TÉCNICO
LISBOA

On the herbrandised interpretation for nonstandard arithmetic

Ana de Almeida Gabriel Vieira Borges

Thesis to obtain the Master of Science Degree in

Mathematics and Applications

Supervisors: Prof. Fernando Ferreira
Prof. Ulrich Kohlenbach
Prof. Carlos Caleiro

Examination Committee

Chairperson: Prof. Cristina Sernadas
Supervisor: Prof. Fernando Ferreira
Member of the Committee: Prof. Reinhard Kahle

November 2016

Either mathematics is too big for the human mind, or the human mind is more than a machine.

Kurt Gödel

Acknowledgements

I am grateful to the many, many people who, in one way or another, made this thesis come to life.

In particular, I would like to thank Professor Fernando Ferreira for introducing me to the realm of proof theory and functional interpretations, for his guidance, for the long hours spent explaining one thing or another, for his patience, and for the several ideas and suggestions which improved this work.

I am also thankful to Professor Ulrich Kohlenbach for receiving me in Darmstadt, for sharing with me his enthusiasm for proof mining and the tools necessary for it, for the discussions of interesting ideas, and for the helpful comments he provided throughout the work.

I would like to thank Professor Carlos Caleiro, for introducing me to the field of functional logic and for nurturing my interest in the area. I am also very grateful for his availability regarding this thesis.

I thank Bruno Dinis for his help in better understanding nonstandard arithmetic, and Florian Steinberg for sharing his knowledge of Weihrauch reducibility.

I am grateful to all the friends who made sure I survived while working on this thesis. In particular, Raúl Penaguião was an invaluable catalyst from the day I started studying these matters, always knowing what to say in order to motivate me. He also had helpful suggestions regarding the readability of this work. Daniel Sousa was a great help during the writing phase, by frequently inquiring about my progress, by the many conversations which helped maintain my sanity, and finally by helping with the final revisions and some LaTeX troubles. Raquel Gonçalves was always available to discuss my work, and to be the (sometimes much needed) voice of reason. João Alírio was the perfect companion while writing, and always ready to debate anything that crossed my mind regarding what I was working on. Carolina Caldeirinha has been someone I can always count on for a long, long time, and corrected numerous grammatical mistakes (I'm sure there are still many more left, fully of my responsibility).

I would also like to thank Pedro Filipe, Marta Cruz, Victor Pecanins, Tomé Ribeiro, Maria do Mar da Câmara Pereira, Ana Galhoz, and Diogo Borges for the shared laughs and conversations, and for working beside me.

Finally, I am grateful to all my family. Each person offered to help and indeed helped by their constant support. In particular, my mother helped with revising the text, and by listening to my worries. My father always had an enthusiastic word. Together, they supported me throughout my education, and particularly made possible my stay in Lisbon during my undergraduate studies, and in Darmstadt while working on this thesis. My brother was always available for anything I needed. He cooked many dinners for me, and above all was always happy to listen to the "mathematical stuff" I found most fascinating at any given time.

Abstract

Functional interpretations are useful tools of proof theory. After Gödel described his *dialectica* interpretation for Heyting arithmetic in 1941, many other interpretations have been proposed, each focusing on different goals. We start with an overview of the interpretations of Gödel and Shoenfield.

We propose a functional interpretation for nonstandard Heyting arithmetic based on previous work by Van den Berg, Briseid and Safarik. This interpretation enables the transformation of proofs in nonstandard arithmetic of internal statements into proofs in standard arithmetic of those same statements. The witnesses for external, existential statements of the interpreting formulas are functions whose output is a finite sequence. Syntactically, the terms representing these functions are called end-star terms. It is possible to define a preorder of end-star terms. Our interpretation is monotone over this preorder: if a certain end-star term is a witness for an existential statement, then any "bigger" term also is. Using this property, we are able to prove a soundness theorem for our interpretation, which eliminates principles recognisable from nonstandard analysis. From this theorem, we get as corollary the conservativity of nonstandard arithmetic over standard arithmetic, as well as a term extraction theorem. It is also possible to prove a characterization theorem for our interpretation. As corollary, we show that the countable saturation principle does not add proof theoretical strength to our intuitionistic nonstandard system.

