is needed. Similar inductions have to be used in Lemmata 3.3.10 and 3.3.11. Their uses are marked in the text.

Lemma 3.3.7. $\Psi^{\rho[\alpha_k:=\sigma]}(x) \equiv \Psi^{\rho}[\mathfrak{X}_k{:=}\Psi^{\sigma}]$.

Proof. Simple induction on ρ. ☐

Remark 3.3.8. It should be noted that, as opposed to Lemma 3.1.8, we actually get syntactical equality in Lemma 3.3.7 and not just provable equivalence. The reason is that for our second order variables we have well behaved sets in mind ("candidates") and not just arbitrary sets as we did in Subsection 3.1, so we do not have to use a closure in the base case.

On the other hand, the nicer formulation comes with nasty side inductions, as we have seen in the proof of Lemma 3.3.5. In our proof of weak normalization, the proof of the closure property (Remark 3.1.7) could be pushed to the meta-level, where we constructed the proof by induction on types.

Lemma 3.3.9. *For every* $\mathrm{F}^{\times}_{n^*}$*-type* σ *with free type variables among* $\{\alpha_\ell\}$ *it holds that* $\mathrm{HA}^2_{\mathrm{n}^*,(+)} \vdash \Psi^{\forall\alpha_k\rho}(t) \to$ *"*$\mathfrak{X}_\ell$ *a candidate"* $\to \Psi^{\rho[\alpha_k:=\sigma]}(t)$.

Proof. Arguing informally in $\mathrm{HA}^2_{\mathrm{n}^*,(+)}$ we instantiate our assumption $\Psi^{\forall\alpha_k\rho}(t)$ to the predicate Ψ^{σ}. By Lemma 3.3.5 we know that Ψ^{σ} is a candidate under the additional assumption on $\mathfrak{X}_\ell$. Hence the claim by Lemma 3.3.7. ☐

Lemma 3.3.10. *If* $\vec{x}\colon \vec{\rho} \vdash t\colon \rho$ *is a* $\mathrm{F}^{\times}_{n^*}$*-derivation then* $\mathrm{HA}^2_{\mathrm{n}^*,(+)}$ *proves, under the assumptions "*$\mathfrak{X}_\ell$ *a candidate" for every* α_ℓ *free in* $\vec{\rho}, \rho$*, the formula*

$$\forall\vec{x}.\Psi^{\rho_1}(x_1) \to \ldots \to \Psi^{\rho_n}(x_n) \to \Psi^{\rho}(\ulcorner t(\dot{\vec{x}})\urcorner).$$

Proof. Induction on the Definition 3.1.4 of $\vec{x}\colon \vec{\rho} \vdash t\colon \rho$. We now argue informally in $\mathrm{HA}^2_{\mathrm{n}^*,(+)}$.

The case of an assumption is trivial.

In the case of a λ-abstraction the induction hypothesis provides a proof of $\forall\vec{x},x.\Psi^{\rho_1}(x_1) \to \ldots \to \Psi^{\rho_n}(x_n) \to \Psi^{\rho}(x) \to \Psi^{\sigma}(\ulcorner t(\dot{\vec{x}},\dot{x})\urcorner)$ and we have to show $\Psi^{\rho\to\sigma}(\ulcorner \lambda x.t(\dot{\vec{x}},x)\urcorner)$ under the assumptions $\Psi^{\rho_1}(x_1),\ldots,\Psi^{\rho_n}(x_n)$. So assume an x with $\Psi^{\rho}(x)$. We have to show $\Psi^{\sigma}(\ulcorner (\lambda x.t(\dot{\vec{x}},x))\dot{x}\urcorner)$, for which it suffices to show that all one-step reducts of $\ulcorner (\lambda x.t(\dot{\vec{x}},x))\dot{x}\urcorner$ belong to Ψ^{σ}, since by Lemma 3.3.5 the predicate Ψ^{σ} is a candidate. But all those one-step reducts are either $\ulcorner t(\dot{\vec{x}},\dot{x})\urcorner$ which belongs to Ψ^{σ} by induction hypothesis, or are of the form $\ulcorner (\lambda x.t'(\dot{\vec{x}},x))\dot{x}\urcorner)$ with $t \rightsquigarrow t'$. Note that by the induction hypothesis t belongs to Ψ^{σ}, hence is strongly normalizing. So a side induction on the height of the reduction tree of t shows the claim. *Here we use induction to show the formula* $\Psi^{\sigma}(\ulcorner (\lambda x.t(\dot{\vec{x}},x))\dot{x}\urcorner)$.

The case of an application is immediate by logic and so are the cases of projection and $\forall\alpha_\ell$-introduction.

In the case of pairing we argue similar as in the case of λ-abstraction. *Here we use induction to show the formulae* $\Psi^{\sigma}(\ulcorner \langle t;s\rangle\mathtt{L}\urcorner)$ *and* $\Psi^{\rho}(\ulcorner \langle t;s\rangle\mathtt{R}\urcorner)$.

The case of $\forall\alpha_\ell$-elimination we have by induction hypothesis $\Psi^{\forall\alpha_\ell\rho}(\ulcorner t(\dot{\vec{x}})\urcorner)$. We want to show $\Psi^{\rho[\alpha_\ell:=\sigma]}(\ulcorner t(\dot{\vec{x}})\urcorner)$. By Lemma 3.3.7 this is the same as showing $\Psi^{\rho}[\mathfrak{X}_\ell{:=}\Psi^{\sigma}]$ which we can show by logic since we have $\Psi^{\forall\alpha_\ell\rho}(\ulcorner t(\dot{\vec{x}})\urcorner)$ and by Lemma 3.3.5 the predicate Ψ^{σ} is a candidate. ☐

Lemma 3.3.11. $\mathrm{HA}^2_{1,(+)} \vdash \forall n.\Psi^{\mathcal{N}}(\ulcorner c_{\dot{n}}\urcorner)$.

Proof. Argue informally in $\mathrm{HA}^2_{1,(+)}$ as follows. Assume a candidate $\mathfrak{X}_0$ and x, y with $\Psi^{\alpha_0\to\alpha_0}(x)$ and $\Psi^{\alpha_0}(y)$. Induction on n shows that $\Psi^{\alpha_0}(\ulcorner \dot{x}^{\dot{n}}\dot{y}\urcorner)$. Hence a side induction on all the subterms $\vec{t}, s$ in $\ulcorner(\lambda y.t_0(t_1(\dots(t_n y))))s\urcorner = \ulcorner \dot{x}^{\dot{n}}\dot{y}\urcorner$ being strongly normalizing shows that all one-step reducts of $\ulcorner(\lambda y.\dot{x}^{\dot{n}}y)\dot{y}\urcorner$ belong to Ψ^{α_0}, so the only one-step reduct of $\ulcorner c_{\dot{n}}\dot{x}\dot{y}\urcorner$ belongs to Ψ^{α_0}.

Here we used induction (on the sum of the reduction heights) to show the formulae $\mathfrak{X}_0\ulcorner(\lambda y.\dot{x}^{\dot{n}}y)\dot{y}\urcorner$ *and* $\mathfrak{X}_0\ulcorner c_{\dot{n}}\dot{x}\dot{y}\urcorner$. ∎

Corollary 3.3.12. *If t is typable in $\mathrm{F}^{\times}_{n^*}$ with type $\mathcal{N}\to\mathcal{N}$ then*

$$\mathrm{HA}^2_{\mathrm{n}^*,(+)} \vdash \forall n.\text{“}\ulcorner tc_{\dot{n}}\urcorner \textit{ is strongly normalizing”}.$$

Proof. By Lemmata 3.3.10 and 3.3.11 we know $\mathrm{HA}^2_{\mathrm{n}^*,(+)} \vdash \Psi^{\mathcal{N}}(\ulcorner tc_{\dot{n}}\urcorner)$. By Lemma 3.3.4 we know that this is a candidate set, hence all members are strongly normalizing by Definition 3.3.1. ∎

3.4 Witnessing Terms for the Functions Provably Recursive in $\widetilde{\mathrm{HA}}^2_{n^*+1}$ via the Direct Method.

We have seen (in Corollary 3.1.13) that all the functions definable in $\mathrm{F}^{\times}_{n^*+1}$ are provably recursive in $\mathrm{HA}^2_{n^*+1}$. We will now work towards the reverse direction and will show that every function provably recursive in $\mathrm{HA}^2_{n^*+1}$ actually *can* be defined in $\mathrm{F}^{\times}_{n^*+1}$.

The idea is that a proof in second order arithmetic can be read as a λ-term, if we ignore all first order quantifications. However, we lack a constructive existential quantifier and hence have to deal with “$\forall\neg\forall\neg$”-proofs, that is, after instantiation to the input we obtain a proof of absurdity from the assumption $\forall y.\mathbb{N}y \to \mathrm{R}(\underline{k}, y) \to \bot$. An investigation of normal such proofs will reveal that in fact an incorrect instance of this assumption has been used. In other words, the proof of $\mathbb{N}y$ that is provided to this assumption *is* the result of the function. We will use this assumption to pass the answer through. So when assigning types to formulae we have to assign to atoms (or at least to the atom $\bot$) the type $\mathcal{N}$ of natural numbers. This idea has been used by various authors [7, 17].

So our atoms will already be interpreted by quantified types. To avoid malformed types due to other quantifications of level 0 we will use our reduction of $\mathrm{HA}^2_{n^*+1}$ to $\widetilde{\mathrm{HA}}^2_{n^*+1}$ and in fact assume the proof of totality to be formalized in $\widetilde{\mathrm{HA}}^2_{n^*+1}$ already. So the quantification of $\mathfrak{X}_0$ will range over atoms of the form $\mathfrak{X}_0 t$ only. And to these atoms we will assign α_0 as corresponding type.

Before we start working out the details we should contemplate the kind of results we can and cannot expect. First of all we obviously cannot expect, within our meta theory, a result stating something like

> “If $\mathrm{HA}^2_{n^*+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg\mathrm{R}(x,y))$ then there is a term t in the corresponding ramified fragment $\mathrm{F}^{\times}_{n^*+1}$ of system F, such that for every x there is a y such that $t\underline{x}$ and $\underline{y}$ are β-equal and $\mathrm{R}(x,y)$ holds.”

The reason is that, in the case $\mathrm{R} = \bot$, this would yield a proof of consistency of $\mathrm{HA}^2_{n^*+1}$, and thus also of $\mathrm{ID}^c_{n^*}$, out of the blue. But $\mathrm{ID}^c_{n^*}$ is a strong extension of our meta theory so that the result would contradict Gödel's incompleteness theorem [24].