Finally, we give a short description of Weihrauch reducibility and comment on an application of Gödel's *dialectica* interpretation to, in certain circumstances, prove that a $\forall\exists$-formula Weihrauch reduces to another one.

Keywords

Functional interpretations, nonstandard arithmetic, Weihrauch reducibility.

Resumo

As interpretações funcionais são ferramentas úteis da teoria da demonstração. Depois de Gödel ter descrito a sua interpretação *dialectica* para a aritmética de Heyting, foram propostas muitas outras interpretações, cada uma com objectivos diferentes. Começamos por apresentar as interpretações de Gödel e Shoenfield.

Propomos uma interpretação funcional para a aritmética de Heyting não *standard*, baseada em trabalho de Van den Berg, Briseid e Safarik. Esta interpretação permite a transformação de provas na aritmética não *standard* de teoremas internos em provas na aritmética *standard* desses mesmos teoremas. As testemunhas para afirmações existenciais externas das fórmulas interpretadoras são funções cujo *output* é uma lista finita. Sintacticamente, os termos que representam estas funções são chamados termos *end-star*. É possível definir uma pré-ordem nos termos *end-star*. A nossa interpretação é monótona nesta pré-ordem: se um dado termo *end-star* é uma testemunha para uma afirmação existencial, então qualquer termo “maior” também o é. Usando esta propriedade, provamos a correcção da nossa interpretação, e eliminamos princípios reconhecíveis da análise não *standard*. Também obtemos como corolário que a aritmética não *standard* é conservativa sobre a aritmética *standard*, bem como um teorema de extracção de termos. É possível provar um teorema de caracterização para a nossa interpretação. Como corolário, mostramos que o princípio da saturação contável não acrescenta força ao nosso sistema intuicionista não *standard*.

Por fim, descrevemos brevemente a redutibilidade de Weihrauch e sugerimos uma aplicação da interpretação *dialectica* de Gödel para, em certas circunstâncias, decidir se uma fórmula-$\forall\exists$ se reduz-Weihrauch a outra.

Palavras-chave

Interpretações funcionais, aritmética não *standard*, redutibilidade de Weihrauch.

Contents

1

Introduction

Contents

1.1 Proof interpretations

Given formal systems $\mathcal{T}_1$ and $\mathcal{T}_2$, a proof interpretation I from $\mathcal{T}_1$ to $\mathcal{T}_2$ is a particular way of obtaining a formula A^I in the language of $\mathcal{T}_2$ from a formula A in the language of $\mathcal{T}_1$. Such an interpretation is defined by induction on the logical structure of A. After defining it, one should be able to prove a soundness theorem. Soundness theorems are loosely of the form:

Theorem (Soundness)**.** Let $\mathcal{C}$ and Δ be collections of principles, and Δ' a suitable modification of the principles Δ. For all formulas A of the language of $\mathcal{T}_1$, if $\mathcal{T}_1 + \mathcal{C} + \Delta \vdash A$, then $\mathcal{T}_2 + \Delta' \vdash A^I$.

The principles $\mathcal{C}$ are called characteristic principles. Characteristic principles of a proof interpretation vanish under it. In other words, for a characteristic principle C, it is possible to prove C^I in $\mathcal{T}_2$ without having access to C itself. The principles Δ are called side principles. Side principles don't vanish under I, and we need Δ' to prove their interpretations. However, the formulas in Δ' are usually similar to the formulas in Δ.

Soundness theorems are very useful. For example, as long as $\perp^I \equiv \perp$, which is usually the case, the soundness theorem for the I-interpretation entails the consistency of $\mathcal{T}_1 + \mathcal{C}$ relative to the consistency of $\mathcal{T}_2$.

There is another class of theorems related to proof interpretations: the characterization theorems. If $\mathcal{T}$ is a theory which encompasses the languages of $\mathcal{T}_1$ and $\mathcal{T}_2$, and $\mathcal{P}$ is a collection of principles, a characterization theorem is of the form:

Theorem (Characterization)**.** For all formulas A of the language of $\mathcal{T}_1$, we have that $\mathcal{T} + \mathcal{P} \vdash A \leftrightarrow A^I$.

It is interesting to look for a minimal collection of principles $\mathcal{P}$ which is enough to prove the characterization theorem for a given interpretation. These principles give a notion of the "distance" between the original formula A and its interpretation A^I.

The functional interpretations that are discussed in this work are much more specific than what is outlined above. The theories $\mathcal{T}_1$ and $\mathcal{T}_2$ are always theories of arithmetic that vary in two important ways: on the one hand, they feature either Heyting (intuitionistic) or Peano (classical) arithmetic; on the other hand, they focus on either standard or nonstandard arithmetic. By mixing these four possibilities, four important theories were defined: WE-HA$^\omega$, which we outline in Section 2.1.1; WE-PA$^\omega$, see Section 2.1.2; and E-HA$^{\omega*}_{\mathsf{st}}$ and E-PA$^{\omega*}_{\mathsf{st}}$, discussed in Section 3.1.2.

The "HA" and "PA" stand for Heyting and Peano arithmetic, respectively. The prefixes "WE" and "E" have to do with the treatment of equality: in the first case the theories are weakly extensional, while in the second case we allow full extensionality. The "ω" indicates that the theories are typed, and that we allow all finite types. The "st" means that the theories in question are in a nonstandard setting. Finally, the "$*$" heralds the use of the so-called star types (types for finite sequences), which will be necessary when we interpret nonstandard arithmetic.

1.2 Interpreting arithmetic

The first use of a functional interpretation was by Gödel [22] (later revised and translated into English as [23]). Gödel's functional interpretation was dubbed *dialectica* after the name of the journal in which it was originally published.

The *dialectica* interpreted Heyting arithmetic in a quantifier-free version of it. The goal was to show the consistency of number theory by "finitistic" means (see [44] for a discussion). Each arithmetical formula $A(\boldsymbol{a})$ was given a translation into a formula of the form $A(\boldsymbol{a})^D \equiv \exists \boldsymbol{x}\, \forall \boldsymbol{y}\, A_D(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a})$, with A_D quantifier-free and without $\vee$. As long as we knew a proof that $A(\boldsymbol{a})$ was intuitionistically true, the soundness theorem guaranteed that there existed closed, higher-typed, computable terms $\boldsymbol{t}$ such that $A_D(\boldsymbol{ta}, \boldsymbol{y}, \boldsymbol{a})$ was provable in a quantifier-free intuitionistic theory. Since $\bot^D \equiv \bot$, the soundness theorem was a consistency proof of Heyting arithmetic, as long as one accepted the consistency of the latter theory.

Nowadays, the consistency of widely used theories such as Heyting arithmetic is hardly in question (see [26] for a discussion). However, the *dialectica* interpretation and others like it are still useful as tools in proof theory for their *term extraction* functionalities: the terms $\boldsymbol{t}$ obtained from the soundness theorem are a codification of the proof of $A(\boldsymbol{a})$, and offer important insights into it. These terms can be used in practical applications in other areas of mathematics. In fact, there is a branch of proof theory, named "proof mining", which uses functional interpretations in a prominent way. These make it possible to extract quantitative information from certain proofs of theorems in analysis, for example. This extra information is often in the form of bounds on the growth rate of specific functions. For surveys of advances made in proof mining see [29, 30, 32].

In Section 2.2, we present the *dialectica* interpretation as an interpretation of WE-HA$^\omega$ in itself, following Chapter 8 of [29]. When seen from this point of view, the main goal is not to obtain a consistency proof, but to extract computable terms from the proof of A, which may give us insights into the said proof. The surveys [44, 13, 1] also give good overviews of the *dialectica* interpretation.