There are two possible ways to overcome this problem. One way is to give up insisting that the statement "t represents the function proven to be recursive" should hold. Instead one could be satisfied if the statement is provable in some appropriate theory. This approach will be pursued in Subsection 3.5 and yields the optimal characterization shown in Corollary 3.5.21. In this subsection we will use a more direct approach and show the above statement, but under the strong proviso that "$\mathrm{F}^{\times}_{n^*+1}$ is weakly normalizing". Although proof theoretically slightly less optimal this proof highlights very nicely the direct connection between proofs in second order arithmetic and computations in the polymorphic lambda calculus.

We assume a canonical choice of type variables $\alpha_0, \alpha_1, \ldots$ of levels $0, 1, \ldots$. As in all our fragments of analysis, second order variables come in different levels. We denote the level of the second order variable $\mathfrak{X}$ by $\{\!\{\mathfrak{X}\}\!\}$. As usual the type $\mathcal{N}$ of Church numerals is defined as $\mathcal{N} = \forall\alpha_0.(\alpha_0 \to \alpha_0) \to \alpha_0 \to \alpha_0$.

We also assume a canonical enumeration of the object variables of F. In this subsection we will write x_n for variable number $2n$ and y_n for variable number $2n+1$. Note that we do not need any assumption on the types of these variables, as our variables are type-free. Instead we will use, as we already did, a type assignment calculus (the so-called "Curry style of system F" [5]).

Definition 3.4.1. By induction on the buildup of formulae we define the realizing type $\{\!\{A\}\!\}$ of a formula A as follows.

- $\{\!\{\bot\}\!\} = \mathcal{N} = \forall\alpha_0.(\alpha_0 \to \alpha_0) \to \alpha_0 \to \alpha_0$
- $\{\!\{R\vec{t}\}\!\} = \mathcal{N}$
- $\{\!\{\mathfrak{X}t\}\!\} = \alpha_{\{\!\{\mathfrak{X}\}\!\}}$
- $\{\!\{A \to B\}\!\} = \{\!\{A\}\!\} \to \{\!\{B\}\!\}$
- $\{\!\{\forall x A\}\!\} = \{\!\{A\}\!\}$
- $\{\!\{\forall \mathfrak{X} A\}\!\} = \forall\alpha_{\{\!\{\mathfrak{X}\}\!\}}\{\!\{A\}\!\}$
- $\{\!\{A \wedge B\}\!\} = \{\!\{A\}\!\} \times \{\!\{B\}\!\}$

Remark 3.4.2. For the later development it is inessential which type we use for *arithmetical* atoms, as long as it is closed and contains at most one quantifier (to ensure type-correctness of the obtained terms). So we could as well have used the empty type $\forall\alpha\alpha$.

However, as discussed in the introduction, it *is* essential that we use $\mathcal{N}$ for the atom $\bot$ in order to be able to pass the answer through that atom. The details of this passing mechanism can be seen in the proof of Theorem 3.4.23.

For a better understanding of the types chosen here one might consult the more "semantical" explanation in Subsection 3.5 and in particular Definition 3.5.1.

Proposition 3.4.3. *If $A \in \widetilde{\mathcal{I}}^2_{n^*+1}$ then $\{\!\{A\}\!\}$ is an $\mathrm{F}^{\times}_{n^*+1}$-type.*

Proof. Induction on $A \in \widetilde{\mathcal{I}}^2_{n^*+1}$. ∎

Note that at this point the restriction of $\widetilde{\mathcal{I}}^2_{n^*+1}$, as compared to $\mathcal{I}^2_{n^*+1}$ is crucial. For if we allowed atoms other than $\mathfrak{X}_0 t$ in the scope of $\forall \mathfrak{X}_0 \dots$ the nesting level would increase by one. However, it is not important, that the $\mathbb{N}$-predicate was the only predicate allowed to be of the form $\forall \mathfrak{X}_0 A$. Arbitrary formulae built of atoms of the form $\mathfrak{X}_0 t$ by means of conjunction, implication and universal quantification would have sufficed for this purpose.

Definition 3.4.4. For t a term, A a formula, and Γ a context, that is, a finite set of pairs $x_\ell{:}A$ of variables from $\{x_\ell \mid \ell \in \mathbb{N}\}$ and formulae such that every variable occurs at most once left of the colon, we define the Σ^0_1-relation $t \gg \Gamma \vdash A$ inductively as follows.

- $y_i \gg \Gamma \vdash A$ if A is an axiom and i the natural number such that $\{\!\{A\}\!\} = \underbrace{\mathcal{N} \to \dots}_{i} \to \mathcal{N}$. Note that $\{\!\{A\}\!\}$ necessarily has to have a type of this shape, as can be seen from Definition 1.1.2.
- $x \gg x{:}A, \Gamma \vdash A$.
- If $t \gg x{:}A, \Gamma \vdash B$ then $\lambda x t \gg \Gamma \vdash A \to B$.
- If $r \gg \Gamma \vdash A \to B$ and $s \gg \Gamma \vdash A$ then $rs \gg \Gamma \vdash B$.
- If $t \gg \Gamma \vdash A$ with x not free in Γ then $t \gg \Gamma \vdash \forall x A$.
- If $t \gg \Gamma \vdash \forall x A(x)$ then $t \gg \Gamma \vdash A(s)$.
- If $t \gg \Gamma \vdash A \wedge B$ then $t\mathtt{L} \gg \Gamma \vdash A$ and $t\mathtt{R} \gg \Gamma \vdash B$.
- If $t \gg \Gamma \vdash A$ and $s \gg \Gamma \vdash B$ then $\langle t; s\rangle \gg \Gamma \vdash A \wedge B$.
- If $t \gg \Gamma \vdash A$ with $\mathfrak{X}$ not free in Γ then $t \gg \Gamma \vdash \forall \mathfrak{X} A$.
- If $t \gg \Gamma \vdash \forall \mathfrak{X} A(\mathfrak{X})$ then $t \gg \Gamma \vdash A(\mathcal{A})$.

If we had used a Church style of typing, the last two rules would have to be formulated with the terms $\Lambda \alpha_{\{\!\{\mathfrak{X}\}\!\}} t$ and $t\{\!\{\mathcal{A}\}\!\}$.

Proposition 3.4.5 (Weakening). *If $t \gg \Gamma \vdash A$ and $\Gamma' \supset \Gamma$ then $t \gg \Gamma' \vdash A$.*

Proof. Induction on $t \gg \Gamma \vdash A$. ∎

Remark 3.4.6. If $t \gg \vec{x}{:}\vec{A} \vdash A$ then the free variables of the term t are among $\{\vec{x}\} \cup \{y_i \mid i \in \mathbb{N}\}$.

Proof. Induction on $t \gg \Gamma \vdash A$. ∎

Proposition 3.4.7. *If $t \gg \vec{x}{:}\vec{A} \vdash A$ then t has (in the sense of Definition 3.1.4) type $\{\!\{A\}\!\}$ in the context*

$$x_1{:}\{\!\{A_1\}\!\}, \dots, x_n{:}\{\!\{A_n\}\!\}, y_{i_1}{:}\underbrace{\mathcal{N} \to \dots}_{i_1} \to \mathcal{N}, \dots, y_{i_k}{:}\underbrace{\mathcal{N} \to \dots}_{i_k} \to \mathcal{N}$$

where $i_1, \dots, i_n$ are the "arities" of the used axioms.

Proof. Induction on $t \gg \Gamma \vdash A$. ∎

Proposition 3.4.8. *If* $\widetilde{\mathrm{HA}}^2_{n^*+1}$ *proves a formula* A *from some assumptions* $\vec{A}$ *then for some term* $t \in \mathrm{F}^{\times}_{n^*+1}$ *it holds that* $t \gg \vec{x}{:}\vec{A} \vdash A$.

Proof. Induction on the notion of proof in HA^2_{n+1}. Note that the proof rules in HA^2_{n+1} are precisely those of Definition 3.4.4. Hence we should think of the term as just a "canonical decoration" of the derivation. The fact that t is typable follows from Proposition 3.4.7 together with Proposition 3.4.3. ∎

Proposition 3.4.9. *If* $t \gg \vec{x}{:}\vec{A} \vdash A$ *then* $t \gg \vec{x}{:}\vec{A}\sigma \vdash A\sigma$ *and* $t \gg \vec{x}{:}\vec{A}\Sigma \vdash A\Sigma$. *Moreover, the number of inferences in the derivation is not changed.*

Proof. Induction on $t \gg \Gamma \vdash A$. ∎

Proposition 3.4.10 (Removal of undecorated redexes). *If* $t \gg \Gamma \vdash A$ *then there is a derivation of this judgment without redexes which are not reflected in the term* t. *In other words, there is a derivation of this judgment where no elimination follows immediately an introduction of a (first or second order) quantifier.*

Proof. Induction on the number of inferences in the derivation of $t \gg \Gamma \vdash A$. To reduce a redex we use Proposition 3.4.9. ∎

The next lemma can be understood as part of a proof that $\mathbb{N}$ is a data type in the sense of Krivine's [29] system AF_2.

Lemma 3.4.11. *If* $x^k z \gg x{:}\forall x.\mathfrak{X}x \to \mathfrak{X}(\mathcal{S}x), z{:}\mathfrak{X}0 \vdash \mathfrak{X}t$ *then* $t = \underline{k}$.

Proof. Induction on k. Using Proposition 3.4.10 we may assume that the derivation of $x^k z \gg x{:}\forall x.\mathfrak{X}x \to \mathfrak{X}(\mathcal{S}x), z{:}\mathfrak{X}0 \vdash \mathfrak{X}\underline{\ell}$ does not contain any undecorated redexes.

In the case $k = 0$ the assumption $\mathfrak{X}0$ cannot be instantiated further and therefore is the conclusion, hence $\ell = 0$. In the case $k = k' + 1$ the assumption $\forall x.\mathfrak{X}x \to \mathfrak{X}(\mathcal{S}x)$ must have been instantiated in such a way that the atoms $\mathfrak{X}(\mathcal{S}x)$ and $\mathfrak{X}t$ coincide. So t must be of the form $t = \mathcal{S}t'$ and the assumption was instantiated to t'. Then $x^{k'} z \gg x{:}\forall x.\mathfrak{X}x \to \mathfrak{X}(\mathcal{S}x), z{:}\mathfrak{X}0 \vdash \mathfrak{X}t'$ and we are done by the induction hypothesis. ∎

Proposition 3.4.12. *If* $t \gg \Gamma, x{:}B \vdash A$ *and* $s \gg \Delta \vdash B$ *then it is the case that* $t[x{:=}s] \gg \Gamma, \Delta \vdash A$.