With the *dialectica* interpretation we are able to obtain information about Heyting arithmetic. What about Peano arithmetic? We could define an interpretation directly for WE-PA$^\omega$ (discussed ahead), or we could first use a negative translation. The latter are a kind of interpretation that we won't discuss in this work. They are a way to go from classical logic to intuitionistic logic, without extraction of terms. Perhaps the best known such interpretation is due to Gödel ([21]) and Gentzen (independently, [18, 19]). Negative translations can be coupled with functional interpretations of an intuitionistic theory to obtain functional interpretations of a classical theory. Hence, first doing a negative translation and then applying the *dialectica* interpretation, we get an interpretation of WE-PA$^\omega$ into WE-HA$^\omega$ with the possibility of term extraction. See Chapter 10 of [29] for more details.

Another way to obtain interpretations for classical theories is to do them directly. In Section 2.3, we present Shoenfield's interpretation, first published in Section 8.3 of his book [38], which interprets WE-PA$^\omega$ into WE-HA$^\omega$. As a side note, it was shown by Streicher and Kohlenbach in [41] (and independently by Avigad, unpublished) that Shoenfield's interpretation is precisely a specific negative translation due to Krivine [34] followed by the *dialectica* interpretation. So in this case, at least, defining an interpretation directly for Peano arithmetic is no different from composing a negative translation with an interpretation for Heyting arithmetic.

1.3 Interpreting nonstandard arithmetic

Interpreting nonstandard arithmetic by way of functional interpretations is a somewhat recent idea due to Benno van den Berg, Eyvind Briseid and Pavol Safarik [2]. In their paper, they describe two functional interpretations for intuitionistic and classical logic: the D_{st}-interpretation, and the S_{st}-interpretation, respectively.

In the theories E-HA$^{\omega *}_{\text{st}}$ and E-PA$^{\omega *}_{\text{st}}$ there is a unary predicate symbol "st", which we see as determining whether a given object is standard. This gives rise to two sets of quantifiers: the internal quantifiers $\forall$ and $\exists$, which quantify over all objects, and the external quantifiers $\forall^{\text{st}}$ and $\exists^{\text{st}}$, which quantify only over standard objects. The interpretations of Van den Berg et al. distinguish the treatment of internal and external quantifiers. The external quantifiers are interpreted as usual, requiring computational witnesses, while the internal quantifiers are exempt from this obligation.

There is another interesting particularity of the D_{st}-interpretation. It is herbrandised, in the sense that the witnesses for the external, existential quantifiers of the interpreting formulas need not be exact; only a finite set of possible witnesses for each statement is required. In this spirit, Van den Berg et al. impose that the external, existential quantifiers of the interpreting formulas quantify only over finite sets, which syntactically translates to quantifying over variables of star type. The star type indicates finite sequences, not sets, but the order of the elements is never relevant.

In accordance with the herbrandised quality of the D_{st}-interpretation, the interpreting formulas have a monotonicity property, *i.e.*, if a sequence s is a witness for an external existential statement, then a "bigger" sequence s' is also a witness for it. When s and s' are of star type, we define the meaning of "bigger" like so: s is contained in s' if and only if all the elements of s are also elements of s'.

In Chapter 3 of this work, we present two different interpretations for nonstandard arithmetic: the H_{st} and S_{st}-interpretations, in Sections 3.3 and 3.4, respectively. These interpretations are adaptations of the ones by Van den Berg et al. The key difference between the H_{st} and D_{st}-interpretations is in the external, existential statements of the interpreting formulas: here these variables don't have to be of star type, they only have to be of *end-star* type. Terms of this type represent functions with several inputs whose last possible output is a finite sequence.

It is possible to extend the preorder of sequences mentioned above to a preorder of the end-star types. The extension is pointwise, following a suggestion of Fernando Ferreira. For example, if f and f' are of type $\rho \to \sigma \to \tau^*$, we say that f is contained in f' if and only if for all x of type ρ and for all y of type σ, fxy is contained in $f'xy$.

The S_{st}-interpretation described here is due to Dinis and Ferreira [10], even though Van den Berg et al. also present an interpretation for E-PA$^{\omega *}_{\text{st}}$ called S_{st} in Section 7 of [2]. They are very similar (and in fact equivalent in a certain sense), but come from different philosophies: while Van den Berg et al. obtain their interpretation by coupling the D_{st}-interpretation with a negative translation and lose the monotonicity property in the process, Dinis and Ferreira define their S_{st}-interpretation from scratch and with monotonicity. Curiously, in the classical setting all external, existential quantifications are naturally of star type, so the notion of containment between end-star types is never necessary.

1.4 Other functional interpretations

There were a lot of functional interpretations proposed over the years. A comprehensive list can be found in the historical comments at the end of Chapter 8 of [29]. Here we mention only a couple of interpretations more.

In [9], Diller and Nahm proposed a modification of the *dialectica* interpretation which doesn't need the decidability of quantifier-free formulas to interpret the $A \to A \wedge A$ axiom. The trick is a kind of herbrandisation. See [8] for a discussion of this interpretation.

In [28], Kohlenbach presented the monotone functional interpretation, which is the *dialectica* interpretation followed by a weakening of the requirement for the witnessing terms. Instead of requiring precise witnesses, it requires only majorants for such witnesses. Since it allows for more side principles, it has been used very successfully as a tool for proof mining. A survey of examples of its use can be found in [30].

Ferreira and Oliva also have interpretations which only require bounds on the witnessing terms, called the bounded functional interpretations: [16] for the intuitionistic case, and [14] for the classical case. Contrary to the monotone functional interpretation, here the bounds are applied at each step of the interpretations, instead of only at the end. These interpretations were adapted to the nonstandard setting by Ferreira and Gaspar for the classical case, [15], and by Dinis and Gaspar for the intuitionistic case, [11].

The bounded functional interpretations share many similarities with the functional interpretations for nonstandard arithmetic discussed before, since both distinguish between two kinds of quantifiers (one kind which must produce explicit witnesses and the one kind interpreted uniformly), and both dispense with precise witnesses, requiring merely some kind of bound.

1.5 Weihrauch reducibility

In Chapter 4, we study a relation between multi-valued operations introduced by Klaus Weihrauch [45, 46], and later modified to its modern definition by Gherardi and Marcone [20]. It has been used in logic to find which theorems can be computationally (or continuously) reduced into which other (see, for example, [5, 6, 7, 20]). To deal with this, we look at theorems as multi-valued operations. In particular, theorems of the form:

$$\forall x \, \exists y \, A_0(x, y)$$

are good candidates to see as multi-valued operations, since to every x it's possible to assign at least a y such that $A_0(x, y)$.

With this notion it is possible to define the so-called Weihrauch degrees, which are classes of inter-reducible theorems, thus obtaining a classification of theorems. In this work we don't classify any theorem, but instead give a suggestion on how to, in certain circumstances, prove that a $\forall\exists$-theorem Weihrauch-reduces to another, by way of the *dialectica* interpretation.

1.6 Thesis outline

We start by introducing the two best known functional interpretations in Chapter 2: Gödel's (Section 2.2) and Shoenfield's (Section 2.3). These are interpretations for Heyting and Peano arithmetic in all finite types, respectively. So the first step is to introduce formal systems for these two theories. That is done in Section 2.1.1 for Heyting arithmetic, and in Section 2.1.2 for Peano arithmetic. Then we give the statements of the soundness, term extraction, and characterization theorems for both interpretations.

In Chapter 3, we move into the realm of nonstandard arithmetic. We start by adapting the systems introduced in Section 2.1: first by adding the star type, meant to describe finite sequences, in Section 3.1.1; then by adding the unary "st" predicate, meant to describe the standardness of a term, in Section 3.1.2.

In Section 3.2, we describe proof principles often used in the context of nonstandard arithmetic. These principles will be eliminated by the functional interpretations we present in the ensuing sections.