Proof. Induction on the length of the derivation of $t \gg \Gamma, B \vdash A$, using Proposition 3.4.9 where necessary (to get fresh variable names where otherwise a conflict would arise because of free variables in Δ).

A typical case would be an inference from $t \gg \Gamma, x{:}B \vdash A(y)$ to $t \gg \Gamma, x{:}B \vdash \forall y A(y)$. In this case we first apply Proposition 3.4.9 to obtain $t \gg \Gamma, z{:}B \vdash A(z)$ for some new z, using that y not free in Γ, B. Since the number of inferences is unchanged, we can use the induction hypothesis to obtain $t[x{:=}s] \gg \Gamma, \Delta \vdash A(z)$, from where we conclude $t[x{:=}s] \gg \Gamma, \Delta \vdash \forall z A(z)$. ∎

Corollary 3.4.13. *If* $(\lambda x t)s \gg \Gamma \vdash A$ *then* $t[x{:=}s] \gg \Gamma \vdash A$.

Proof. Using Proposition 3.4.10 we may assume that the derivation of $(\lambda xt)s \gg \Gamma \vdash A$ does not contain any decorated redex. As applications occur only as decorations of cut, the abstraction λxt must be the decoration of a $\to$-formula, concluded by an $\to$-introduction (since we have no undecorated redexes). So we can apply Proposition 3.4.12. □

Proposition 3.4.14. *If $\langle t; s\rangle \mathtt{L} \gg \Gamma \vdash A$ then $t \gg \Gamma \vdash A$ and if $\langle t; s\rangle \mathtt{R} \gg \Gamma \vdash A$ then $s \gg \Gamma \vdash A$.*

Proof. Using Proposition 3.4.10 we may assume that the derivation of $(\lambda xt)s \gg \Gamma \vdash A$ contains no undecorated redex. An inspection of the proof rules in Definition 3.4.4 shows that $\langle t; s\rangle \mathtt{L} \gg \Gamma \vdash A$ must have been concluded from $\langle t; s\rangle \gg \Gamma \vdash A \wedge B$ for some formula B. And this must have been concluded from $t \gg \Gamma \vdash A$ and $s \gg \Gamma \vdash B$, so we are done. Similar for the right projection. □

Corollary 3.4.15 (Subject reduction). *If $\mathrm{F}^{\times}_{n^*+1}$ is weakly normalizing and $t \gg \Gamma \vdash A$ then there is a term t' which is normal and β-equal to t such that $t' \gg \Gamma \vdash A$.*

Proof. Immediate from Corollary 3.4.13 and Proposition 3.4.14. □

Lemma 3.4.16. *If there is a normal derivation of $s = t$ with at most open assumptions ending in $\perp$ then $s = t$ holds.*

Proof. Induction on the proof figure shows that *none* of the assumptions has been used: The conclusion is an equation, not the atom $\perp$, and hence cannot be concluded immediately from one of the assumptions. So an axiom was used. Inspection of the axioms (Definition 1.1.2) shows that again only equations occur negatively. The induction also shows that only quantifier-free formulae occur. Defining Δ^0_0-truth predicates for those formulae shows that the conclusion is true.

Note the similarity of this proof to that of Lemma 1.5.7. □

Remark 3.4.17. The following argument constitutes a semantical proof of Lemma 3.4.16. Note that the axioms and assumptions become true if we interpret $\perp$ as truth (which is possible since we work in minimal logic only). All conclusions therefore have to be true.

However, in order to formalize this argument, we have to get a bound on the logical complexity of the occurring formulae (in fact, we even need that all occurring formulae are Δ^0_0, as we are working in PRA), as is done in the above proof. As a side-effect we obtain that none of the assumptions has been used, which might be of independent interest.

Lemma 3.4.18. *Every normal derivation $r \gg z{:}\forall y.\mathbb{N}y \to R(t,y) \to \perp\ \vdash \perp$ has the shape, that $\perp$ was obtained by instantiating the assumption $\forall y.\mathbb{N}y \to R(t,y) \to \perp$ to a term s (and proofs of $\mathbb{N}s$ and $R(t,s)$). Moreover, for that term $R(s,t)$ holds.*

Proof. Inspection of the axioms (Definition 1.1.2) shows that the only axiom which $\perp$ might have been concluded from is $\mathcal{S}t' = 0 \to \perp$. This would require a normal proof of $\mathcal{S}t' = 0$ from assumptions ending in $\perp$, which cannot exist by Lemma 3.4.16. Hence the assumption $\forall y.\mathbb{N}y \to R(t,y) \to \perp$ must have been used and instantiated to a term s and normal proofs of $\mathbb{N}s$ and $R(t,s)$. Lemma 3.4.16 yields that $R(t,s)$ holds. □

Proposition 3.4.19. *For every k it holds that $c_k \gg \emptyset \vdash \mathbb{N}\underline{k}$.*

Proof. By induction on k one easily constructs a proof of

$$x^k y \gg x{:}\forall x.\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}x), y{:}\mathfrak{X}_0 0 \vdash \mathfrak{X}_0\underline{k},$$

which suffices. □

Proposition 3.4.20. *Every closed normal term of $\mathrm{F}^{\times}_{n^*+1}$ of type $\mathcal{N}$ is η-equal to a numeral.*

Proof. As the term is closed (and no products occur in the type) it has to start with a λ-abstraction and thus is of the form λxt. By the same argument, the term either is $\lambda x.x$, which is η-equal to c_1, or is of the form $\lambda xy.t'$. Induction on t' shows that it is of the form $t' = x^k y$ for some k. □

Lemma 3.4.21. *Every normal proof $t \gg \Gamma, x{:}\forall x.\mathfrak{X}x \to \mathfrak{X}(\mathcal{S}x), z{:}\mathfrak{X}0 \vdash \mathfrak{X}s$, where the context Γ contains only formulae ending in $\bot$, uses neither an axiom nor an assumption of the context Γ.*

Proof. Induction on the proof figure. No axiom or formula in Γ ends in $\mathfrak{X}s$, hence one of the last two assumptions has been used. The claim is trivial, if the last assumption has been used. If the other assumption has been used, we conclude by induction hypothesis, since the only subformula occurring negatively is again of the form $\mathfrak{X}s$. □

The following corollary shows that, in order to prove that a particular term is a numeral, the proof itself has to be that particular numeral. That is, for our proofs we have the stricter property than the usual, that the proof of a witness just has to be big enough [35, p.241].

Corollary 3.4.22. *If $t \gg \Gamma \vdash \mathbb{N}s$ for some normal term t and some context Γ with formulae ending in $\bot$, then for some ℓ it is the case that $t = c_\ell$ and $s = \underline{\ell}$.*

Proof. Using Proposition 3.4.10 we conclude that the last rule of the derivation must have been $\forall\mathfrak{X}$-introduction and the two rules before that must have been $\to$-introductions. Hence $t' \gg \Gamma, x{:}\forall x.\mathfrak{X}x \to \mathfrak{X}(x+1), z{:}\mathfrak{X}0 \vdash \mathfrak{X}s$ where $t = \lambda xz.t'$. By Lemma 3.4.21 we conclude that the free variables of t' are among $\{x,y\}$. Hence t is a closed term of type $\mathcal{N}$, so by Proposition 3.4.20 a numeral and we are in the situation of Lemma 3.4.11. □

Theorem 3.4.23. *Suppose $\mathrm{F}^{\times}_{n^*+1}$ is weakly normalizing. If $\widetilde{\mathrm{HA}}^2_{n^*+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$ then there is a $t \in \mathrm{F}^{\times}_{n^*+1}$ computing this function on Church numerals, that is, for every n the term tc_n reduces to a Church numeral c_ℓ and $R(\underline{n},\underline{\ell})$ holds.*

Proof. From $\widetilde{\mathrm{HA}}^2_{n^*+1} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$ we get by Proposition 3.4.8 a term r with $r \gg x{:}\mathbb{N}x, y{:}\forall y(\mathbb{N}y \to \neg R(x,y)) \vdash \bot$. For every k we have $r[x{:=}c_k] \gg y{:}\forall y(\mathbb{N}y \to \neg R(\underline{k},y)) \vdash \bot$, using Propositions 3.4.9, 3.4.19 and 3.4.12.

By Lemma 3.4.15 and proposition 3.4.10 we get a normal derivation of this statement with β-equal decoration r'.

By Proposition 3.4.18 we know that r' has the form $r' = yt's$ where $t' \gg y{:}\forall y(\mathbb{N}y \to \neg R(\underline{k}, y)) \vdash \mathbb{N}\hat{t}$ and $R(\underline{k}, \hat{t})$ holds. Moreover, by Corollary 3.4.22 we know that, for some natural number ℓ, it is the case that $\hat{t} = \underline{\ell}$ and $t' = c_\ell$.

So we found a realizing term: $\lambda x.r[y{:=}\lambda z'z''.z']$. Note that if we set $\tilde{t} = \lambda xy.r$ then $\tilde{t} \gg \emptyset \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$ and, by Proposition 3.4.7, the term $\tilde{t}$, in some appropriate context, has type $\{\!\{\forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))\}\!\} = \mathcal{N} \to (\mathcal{N} \to \mathcal{N} \to \mathcal{N}) \to \mathcal{N}$. The proof shows that the term computing the witnesses is β-equal to $\lambda x.\tilde{t}x(\lambda z'z''.z')$ which has type $\mathcal{N} \to \mathcal{N}$, as desired. □

Remark 3.4.24. It is worth noting that in the proof of Theorem 3.4.23 it is the proof figure that is used as a program where for the axioms *variables* are plugged in. So all the axioms have no "computational content" whatsoever, as they are completely discarded during the computation. Still the precise choice of the axioms was important to ensure correctness of the computed result.

Corollary 3.4.25. *The statements of Corollary 1.6.19 also imply*

($*$) *Primitive Recursive Arithmetic plus "*$\mathrm{F}^{\times}_{n+1}$ *weakly normalizing" proves:*

"There is a function $\mathfrak{f}$ *definable in* $\mathrm{F}^{\times}_{n+1}$ *such that* $R(n, \mathfrak{f}n)$ *for all* n*."*

Proof. $(\gamma)\Rightarrow(*)$ is Theorem 3.4.23. □

Remark 3.4.26. Recalling the proof of $(\alpha)\Rightarrow(*)$ we note the following: $(\alpha)\Rightarrow(\beta)$ is a double negation translation, formally defined by induction on the buildup of the proof; $(\beta)\Rightarrow(\gamma)$ is an embedding, also defined by induction on the buildup of the proof; and $(\gamma)\Rightarrow(*)$ works by leaving out parts of the proof figure and interpreting the rest as an $\mathrm{F}^{\times}_{n+1}$-term. So all in all we have a functional interpretation for ID^c_n, the classical system of (natural numbers and) n positive inductive definitions.