In Section 3.3, we introduce the H_{st}-interpretation. This is a functional interpretation for nonstandard Heyting arithmetic, based on the D_{st}-interpretation of Van den Berg et al. [2]. The key difference is in the allowed types of the witnesses for the external, existential statements of the interpreting formulas. We prove soundness, term extraction, and characterization theorems for this new interpretation.

In Section 3.4, we describe the S_{st}-interpretation of Dinis and Ferreira [10], a functional interpretation for nonstandard Peano arithmetic. We extend the soundness theorem of this interpretation to a theory of full arithmetic, since in [10] the theory considered only allows induction for quantifier-free formulas. We also prove a characterization theorem for this interpretation.

In Section 3.5, we put the H_{st}-interpretation to use. We start by showing, in Section 3.5.1, that the principle of countable saturation, CSAT^{ω}, does not add power to an intuitionistic theory. Then, in Section 3.5.2, we note that that, when CSAT^{ω} is added to the classical theory of nonstandard arithmetic we use, the ensemble interprets full second-order arithmetic.

In Section 4.1, we introduce the notion of Weihrauch reducibility between multi-valued operations. Finally, Section 4.2 is a short note describing how to use Gödel's interpretation of Section 2.2 to, in certain circumstances, prove that a $\forall\exists$-theorem Weihrauch-reduces to another one.

2

Functional interpretations for arithmetic

Contents

2.1 Systems for Heyting and Peano arithmetic in all finite types

We use several formalizations of arithmetic throughout this work, both in the intuitionistic and the classical settings. This section describes what we mean by Heyting and Peano arithmetic in all finite types. Along the way, we mention various useful results, mostly without proof. For an overview of these topics, see Chapter 3 of [29].

In order to present the typed systems of this section, we need to first define what we mean by type.

Definition 2.1.1 (Finite types)**.** The finite types are described inductively as:

- 0 is a finite type;
- If ρ, τ are finite types, then $(\rho \to \tau)$ is a finite type.

The parenthesis in a type associate to the right, and we omit them when possible, to simplify the notation.

The type 0 should be thought of as the natural numbers, and the type $\rho \to \tau$ as the functions from objects of type ρ to objects of type τ. As a side note, the type $0 \to 0$, also represented by the symbol 1, is the type of the functions from the natural numbers to the natural numbers. It is also the appropriate type for the real numbers, which can be represented by Cauchy sequences of rational numbers. See [29] for a description on how to formalize analysis.

Remark 2.1.2. Any type $\rho \neq 0$ can be uniquely written as $\rho = \rho_1 \to \rho_2 \to \cdots \to \rho_k \to 0$, for some natural number k and types $\rho_1, \ldots, \rho_k$.

2.1.1 Extensional and weakly extensional Heyting arithmetic, E-HA$^\omega$ and WE-HA$^\omega$

Extensional and weakly extensional Heyting arithmetic in all finite types, represented by E-HA$^\omega$ and WE-HA$^\omega$ respectively, share the same terms, formulas and most axioms. The distinction is made in the axioms of equality for types higher than 0, as we will see shortly.

Terms

- A numerable set of variables for each type: $x^\sigma, y^\sigma, z^\sigma, \ldots$;
- Logical constants or combinators:
  - $\Pi_{\rho,\tau}$ of type $\rho \to \tau \to \rho$, for all types ρ, τ;
  - $\Sigma_{\delta,\rho,\tau}$ of type $(\delta \to \rho \to \tau) \to (\delta \to \rho) \to \delta \to \tau$, for all types δ, ρ, τ;
- Arithmetical constants:
  - 0 of type 0 (zero);
  - S of type $0 \to 0$ (successor);

– $(\boldsymbol{R}_{\boldsymbol{\rho}}) = (R_1)_{\boldsymbol{\rho}}, \dots, (R_k)_{\boldsymbol{\rho}}$ where $\boldsymbol{\rho} = \rho_1, \dots, \rho_k$ and each R_i has type:

$$0 \to \rho_1 \to \cdots \to \rho_k \to (0 \to \rho_1 \to \cdots \to \rho_k \to \rho_1) \to \cdots \to (0 \to \rho_1 \to \cdots \to \rho_k \to \rho_k) \to \rho_i$$

for all natural numbers k, and types $\rho_1, \dots, \rho_k$ (simultaneous recursors);

- If $t^\tau, f^{\tau \to \sigma}$ are terms, then $(ft)^\sigma$ is a term.