In other words, we have a new form of program extraction for ID^c_n. The program can be calculated primitive recursively from the proof. Strong principles are only needed to validate correctness.

3.5 An Approach Based on Realizability

In Subsection 3.4 we exhibited a method to translate a totality proof into a term of $\mathrm{F}^{\times}_{n^*+1}$ so that it represents the function shown to be recursive. By studying the shape of normal proofs we adopted an external view on the process of interpreting proofs as terms. The necessary proof theoretic strength was obtained by stating theorems under the proviso "$\mathrm{F}^{\times}_{n^*+1}$ weakly normalizing" wherever necessary.

In this subsection we aim at a proof theoretically equivalent characterization. To do so we use a more internal view of the same process and translate proofs into proofs of other statements using a realizability technique. Realizability was first studied by Kleene [28] under the name of "recursive realizability".

In this subsection let $n^* \geq 1$ be a natural number. We use $\sim$ as notation for an appropriate formalization of β-equality. We will again use only the properties of β-equality summarized in Remark 3.1.3.

Large parts of this subsection will be relative to a comprehension formula $\mathcal{A}(z)$, specifying what it should mean to realize absurdity. Later we will set $\mathcal{A}(z) \equiv \exists y.z \;\underline{\mathbf{r}}\; \mathbb{N}y \wedge R(x,y)$, but we develop the theory as far as possible and reasonable by only assuming that $\mathcal{A} \in \mathcal{I}_1^2$ and ends in $\perp$. Nevertheless we will suppress $\mathcal{A}$ in the notation of the relation $\underline{\mathbf{r}}$.

In the following definition we should think of t as the code for a term in $\mathrm{F}_{n^*}^{\times}$ of type $\{\!\{A\}\!\}$. Morally we would like to identify β-equal terms. However, this is undecidable and therefore our theory does not allow such an identification. So we make explicit where we argue about β-equality. This explains the somewhat strange base case in the following definition. Moreover, since in this case only a classical existential quantifier can be used, we have to provide a "classical environment". Hence the double-negation of atoms.

Definition 3.5.1 ($t \;\underline{\mathbf{r}}\; A$). By induction on A we define a formula $t \;\underline{\mathbf{r}}\; A$ relative to a given formula $\mathcal{A}(x) \in \mathcal{I}_1^2$ ending in $\perp$ as follows.

- $t \;\underline{\mathbf{r}}\; \perp \equiv \mathcal{A}(t)$
- $t \;\underline{\mathbf{r}}\; R(\vec{t}) \equiv \neg\neg R(\vec{t})$ for atomic formulae $R(\vec{t})$ of $\mathcal{L}_0$.
- $t \;\underline{\mathbf{r}}\; \mathfrak{X}_i s \equiv \exists x.x \sim t \wedge \mathfrak{X}_i(\langle x, s\rangle)$
- $t \;\underline{\mathbf{r}}\; \forall \mathfrak{X}_i A \equiv \forall \mathfrak{X}_i.t \;\underline{\mathbf{r}}\; A$
- $t \;\underline{\mathbf{r}}\; (A \to B) \equiv \forall x.(x \;\underline{\mathbf{r}}\; A) \to (\ulcorner \dot{t}\dot{x} \urcorner \;\underline{\mathbf{r}}\; B)$
- $t \;\underline{\mathbf{r}}\; (A \wedge B) \equiv (\ulcorner \dot{t}\mathtt{L} \urcorner \;\underline{\mathbf{r}}\; A) \wedge (\ulcorner \dot{t}\mathtt{R} \urcorner \;\underline{\mathbf{r}}\; B)$
- $t \;\underline{\mathbf{r}}\; \forall x.A \equiv \forall x.t \;\underline{\mathbf{r}}\; A$

Remark 3.5.2. It should be noted that $t \;\underline{\mathbf{r}}\; \mathbb{N}y$, and in fact $t \;\underline{\mathbf{r}}\; A$ for any A without $\perp$, is independent of the choice of the underlying $\mathcal{A}(x)$. Moreover

$$\begin{aligned} t \;\underline{\mathbf{r}}\; \mathbb{N}s \equiv \forall \mathfrak{X}_0 \forall y.(\forall x \forall z.(z \;\underline{\mathbf{r}}\; \mathfrak{X}_0 x) \to (\ulcorner \dot{y}\dot{z} \urcorner \;\underline{\mathbf{r}}\; \mathfrak{X}_0(\mathcal{S}x))) \\ \to \forall z.(z \;\underline{\mathbf{r}}\; \mathfrak{X}_0 0) \to (\ulcorner \dot{t}\dot{y}\dot{z} \urcorner \;\underline{\mathbf{r}}\; \mathfrak{X}_0 s) \end{aligned}$$

Proof. $t \;\underline{\mathbf{r}}\; \mathbb{N}s \equiv t \;\underline{\mathbf{r}}\; (\forall \mathfrak{X}_0.(\forall x.\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}x)) \to \mathfrak{X}_0 0 \to \mathfrak{X}_0 s) \equiv \forall \mathfrak{X}_0 \forall y.(y \;\underline{\mathbf{r}}\; (\forall x.\mathfrak{X}_0 x \to \mathfrak{X}_0(\mathcal{S}x))) \to (\ulcorner \dot{t}\dot{y} \urcorner \;\underline{\mathbf{r}}\; (\mathfrak{X}_0 0 \to \mathfrak{X}_0 s)) \equiv \forall \mathfrak{X}_0 \forall y.(\forall x \forall z.(z \;\underline{\mathbf{r}}\; \mathfrak{X}_0 x) \to (\ulcorner \dot{y}\dot{z} \urcorner \;\underline{\mathbf{r}}\; \mathfrak{X}_0(\mathcal{S}x))) \to \forall z.(z \;\underline{\mathbf{r}}\; \mathfrak{X}_0 0) \to (\ulcorner \dot{t}\dot{y}\dot{z} \urcorner \;\underline{\mathbf{r}}\; \mathfrak{X}_0 s)$ ∎

Remark 3.5.3. All formulae of the form $t \;\underline{\mathbf{r}}\; A$ end in absurdity and hence enjoy stability over minimal logic.

By a simple induction on A one easily shows

Proposition 3.5.4. *If $A \in \widetilde{\mathcal{I}}_{n^*}^*$ then $(t \;\underline{\mathbf{r}}\; A) \in \mathcal{I}_{n^*}^*$.*

Remark 3.5.5. It should be noted that in Proposition 3.5.4 we essentially used that $\perp$ does not occur within the scope of $\forall \mathfrak{X}_0$. Nevertheless the resulting formulae are only in $\mathcal{I}_{n^*}^*$ and not necessarily in $\widetilde{\mathcal{I}}_{n^*}^*$. But as we do have a (albeit rather indirect) proof theoretic reduction (Corollary 1.6.19 and Corollary 2.2.22) of $\mathrm{HA}^2_{\mathrm{n}^*,(2)}$ to $\widetilde{\mathrm{HA}}^2_{n^*}$, we do not lose generality.

Proposition 3.5.6. *If the underlying formula $\mathcal{A}$ is invariant under β-equality, that is, if* $\mathrm{HA}^2_{n^*}1 \vdash x \sim x \to \mathcal{A}(x) \to \mathcal{A}(y)$ *then for every formula* $A \in \widetilde{\mathcal{I}}^*_{n^*}$ *we have*

$$\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash x \sim y \to (x \ \underline{\mathbf{r}} \ A) \to (y \ \underline{\mathbf{r}} \ A)$$

Proof. By induction on A we construct a proof in $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$. The only cases that are not immediate from the induction hypothesis are the case $A \equiv \bot$, where we have made an appropriate assumption, and the case of a second order variable where we only have to use transitivity of β-equality, due to our definition of realizability. □

A simply induction on A shows

Proposition 3.5.7. *If A contains neither a (free or bound) second-order variable nor $\bot$ then $x \ \underline{\mathbf{r}} \ A$ does* not *contain x free, in particular $t \ \underline{\mathbf{r}} \ A \equiv s \ \underline{\mathbf{r}} \ A$.*

Proposition 3.5.8. *If both A and B contain neither a (free or bound) second-order variable nor $\bot$ then,* $\mathrm{HA}^2_{n^*} \vdash (x \ \underline{\mathbf{r}} \ (A \to B)) \leftrightarrow ((x \ \underline{\mathbf{r}} \ A) \to (x \ \underline{\mathbf{r}} \ B))$.

Proof. We have $x \ \underline{\mathbf{r}} \ (A \to B) \equiv \forall z.(z \ \underline{\mathbf{r}} \ A) \to (\ulcorner \dot{x}\dot{z} \urcorner \ \underline{\mathbf{r}} \ B) \equiv \forall z.(x \ \underline{\mathbf{r}} \ A) \to (x \ \underline{\mathbf{r}} \ B)$, where the last equality is due to Proposition 3.5.7. Since z does not occur in the scope of the $\forall z$-quantification, the claimed equivalence is a logical one. □

Proposition 3.5.9. *If A is one of the arithmetical axioms of Definition 1.1.2, but different from $\mathcal{S}t = 0 \to \bot$, then $A \to (x \ \underline{\mathbf{r}} \ A)$.*

Proof. By Proposition 3.5.8 and minimal logic. □

Proposition 3.5.10. $\mathrm{HA}^2_{0,(1)} \vdash (\mathcal{S}t = 0 \to \bot) \to (x \ \underline{\mathbf{r}} \ (\mathcal{S}t = 0 \to \bot))$.

Proof. $x \ \underline{\mathbf{r}} \ (\mathcal{S}t = 0 \to \bot) \equiv \forall z.\neg\neg(\mathcal{S}t = 0) \to \mathcal{A}(\ulcorner \dot{t}\dot{z} \urcorner)$ and the claim follows by minimal logic from $\mathcal{S}t = 0 \to \bot$, since $\mathcal{A}(z)$ was assumed to end in $\bot$. □

The next lemma should be thought of as an "internal version" of Corollary 3.4.22.

Lemma 3.5.11. $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash (t \ \underline{\mathbf{r}} \ \mathbb{N}x) \to \ulcorner \dot{t}(\lambda \mathtt{nfs.f(nfs)})c_0 \urcorner \sim \ulcorner c_{\dot{x}} \urcorner$.

Proof. See Remark 3.5.2 for the expanded formula. We argue informally within $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$.

Assume $t \ \underline{\mathbf{r}} \ \mathbb{N}x$. Instantiate $\mathfrak{X}_0$ to $\mathcal{B}(x) \equiv (x)_0 \sim \ulcorner c_{(\dot{x})_1} \urcorner$, that is, to the formula expressing that the realizer should be the Church numeral corresponding to the realized number. If for y we chose $\ulcorner \lambda \mathtt{nfs.f(nfs)} \urcorner$, that is, the successor on Church numerals, then the first premise becomes true. Here we used $\ulcorner c_{\dot{n}}\dot{y} \urcorner \sim \ulcorner \lambda x.\dot{y}^{\dot{n}}x \urcorner$. If for z we chose $\ulcorner c_0 \urcorner$ the second premise also becomes true, hence the claim. □

Proposition 3.5.12. $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash \mathbb{N}n \to \ulcorner c_{\dot{n}} \urcorner \ \underline{\mathbf{r}} \ \mathbb{N}n$.

Proof. See Remark 3.5.2 for the expanded formula. Argue informally within $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$ and use the induction principle provided by the assumption $\mathbb{N}n$. The claim follows from $\ulcorner c_0\dot{y}\dot{z} \urcorner \sim z$ and $\ulcorner c_{\dot{n}+1}\dot{y}\dot{z} \urcorner \sim \ulcorner \dot{y}(c_{\dot{n}}\dot{y}\dot{z}) \urcorner$. For the second β-equality note that both terms are in fact β-equal to $\ulcorner \dot{y}^{\dot{n}+1}\dot{z} \urcorner$. □

A simple induction on A shows

Lemma 3.5.13. *$(t \underline{\mathbf{r}}\, A)\sigma \equiv t\sigma \,\underline{\mathbf{r}}\, A\sigma$, provided the free variables of $\mathcal{A}$ are not in the domain of σ.*

Due to the subtleties of β-equality we do not get syntactical equality in the next lemma. But at least we do get provable equivalence.

Lemma 3.5.14. *For formulae A and $\mathcal{F}(x)$, provably in $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$, we have $(t \,\underline{\mathbf{r}}\, A)[\mathfrak{X}{:=}(\cdot)_0 \,\underline{\mathbf{r}}\, \mathcal{F}((\cdot)_1)] \leftrightarrow t \,\underline{\mathbf{r}}\, (A[\mathfrak{X}{:=}\mathcal{F}])$ provided that $\mathfrak{X}$ is not free in $\mathcal{A}(z)$.*

Proof. By induction on A we construct a proof in $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$.

The cases $\bot$ and $R(\vec{t})$ are trivial, as the substitution has no effect.

In the case $A \equiv \mathfrak{X}s$ the left hand side becomes $\exists z.x \sim t \wedge (x \,\underline{\mathbf{r}}\, \mathcal{F}(s))$ while the right hand side becomes $t \,\underline{\mathbf{r}}\, \mathcal{F}(s)$. The implication "$\leftarrow$" is trivial while for the implication "$\rightarrow$" we use Proposition 3.5.6 and Remark 3.5.3 to handle the classical existential quantifier.

The remaining cases are immediate from the induction hypotheses. ☐

Corollary 3.5.15. $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash (t \,\underline{\mathbf{r}}\, \forall\mathfrak{X}A) \rightarrow (t \,\underline{\mathbf{r}}\, (A[\mathfrak{X}{:=}\mathcal{F}]))$.

Proof. $(t \,\underline{\mathbf{r}}\, \forall\mathfrak{X}A) \equiv \forall\mathfrak{X}.t \,\underline{\mathbf{r}}\, A$. Arguing informally within $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$ the assumption $t \,\underline{\mathbf{r}}\, \forall\mathfrak{X}A$ implies $(t \,\underline{\mathbf{r}}\, A)[\mathfrak{X}{:=}(\cdot)_0 \,\underline{\mathbf{r}}\, \mathcal{F}((\cdot)_1)]$ and Lemma 3.5.14 yields the claim. ☐

Theorem 3.5.16. *If $t \gg \vec{x} \colon \vec{A} \vdash A$ then*

$$\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash (\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow (\ulcorner t[\vec{x}{:=}\dot{\vec{y}}]\urcorner \,\underline{\mathbf{r}}\, A)$$

where $\vec{y}$ are pairwise distinct variables not occurring in $A, \vec{A}$.

Proof. Induction on the Definition 3.4.4 of $t \gg \vec{x} \colon \vec{A} \vdash A$.

The case of an axiom is handled by Propositions 3.5.9 and 3.5.10. The case of an assumption is trivial.

In the case of $\rightarrow$-introduction we have $\forall y, \vec{y}.(y \,\underline{\mathbf{r}}\, A) \rightarrow (\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow (\ulcorner t[x, \vec{x}{:=}\dot{x}, \dot{\vec{y}}]\urcorner \,\underline{\mathbf{r}}\, B)$ by induction hypothesis. By logic we obtain $\forall\vec{y}.(\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow \forall y.(y \,\underline{\mathbf{r}}\, A) \rightarrow (\ulcorner t[x, \vec{x}{:=}\dot{x}, \dot{\vec{y}}]\urcorner \,\underline{\mathbf{r}}\, B)$. Since $\ulcorner t[x, \vec{x}{:=}\dot{y}, \dot{\vec{y}}]\urcorner \sim \ulcorner(\lambda x.t[\vec{x}{:=}\dot{\vec{y}}])\dot{y}\urcorner$ the claim follows from Proposition 3.5.6.

In the case of an $\rightarrow$-elimination we have $(\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow \forall z.(z \,\underline{\mathbf{r}}\, A) \rightarrow (\ulcorner t[\vec{x}{:=}\dot{\vec{y}}]\dot{z}\urcorner \,\underline{\mathbf{r}}\, B)$ and $(\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow (\ulcorner s[\vec{x}{:=}\dot{\vec{y}}]\urcorner \,\underline{\mathbf{r}}\, A)$ by the induction hypotheses, so the claim $(\vec{y} \,\underline{\mathbf{r}}\, \vec{A}) \rightarrow (\ulcorner (ts)[\vec{x}{:=}\dot{\vec{y}}]\urcorner \,\underline{\mathbf{r}}\, B)$ follows purely logically, since $\ulcorner(t[\vec{x}{:=}\dot{\vec{y}}])(s[\vec{x}{:=}\dot{\vec{y}}])\urcorner = \ulcorner(ts)[\vec{x}{:=}\dot{\vec{y}}]\urcorner$.

The case of a $\forall x$-introduction is purely logical. The case of a $\forall x$-elimination is handled by Lemma 3.5.13. The cases of $\wedge$-introduction and $\wedge$-elimination are handled similar to those of $\rightarrow$-introduction and $\rightarrow$-elimination. The case of $\forall\mathfrak{X}$-introduction follows by logic and the case of an $\forall\mathfrak{X}$-elimination is handled by Lemma 3.5.14. ☐

Remark 3.5.17. $t \,\underline{\mathbf{r}}\, (\neg\forall y(\mathbb{N}y \rightarrow \neg R(s, y))) \equiv$
$\forall z.(\forall y \forall z'.(z' \,\underline{\mathbf{r}}\, \mathbb{N}y) \rightarrow \forall z''.\neg\neg R(s, y) \rightarrow \mathcal{A}(\ulcorner\dot{z}\dot{z}'\dot{z}''\urcorner)) \rightarrow \mathcal{A}(\ulcorner t\dot{z}\urcorner)$.

Proof. $t \ \underline{\mathbf{r}}\ (\neg\forall y(\mathbb{N}y \to \neg R(s,y))) \equiv \forall z.(\forall y.z\ \underline{\mathbf{r}}\ (\mathbb{N}y \to \neg R(s,y))) \to \mathcal{A}(\ulcorner t\dot{z}\urcorner) \equiv \forall z.(\forall y \forall z'.(z'\ \underline{\mathbf{r}}\ \mathbb{N}y) \to (\ulcorner \dot{z}\dot{z}'\urcorner\ \underline{\mathbf{r}}\ \neg R(s,y))) \to \mathcal{A}(\ulcorner t\dot{z}\urcorner) \equiv \forall z.(\forall y \forall z'.(z'\ \underline{\mathbf{r}}\ \mathbb{N}y) \to \forall z''.\neg\neg R(s,y) \to \mathcal{A}(\ulcorner \dot{z}\dot{z}'\dot{z}''\urcorner)) \to \mathcal{A}(\ulcorner t\dot{z}\urcorner)$. ∎

Proposition 3.5.18. $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$ *proves*
$\forall y \forall z'.(z'\ \underline{\mathbf{r}}\ \mathbb{N}y) \to \forall z''.\neg\neg R(s,y) \to \exists y'.(\ulcorner(\lambda z'z''.z')\dot{z}'\dot{z}''\urcorner\ \underline{\mathbf{r}}\ \mathbb{N}y') \wedge R(s,y')$.

Proof. We argue informally in $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$. Since $\ulcorner(\lambda z'z''.z')\dot{z}'\dot{z}''\urcorner \sim z'$ by Proposition 3.5.6, the statement $\ulcorner(\lambda z'z''.z')\dot{z}'\dot{z}''\urcorner\ \underline{\mathbf{r}}\ \mathbb{N}y'$ follows from $z'\ \underline{\mathbf{r}}\ \mathbb{N}y'$. Hence under the given assumptions we can use y as a witness. The rest follows by logic, noting that the existential quantifier is a defined one. ∎

Note that in the following proposition the variable x occurs free in the formula $\mathcal{A}$, the premise, and the concluded formula.

Proposition 3.5.19. *If we choose* $\mathcal{A}(z)$ *to be* $\exists y.(z\ \underline{\mathbf{r}}\ \mathbb{N}y) \wedge R(x,y)$ *then* $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$ *proves* $(t\ \underline{\mathbf{r}}\ (\neg\forall y(\mathbb{N}y \to \neg R(x,y)))) \to \exists y.(\ulcorner \dot{t}(\lambda z'z''.z')\urcorner\ \underline{\mathbf{r}}\ \mathbb{N}y) \wedge R(x,y)$.

Proof. By Remark 3.5.17 we know the shape of the given formula. Instantiating it to $\ulcorner \lambda z'z''.z'\urcorner$, the premise of the remaining implication becomes provable, as shown in Proposition 3.5.18. Hence the claim. ∎

Theorem 3.5.20. *If* $\widetilde{\mathrm{HA}}^2_{n^*} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$ *then there is a* $t \in \mathrm{F}^{\times}_{n^*}$ *such that* $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash \forall x.\mathbb{N}x \to \exists y.\ulcorner tc_{\dot{x}}\urcorner \sim \ulcorner c_{\dot{y}}\urcorner \wedge R(x,y)$.