The type of a term is often omitted to avoid heavy notation. But every time a term is mentioned, it should be assumed that it has a type consistent with the context in which it appears, obeying the application of terms rule.

We think of $(f^{\tau\to\sigma} t^\tau)$ as "f applied to t", as if f were a unary function and t an argument. However, if $\sigma = \rho \to \delta$, the same term f could appear in the expression $((f^{\tau\to\rho\to\delta} t^\tau) u^\rho)$, and now it looks like it should be a binary function. In reality, all of the options above are valid term constructions, as long as the types are correct. The parenthesis associate to the left, which means that ftu is the same as $((ft)u)$.

The expression $\boldsymbol{t}$ is a (possibly empty) tuple of terms $t_1, \dots, t_k$, where $|\boldsymbol{t}| := k$. In particular, we use $\boldsymbol{t}, \boldsymbol{s} := t_1, \dots, t_k, s_1, \dots, s_l$, $f\boldsymbol{t} := ft_1 \cdots t_k$ and $\boldsymbol{f}\boldsymbol{t} := f_1\boldsymbol{t}, \dots, f_n\boldsymbol{t}$.

The variables of a term t, denoted by var(t), are defined inductively in the following way: if x is a variable, var$(x) := \{x\}$; if c is a constant (logical or arithmetical), var$(c) := \emptyset$; if $t^\tau, f^{\tau\to\sigma}$ are terms, var$(ft) :=$ var$(f) \cup$ var(t). If a term t has no variables, *i.e.*, if var$(t) = \emptyset$, we say that t is closed.

Formulas

- $\bot$ is an atomic formula (*falsum*);
- If t^0, u^0 are terms, then $t =_0 u$ is an atomic formula (equality of type 0);
- If A and B are formulas and x^τ is a variable, then $(A \wedge B)$, $(A \vee B)$, $(A \to B)$, $(\forall x^\tau A)$ and $(\exists x^\tau A)$ are formulas.

We use $(\neg A)$ as shorthand for $(A \to \bot)$, $(t \neq_0 u)$ as shorthand for $\neg(t =_0 u)$ and $(A \leftrightarrow B)$ as shorthand for $((A \to B) \wedge (B \to A))$.

In order to facilitate the reading of formulas, it is possible to omit some sets of parenthesis. This is governed by the following priority of symbols, from higher (left) to lower (right):

- $\neg$, $\forall$, $\exists$
- $\wedge$, $\vee$
- $\to$, $\leftrightarrow$

Furthermore, $\to$ associates to the right, *i.e.*, $A \to B \to C$ is to be interpreted as $A \to (B \to C)$, which makes for even lighter notation.

Formulas without both $\forall$ and $\exists$ are said to be quantifier-free, and are sometimes marked as so with an underscore "0", *i.e.*, $A_0, B_0, \dots$ represent quantifier-free formulas.

If A_0 is a quantifier-free formula, $B \equiv \forall \boldsymbol{x}\, A_0$ and $C \equiv \exists \boldsymbol{x}\, A_0$ for some tuple of variables $\boldsymbol{x}$, we say that B is a purely universal formula, and that C is a purely existential formula.