Proof. Assume $\widetilde{\mathrm{HA}}^2_{n^*} \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$. By Proposition 3.4.8, there is a term $t \in \mathrm{F}^{\times}_{n^*}$ such that $t \gg \emptyset \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$. As in Proposition 3.5.19 we choose the formula $\mathcal{A}(z)$ in Definition 3.5.1 to be $\exists y.(z\ \underline{\mathbf{r}}\ \mathbb{N}y) \wedge R(x',y)$. Note that this formula ends in $\bot$.

By Theorem 3.5.16 we get $\mathrm{HA}^2_{n^*} \vdash t\ \underline{\mathbf{r}}\ (\forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y)))$. Note that $t\ \underline{\mathbf{r}}\ (\forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))) \equiv \forall xz.(z\ \underline{\mathbf{r}}\ \mathbb{N}x) \to \ulcorner t\dot{z}\urcorner\ \underline{\mathbf{r}}\ (\neg\forall y(\mathbb{N}y \to \neg R(x,y)))$. By Proposition 3.5.12 we obtain $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash \forall x.\mathbb{N}x \to \ulcorner tc_{\dot{x}}\urcorner\ \underline{\mathbf{r}}\ (\neg\forall y(\mathbb{N}y \to \neg R(x,y)))$.

Note that the proven formula still contains the variable x' free. Instantiating the outermost quantifier to x' and then universally quantifying binds the x' to be the outermost $\forall x \ldots$ and by Proposition 3.5.19 we know $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash \forall x.\mathbb{N}x \to \exists y.\ulcorner tc_{\dot{x}}(\lambda z'z''.z')\urcorner\ \underline{\mathbf{r}}\ \mathbb{N}y \wedge R(x,y)$ and so, by Proposition 3.5.11 we obtain $\mathrm{HA}^2_{\mathrm{n}^*,(2)} \vdash \forall x.\mathbb{N}x \to \exists y.\ulcorner tc_{\dot{x}}(\lambda z'z''.z')(\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0\urcorner \sim \ulcorner c_{\dot{y}}\urcorner \wedge R(x,y)$.

So we have found a term for the computation. Concerning its typing we know by Proposition 3.4.7 that, in an appropriate context, t is typable and has type $\{\!\!\{\forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))\}\!\!\} = \mathcal{N} \to (\mathcal{N} \to \mathcal{N} \to \mathcal{N}) \to \mathcal{N}$, thus $tc_{\dot{x}}(\lambda z'z''.z')$ has type $\mathcal{N}$. Since $(\lambda\mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))$ has type $\mathcal{N} \to \mathcal{N}$ and c_0 has type $\mathcal{N}$, we can assign $tc_{\dot{x}}(\lambda z'z''.z')(\lambda\mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0$ the type $\mathcal{N}$. So the term $\lambda x.tx(\lambda z'z''.z')(\lambda\mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0$, computing the witnesses has type $\mathcal{N} \to \mathcal{N}$ as desired. ∎

Corollary 3.5.21. *The statements of Corollary 1.6.19 are also equivalent to*

$(**)$ $\mathrm{HA}^2_{\mathrm{n}^*,(1)} \vdash \forall x.\mathbb{N}x \to \exists y.\ulcorner tc_{\dot{x}}\urcorner \sim \ulcorner c_{\dot{y}}\urcorner \wedge R(x,y)$ *for some* $t \in \mathrm{F}^{\times}_{n^*}$.

Proof. $(\gamma)\Rightarrow(**)$ by Theorem 3.5.20. $(**)\Rightarrow(*)$ by logic, where $(*)$ is the equivalent statement of Corollary 2.2.22. ∎

Remark 3.5.22. In perfect analogy to Remark 3.4.24 we note that, as a theorem of $\mathrm{HA}^2_{\mathrm{n}^*,(1)}$, the realizing term, although containing free variables for the axioms, maps Church numerals to Church numerals, that is, to closed terms. So the axioms have no computational content, as they are completely discarded during the computation (for otherwise the result would contain a free variable).

Remark 3.5.23. $\forall n. \ulcorner c_{\dot{n}}(\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0 \urcorner \sim \ulcorner c_{\dot{n}} \urcorner$.

Proof. $\ulcorner c_{\dot{n}}(\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0 \urcorner \sim \ulcorner (\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))^{\dot{n}} c_0 \urcorner$. Induction on n shows $\ulcorner (\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))^{\dot{n}} c_0 \urcorner \sim \ulcorner c_{\dot{n}} \urcorner$. ∎

Remark 3.5.24 (Extensional equality of the realizers). In Corollaries 3.4.25 and 3.5.21 we have shown two statements claiming the existence of an $\mathrm{F}^{\times}_{n+1}$-term that computes a function shown to be recursive in HA^2_{n+1}. It is worth noting that the two approaches yield extensionally the same realizing function.

In fact, the proof of Theorem 3.4.23 shows that, if $\tilde{t}$ is the term canonically representing the proof, that is, if

$$\tilde{t} \gg \emptyset \vdash \forall x.\mathbb{N}x \to \neg\forall y(\mathbb{N}y \to \neg R(x,y))$$

then $\lambda x.\tilde{t}x(\lambda z'z''.z')$ is a realizer. On the other hand, the realizer constructed in the proof of Theorem 3.5.20 is $\lambda x.\tilde{t}x(\lambda z'z''.z')(\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0$, again with $\tilde{t}$ a notation of the proof.

So, if $\lambda x.\tilde{t}x(\lambda z'z''.z')$ maps Church numerals to Church numerals, then it denotes the same function as $\lambda x.\tilde{t}x(\lambda z'z''.z')(\lambda \mathtt{nfs}.\mathtt{f}(\mathtt{nfs}))c_0$ by Remark 3.5.23. Moreover, by Theorem 3.4.23 PRA + "$\mathrm{F}^{\times}_{n+1}$ weakly normalizing" proves that $\lambda x.\tilde{t}x(\lambda z'z''.z')$ actually maps Church numerals to Church numerals.

4 The One Variable Fragment $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$

The limit of the systems studied is $\mathrm{ID}^c_{<\omega}$. We will now give a simple description of a fragment of analysis of that strength. The only restriction we add is that the language should have just one set variable $\widehat{\mathfrak{X}}$. Specifically, a formula $\forall\widehat{\mathfrak{X}}A$ has *no* free second order variable. On the other hand, arbitrary *nestings* of second order variables are possible.

Axioms and rules are those of second order arithmetic for the restricted language. In particular we have full comprehension for this language. As this system has strength $\mathrm{ID}^c_{<\omega}$ we have obtained a proof theoretically manageable fragment of second order arithmetic by restricting the *language* rather than comprehension. Nevertheless, comprehension is implicitly restricted due to our restriction of language: we cannot even think about weird comprehension formulae!

Conceptually, this result provides some justification, why an extra amount of impredicativity appears when stepping beyond $\mathrm{ID}^c_{<\omega}$, which sometimes is considered as the supremum of theories constructed "from below". As soon as one is no longer obliged to speak about one set at a time, but has the possibility to compare two sets, one definitely has to conceive them as completed totalities. This should be compared to Gentzen's distinction [21, S. 524] between „aktuale Unendlichkeit" and „potentielle Unendlichkeit".

Mathematically, this result disproves the tempting conjecture that the chosen approach trivially extends to transfinitely iterated inductive definitions by just adding more (transfinite) levels. The reason is, that adding further levels will not lead beyond the one variable fragment.

4.1 Definition and Lower Bound

Definition 4.1.1. $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ is second order arithmetic based on minimal logic with the following restriction of the language: there is only one second order variable $\widehat{\mathfrak{X}}$. Let $\mathrm{HA}^2_{\widehat{\mathfrak{X}},(k)}$ be $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ with induction for Π^0_k-formulae added.

Remark 4.1.2. An "α-conversion free" formulation of this definition would require, that the de Bruijn representation [10] of every formulae uses just the index 0 for second order variables.

First we show that all the theories considered so far can indeed be formalized in $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$. This, is not as trivial as it may seem. Just replacing $\mathfrak{X}_k$ by $\widehat{\mathfrak{X}}$ in a proof of HA^2_n will, in general, not yield a valid proof again, since in the premise of a $\forall\widehat{\mathfrak{X}}$-introduction rule the variable condition might be violated through the identification of second order variables. Fortunately, we know an equivalent system, at least for Π^0_2-formulae. Recalling the proofs of the embedding of ID^-_n into HA^2_{n+1} we note that replacing all $\mathfrak{X}_k$'s by $\widehat{\mathfrak{X}}$ in the proofs obtained *that way*, *will* indeed yield a valid proof.

In the following proposition we use A^* in the sense of Definition 1.4.1, however read as α-variant in the sense of Remark 1.3.6.

Proposition 4.1.3. *If* $\mathrm{ID}^-_n \vdash A$ *then* $\mathrm{HA}^2_{\widehat{\mathfrak{X}}} \vdash A^*$.

Proof. Inspection of the proofs of 1.3.9, 1.3.10, 1.4.5, 1.4.7 and 1.4.9 shows that they remain valid if every $\mathfrak{X}_i$ is replaced by $\widehat{\mathfrak{X}}$. ∎

4.2 Leveling

We now aim at the converse implication. More precisely, we will show (in Corollary 4.2.11) that every formula provable in $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ is already provable in some HA^2_n. To do so, we assign levels to the different occurrences of the second order variables. In order to avoid arguing "up to α-conversion" we assume some canonical choice $\mathfrak{X}_0$, $\mathfrak{X}_1$, ... of set variables of levels 0, 1, ...

Definition 4.2.1. By induction on the $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$-formula A we define its level $\mathrm{lv}(A)$.

- $\mathrm{lv}(\mathrm{R}\vec{t}) = \mathrm{lv}(\widehat{\mathfrak{X}}t) = 0$
- $\mathrm{lv}(\forall x A) = \mathrm{lv}(A)$
- $\mathrm{lv}(A \wedge B) = \mathrm{lv}(A \to B) = \max\{\mathrm{lv}(A), \mathrm{lv}(B)\}$
- $\mathrm{lv}(\forall \widehat{\mathfrak{X}} A) = 1 + \mathrm{lv}(A)$

Definition 4.2.2. By induction on the $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$-formula A we define its leveled α-variant $\widetilde{A}$.

- $\widetilde{\widehat{\mathfrak{X}}t} = \widehat{\mathfrak{X}}t$
- $\widetilde{\mathrm{R}\vec{t}} = \mathrm{R}\vec{t}$
- $\widetilde{A \wedge B} = \widetilde{A} \wedge \widetilde{B}$
- $\widetilde{A \to B} = \widetilde{A} \to \widetilde{B}$
- $\widetilde{\forall x A} = \forall x \widetilde{A}$
- $\widetilde{\forall \widehat{\mathfrak{X}} A} = \forall \mathfrak{X}_{\mathrm{lv}(A)}.\widetilde{A}[\widehat{\mathfrak{X}} := \mathfrak{X}_{\mathrm{lv}(A)}]$

By a simple induction on A we obtain

Proposition 4.2.3. $\widetilde{A} \in \mathcal{I}^2_{\mathrm{lv}(A)}[\widehat{\mathfrak{X}}]$.

Next we define a level-aware formalization of the proofs of $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$. In the following definition one should think of the n as some canonical decoration of the usual formal proofs in $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$. In fact, such a decoration always exists, as we will show in Lemma 4.2.9.

Definition 4.2.4. The relation $\Gamma \vdash A :: [n]$ is inductively defined as follows.

- If $A \in \mathcal{I}^2_n[\widehat{\mathfrak{X}}]$, $\Gamma \subset \mathcal{I}^2_n[\widehat{\mathfrak{X}}]$, and A is an axiom or $A \in \Gamma$ then $\Gamma \vdash A :: [n]$.
- If $\Gamma \vdash A :: [n]$ and $\Gamma \vdash B :: [n]$ then $\Gamma \vdash A \wedge B :: [n]$.
- If $\Gamma \vdash A \wedge B :: [n]$ then $\Gamma \vdash A :: [n]$ and $\Gamma \vdash B :: [n]$.
- If $\Gamma \vdash A :: [n]$ with x not free in Γ then $\Gamma \vdash \forall x A :: [n]$.
- If $\Gamma \vdash \forall x A(x) :: [n]$ then $\Gamma \vdash A(t) :: [n]$.
- If $\Gamma, A \vdash B :: [n]$ then $\Gamma \vdash A \to B :: [n]$.
- If $\Gamma \vdash A \to B :: [n]$ and $\Gamma \vdash A :: [n]$ then $\Gamma \vdash B :: [n]$.

- If $\Gamma \vdash A(\widehat{\mathfrak{X}}) :: [n]$ with $\widehat{\mathfrak{X}}$ not free in Γ, $n > \mathrm{lv}(A)$ and $\Delta \subset \mathcal{I}^2_n[\widehat{\mathfrak{X}}]$ then $\Gamma, \Delta \vdash \forall \mathfrak{X}_{\mathrm{lv}(A)}.A(\mathfrak{X}_{\mathrm{lv}(A)}) :: [n]$.
- If $\Gamma \vdash \forall \mathfrak{X}_m A(\mathfrak{X}_m) :: [n]$, $\mathcal{A} \in \mathcal{I}^2_k[\widehat{\mathfrak{X}}]$ and $n' \geq n, m+1, k$ then $\Gamma \vdash A(\mathcal{A}) :: [n']$.

Remark 4.2.5. Definition 4.2.4 simultaneously handles $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ and all the systems $\mathrm{HA}^2_{\widehat{\mathfrak{X}},(k)}$, depending on how the notion "A is an axiom" is read: either only restricted to one of the axioms of Definition 1.1.2 or including the induction axioms of Definition 2.2.1.

As we will only establish relations between fragments of analysis and their corresponding notion of $\cdot \vdash \cdot :: [\cdot]$, no confusion between the various readings of this relation should arise.

By a simple induction on the definition we obtain

Proposition 4.2.6. *If $\Gamma \vdash A :: [n]$ and $n \leq n'$ then $\Gamma \vdash A :: [n']$.*

Since our axioms are closed under substitution, a simple induction shows

Proposition 4.2.7. *If $\Gamma \vdash A :: [n]$ then $\Gamma\sigma \vdash A\sigma :: [n]$.*

Proposition 4.2.8. *If $\Gamma \vdash A :: [n]$ and $\Gamma \subset \Gamma' \subset \mathcal{I}^2_n[\widehat{\mathfrak{X}}]$ then $\Gamma' \vdash A :: [n]$.*

Proof. Induction on Definition 4.2.4 of $\cdot \vdash \cdot :: [\cdot]$, using Proposition 4.2.7 for a possibly necessary renaming of free object variables in the case of $\forall x$-introduction. ∎

Lemma 4.2.9. *If $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ (or $\mathrm{HA}^2_{\widehat{\mathfrak{X}},(k)}$ respectively) proves the sequence $\Gamma \vdash A$ then there is an n such that $\widetilde{\Gamma} \vdash \widetilde{A} :: [n]$.*

Proof. Induction on the notion of a proof in $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$. For the case of an axiom or assumption use Proposition 4.2.3 to show that $\widetilde{A}$ is in the correct language. Then every proof rule can be reproduced identically, using weakening (provided by Propositions 4.2.6 and 4.2.8) where necessary. ∎

Proposition 4.2.10. *If $\Gamma \vdash A :: [n]$ then HA^2_n (or $\mathrm{HA}^2_{\mathrm{n},(\mathrm{k})}$ respectively) proves the sequence $\Gamma[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n] \vdash A[\widehat{\mathfrak{X}}{:=}\mathfrak{X}_n]$.*

Proof. Induction on $\Gamma \vdash A :: [n]$.

The only non-trivial case is that of an $\forall\widehat{\mathfrak{X}}$-introduction. From the premise $\Gamma \vdash A(\widehat{\mathfrak{X}}) :: [n]$ with $\widehat{\mathfrak{X}}$ not free in Γ, we obtain a HA^2_n-proof (or a $\mathrm{HA}^2_{\mathrm{n},(\mathrm{k})}$-proof) of $\Gamma \vdash A(\mathfrak{X}_n)$ by induction hypothesis. Arguing within HA^2_n (or $\mathrm{HA}^2_{\mathrm{n},(\mathrm{k})}$) we conclude $\Gamma \vdash \forall \mathfrak{X}_n.A(\mathfrak{X}_n)$. Since $A(\mathfrak{X}_{\mathrm{lv}(A)})$ is a legal formula we may conclude $\Gamma \vdash A(\mathfrak{X}_{\mathrm{lv}(A)})$ and from that $\Gamma \vdash \forall \mathfrak{X}_{\mathrm{lv}(A)}.A(\mathfrak{X}_{\mathrm{lv}(A)})$. Finally, weakening yields a proof of $\Gamma, \Delta \vdash \forall \mathfrak{X}_{\mathrm{lv}(A)}.A(\mathfrak{X}_{\mathrm{lv}(A)})$. ∎

Corollary 4.2.11. $\mathrm{HA}^2_{\widehat{\mathfrak{X}}} \vdash A$ *if and only if there is an n such that* $\mathrm{HA}^2_n \vdash \widetilde{A}$. *Moreover,* $\mathrm{HA}^2_{\widehat{\mathfrak{X}},(k)} \vdash A$ *if and only if there is an n such that* $\mathrm{HA}^2_{\mathrm{n},(\mathrm{k})} \vdash \widetilde{A}$.

Proof. From 4.2.9 and 4.2.10. For the other direction use 4.1.3. ∎

Corollary 4.2.12. *The functions provably recursive in* $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ *or* $\mathrm{HA}^2_{\widehat{\mathfrak{X}},(k)}$ *are precisely those provably recursive in* $\mathrm{ID}^c_{<\omega}$.

Proof. Using Corollary 1.6.19 and Corollary 2.2.22 the claim follows from Corollary 4.2.11. □

4.3 The Corresponding Fragment of System F

Adapting the proofs of Subsection 3.1 one can easily come up with a proof of normalization within $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$ for the type system

$$\rho ::= \alpha \mid \rho \to \sigma \mid \rho \times \sigma \mid \forall \alpha \rho$$

where the case $\forall\alpha\rho$ is subject to the proviso $\mathrm{FV}(\rho) \subset \{\alpha\}$.

However, we can do even slightly better. Inspecting Definition 3.1.5 of the computability predicates and the way they are used, one notes that the only property necessary to ensure formalizability in the one variable fragment was that universal types are closed.

Thus we can also allow the slightly liberalized type grammar

$$\rho ::= \alpha \mid \rho \to \sigma \mid \rho \times \sigma \mid \forall \vec{\alpha} \rho$$

where in the case of universal types we have the proviso $\mathrm{FV}(\forall\vec{\alpha}\rho) = \emptyset$.

The typing rules are essentially usual ones, however confined to the restricted language. That is, a block of universal quantifiers has to be instantiated simultaneously, and, when universally quantifying, all the free type variables have to be bound simultaneously.

We can then define computability predicates in the following way.

$$\begin{array}{lcl}
\Phi^{\alpha}(x) & = & \exists y. y \sim x \wedge \widehat{\mathfrak{X}}\langle \ulcorner \alpha \urcorner, y\rangle \\
\Phi^{\rho\times\sigma}(x) & = & \Phi^{\rho}(\ulcorner \dot{x}\mathrm{L} \urcorner) \wedge \Phi^{\sigma}(\ulcorner \dot{x}\mathrm{R} \urcorner) \\
\Phi^{\rho\to\sigma}(x) & = & \forall y. \Phi^{\rho}(y) \to \Phi^{\sigma}(\ulcorner \dot{x}\dot{y} \urcorner) \\
\Phi^{\forall\vec{\alpha}.\rho}(x) & = & \forall \widehat{\mathfrak{X}} \Phi^{\rho}(x)
\end{array}$$

The reader is invited to check that all the proofs can be easily adapted using predicates of the form $\widehat{\mathfrak{X}}x \equiv ((x)_0 = \ulcorner \alpha \urcorner \to \Phi^{\rho}((x)_1)) \wedge \ldots$

As, however, this more liberal way of typing lacks interesting examples we do not exhibit further details.

4.4 On Counting Variables

After the analysis of second order arithmetic restricted to a single set variable it might be tempting to study fragments with two, three, ... variables. Yet, as opposed to $\mathrm{HA}^2_{\widehat{\mathfrak{X}}}$, this would not lead to natural systems, since the property of containing only two variables is *not* closed under substitution. Even if we restrict our formulae to contain only two fixed variables $\widehat{\mathfrak{X}}$ and $\widehat{\mathfrak{Y}}$, the renaming of bound variables that occurs due to substitution gives rise to new variables. It should be noted, however, that it is very easy to define computability predicates (in the style of Definition 3.1.5) for full system F with only these two variables, and in fact even with only the indices 0 and 1 for second order variables occurring in de Bruijn [10] notation:

By induction on ρ we define formulae $\Phi^{\rho}_{\widehat{\mathfrak{X}}}(x)$ and $\Phi^{\rho}_{\widehat{\mathfrak{Y}}}(x)$ with free second order variables $\{\widehat{\mathfrak{X}}\}$ and $\{\widehat{\mathfrak{Y}}\}$ respectively.

$$
\begin{array}{lcl}
\Phi^{\alpha}_{\widehat{\mathfrak{X}}}(x) & = & \exists y . y \sim x \wedge \widehat{\mathfrak{X}} \left\langle \ulcorner \alpha \urcorner, y \right\rangle \\
\Phi^{\rho \times \sigma}_{\widehat{\mathfrak{X}}}(x) & = & \Phi^{\rho}_{\widehat{\mathfrak{X}}}(\ulcorner \dot{x}\mathtt{L} \urcorner) \wedge \Phi^{\sigma}_{\widehat{\mathfrak{X}}}(\ulcorner \dot{x}\mathtt{R} \urcorner) \\
\Phi^{\rho \to \sigma}_{\widehat{\mathfrak{X}}}(x) & = & \forall y . \Phi^{\rho}_{\widehat{\mathfrak{X}}}(y) \to \Phi^{\sigma}_{\widehat{\mathfrak{X}}}(\ulcorner \dot{x}\dot{y} \urcorner) \\
\Phi^{\forall \alpha \rho}_{\widehat{\mathfrak{X}}}(x) & = & \forall \widehat{\mathfrak{Y}} . (\forall z . (z)_0 \neq \ulcorner \alpha \urcorner \to (\widehat{\mathfrak{X}} z \leftrightarrow \widehat{\mathfrak{Y}} z)) \to \Phi^{\rho}_{\widehat{\mathfrak{Y}}}(x)
\end{array}
$$

And $\Phi^{\rho}_{\widehat{\mathfrak{Y}}}(x)$ is defined as $\Phi^{\rho}_{\widehat{\mathfrak{X}}}(x)$ with the role of $\widehat{\mathfrak{X}}$ and $\widehat{\mathfrak{Y}}$ interchanged.

Still, the "obvious" proofs of basic properties such as coincidence (a predicate $\Phi^{\rho}_{\widehat{\mathfrak{X}}}$ depends only on the values of $\widehat{\mathfrak{X}} \left\langle \ulcorner \alpha \urcorner, \cdot \right\rangle$ with $\alpha \in \mathrm{FV}(\rho)$) or independence of the chosen α-variant need instantiations of $\forall \widehat{\mathfrak{X}}$-quantifications in such a way that the invoked substitutions require renaming of the bound variables. Modifications of the formulation of these properties or variants of the proofs suffer the same deficit.

References

[1] P. Aczel. An introduction to inductive definitions. In Barwise [6], chapter C.7, pages 739–782.

[2] T. Altenkirch and T. Coquand. A finitary subsystem of the polymorphic lambda-calculus. In S. Abramski, editor, *Proceedings of the 5th International Conference on Typed Lambda Caculi and Applications (TLCA 2001)*, volume 2044 of *Lecture Notes in Computer Science*, pages 22–28. Springer Verlag, 2001.

[3] T. Arai. A slow growing analogue to Buchholz' proof. *Annals of Pure and Applied Logic*, 54:101–120, 1991.

[4] H. Barendregt. The type free lambda calculus. In Barwise [6], chapter D.7, pages 1091–1132.

[5] H. Barendregt. Lambda calculi with types. In S. Abramsky, D. Gabbay, and T. Maibaum, editors, *Handbook of Logic in Computer Science*, volume 2, pages 117–309. Oxford University Press, 1992.

[6] J. Barwise, editor. *Handbook of Mathematical Logic*, volume 90 of *Studies in Logic and the Foundations of Mathematics.* North-Holland Publishing Company, 1977.

[7] S. Berardi, M. Bezem, and T. Coquand. On the computational content of the axiom of choice. *The Journal of Symbolic Logic*, 63(2):600–622, 1998.

[8] U. Berger and H. Schwichtenberg. An inverse of the evaluation functional for typed λ–calculus. In R. Vemuri, editor, *Proceedings of the Sixth Annual IEEE Symposium on Logic in Computer Science (LICS '91)*, pages 203–211, 1991.

[9] L. E. J. Brouwer. Zur Begründung der intuitionistischen Mathematik III. *Mathematische Annalen*, 96:451–488, 1927.

[10] N. G. de Bruijn. Lambda calculus notation with nameless dummies, a tool for automatic formula manipulation, with application to the Church–Rosser theorem. *Indagationes Mathematicae*, 34:381–392, 1972.

[11] W. Buchholz. The $\Omega_{\mu+1}$-rule. In *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies* [12], chapter IV, pages 189–233.

[12] W. Buchholz, S. Feferman, W. Pohlers, and W. Sieg. *Iterated Inductive Definitions and Subsystems of Analysis: Recent Proof-Theoretical Studies*, volume 897 of *Lecture Notes in Mathematics.* Springer Verlag, 1981.

[13] W. Buchholz and K. Schütte. *Proof Theory of Impredicative Subsystems of Analysis.* Studies in Proof Theory. Bibliopolis, 1988.

[14] S. Buss, editor. *Handbook of Proof Theory*, volume 137 of *Studies in Logic and the Foundation of Mathematics.* Elsevier, 1998.

[15] A. Church. An unsolvable problem of elementary number theory. *American Journal of Mathematics*, 58:345–363, 1936.

[16] A. Church and S. C. Kleene. Formal definitions in the theory of ordinal numbers. *Fundamenta Mathematicae*, 28:11–21, 1937.

[17] M. Fairtlough and S. Wainer. Hierarchies of provably recursive functions. In Buss [14], chapter III, pages 149–207.

[18] S. Feferman. Formal theories for transfinite iterations of generalized inductive definitions. In A. Kino, J. Myhill, and R. E. Vesley, editors, *Intuitionism and Proof Theory. Proceedings of the Summer Conference at Buffalo N.Y. 1968*, Studies in Logic and the Foundation of Mathematics, pages 303–326. North-Holland Publishing Company, 1970.

[19] G. Gentzen. Untersuchungen über das logische Schließen. I. *Mathematische Zeitschrift*, 39:176–210, 1935.

[20] G. Gentzen. Untersuchungen über das logische Schließen. II. *Mathematische Zeitschrift*, 39:405–431, 1935.

[21] G. Gentzen. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen*, 112:493–565, 1936.

[22] J.-Y. Girard. *Interprétation fonctionnelle et élimination des coupures dans l'arithmétique d'ordre supérieur.* Thèse de Doctorat d'État, Université de Paris VII, 1972.

[23] J.-Y. Girard, Y. Lafont, and P. Taylor. *Proofs and Types*, volume 7 of *Cambridge Tracts in Theoretical Computer Science*. Cambridge University Press, 1989.

[24] K. Gödel. Über formal unentscheidbare Sätze der Principia mathematica und verwandter Systeme I. *Monatshefte für Mathematik und Pysik*, 38:173–198, 1931.

[25] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunkts. *Dialectica*, 12:280–287, 1958.

[26] A. D. Gordon. A mechanisation of name-carrying syntax up to alpha-conversion. In J. J. Joyce and C.-J. H. Seger, editors, *Higher Order Logic Theorem Proving and Its Applications. 6th International Workshop, HUGS'93, Vancouver, B. C., Canada, August 1993*, volume 780 of *Lecture Notes in Computer Science*, pages 413–425. Springer Verlag, 1994.

[27] A. Grzegorczyk. Some classes of recursive fuctions. *Rozprawy Matematyczne*, 4, 1953.

[28] S. C. Kleene. The formal interpretation of intuitionistic number theory. *The Journal of Symbolic Logic*, 10(4):109–124, 1945.

[29] J. Krivine. *Lambda-calcul, types et modèles.* Masson, Paris, 1990.

[30] D. Leivant. Finitely stratified polymorphism. *Information and Computation*, 93:93–113, 1991.

[31] P. Martin-Löf. Hauptsatz for the intuitionistic theory of iterated inductive definitions. In J. E. Fenstad, editor, *Proceedings of the Second Scandinavian Logic Symposium (Oslo 1970)*, volume 63 of *Studies in Logic and the Foundation of Mathematics*, pages 179–216. North-Holland Publishing Company, 1971.

[32] R. Matthes. *Extensions of System F by Iteration and Primitive Recursion on Monotone Inductive Types*. PhD thesis, Fakultät für Mathematik und Informatik der Ludwig-Maximilians-Universität München, May 1998.

[33] P. Odifreddi. *Classical Recursion Theory. The Theory of Functions and Sets of Natural Numbers*, volume 125 of *Studies in Logic and the Foundation of Mathematics*. North-Holland Publishing Company, 1989.

[34] E. Palmgren. Type-theoretic interpretation of iterated, strictly positive inductive definitions. *Archive for Mathematical Logic*, 32:75–99, 1992.

[35] W. Pohlers. Subsystems of set theory and second order number theory. In Buss [14], chapter IV, pages 209–335.

[36] E. L. Post. Formal reductions of the general combinatorial decision problem. *American Journal of Mathematics*, 65:197–215, 1943.

[37] J. C. Reynolds. Towards a theory of type structure. In G. Goos and J. Harmanis, editors, *Programming Symposium*, volume 19 of *Lecture Notes in Computer Science*, pages 408–425. Springer Verlag, 1974.

[38] K. Schütte. Die unendliche Induktion in der Zahlentheorie. *Mathematische Annalen*, 122:369–389, 1951.

[39] W. W. Tait. Intensional interpretations of functionals of finite type. *The Journal of Symbolic Logic*, 32(2):198–212, 1967.

[40] G. Takeuti. On a generalized logic calculus. *Japanese Journal of Mathematics*, 23, 1953.

[41] G. Takeuti. Consistency proofs of subsystems of classical analysis. *Annals of Mathematics*, 86:299–348, 1967.

[42] A. S. Troelstra. Intuitionistic formal systems. In *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis* [43], chapter I, pages 1–96.

[43] A. S. Troelstra, editor. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, volume 344 of *Lecture Notes in Mathematics*. Springer Verlag, 1973.

[44] J. I. Zucker. Iterated inductive definitions, trees and ordinals. In Troelstra [43], chapter VI, pages 392–453.