The free variables of a formula A, denoted by fv(A), are defined inductively in the following way: fv$(\bot) := \emptyset$; if t^0, u^0 are terms, fv$(t =_0 u) :=$ var$(t) \cup$ var(u); if A, B are formulas, fv$(A \,\Box\, B) :=$ fv$(A) \cup$ fv(B), with $\Box \in \{\wedge, \vee, \to\}$; if A is a formula and x is a variable, fv$(\triangle\, x\, A) :=$ fv$(A) \setminus \{x\}$, with $\triangle \in \{\forall, \exists\}$. Variables which are not free but do appear in the formula are said to be bound.

If a formula has no free variables, we say that it is a closed formula, or a sentence. When we write $A(x)$, we mean that x might be a free variable of A (but never a bound variable of A), and we wish to bring it to attention.

Logical axioms and rules

Heyting arithmetic is based on intuitionistic logic. The axioms that we use for it are:

Contraction:	$A \to A \wedge A$	$A \vee A \to A$	
Weakening:	$A \wedge B \to A$	$A \to A \vee B$	
Symmetry:	$A \wedge B \to B \wedge A$	$A \vee B \to B \vee A$	
Ex falso quodlibet:	$\bot \to A$		
Quantifier:	$\forall\, x\, A \to A[t/x]$	$A[t/x] \to \exists\, x\, A$	where t is free for x in A

The notation $A[t/x]$ represents formula A where variable x is replaced by t in every place where x appears free, *i.e.*, in every place where it is not in the scope of a quantification over x. The substitution can only be made when it doesn't lead to previously free variables in t becoming bound in $A[t/x]$. If the substitution is possible, we say that t is free for x in A. It is conventional to assume that, in a context where we wish to replace x for t in formula A, but t is not free for x in A, we just rename the bound variables in A until the substitution causes no problems, obtaining A'. This is possible because the name of a bound variable is for all intents and purposes irrelevant, meaning that A and A' are equivalent. After implementing this convention, there is no longer a need to mention whether a term is free for a variable in a formula or not.

When we use the axiom $\forall\, x\, A \to A[t/x]$, we say that we are instantiating x by t, and often write $x := t$.

The logical rules are:

Modus ponens:

$$\frac{A \quad A \to B}{B}$$

Syllogism:

$$\frac{A \to B \quad B \to C}{A \to C}$$

Exportation:

$$\frac{A \wedge B \to C}{A \to B \to C}$$

Importation:

$$\frac{A \to B \to C}{A \wedge B \to C}$$

Expansion:

$$\frac{A \to B}{C \vee A \to C \vee B}$$

∀-introduction:

$$\frac{B \to A}{B \to \forall x\, A}, x \notin \mathsf{fv}(B)$$

∃-introduction:

$$\frac{A \to B}{\exists x\, A \to B}, x \notin \mathsf{fv}(B)$$

There are many other descriptions of intuitionistic logic, using other axioms and rules. This one, first alluded to by Gödel in [22], is particularly useful for proving theorems about it, which is our purpose. Still, it would be possible to use a natural deduction description to obtain the same goals, as is done in [25]. Natural deduction makes proving assertions inside the language much more straightforward. For an overview of a natural deduction calculus for intuitionistic logic, including descriptions of semantics for it, see, for example, Chapters 2 and 9 of [42].

Equality

Following Chapter 3 of [29], we take type 0 equality as primitive with the following axioms:

Reflexivity: $x =_0 x$

Symmetry: $x =_0 y \to y =_0 x$

Transitivity: $x =_0 y \wedge y =_0 z \to x =_0 z$

Equality at higher types is defined: let $\rho = \rho_1 \to \cdots \to \rho_k \to 0$ be a type, and t, u terms of type ρ. Then:

$$(t =_\rho u) :\equiv (\forall\, y_1^{\rho_1}, \ldots, y_k^{\rho_k}\; ty_1 \ldots y_k =_0 uy_1 \ldots y_k)$$

where $y_1, \ldots, y_k$ are not variables of either t or u.

Sometimes we omit the subscript of the equality symbol to lighten the notation. It can be inferred from the types of the terms being compared. Conversely, if we write $t =_\rho u$, then it is implicit that both t and u have type ρ.

Reflexivity, symmetry and transitivity of $=_\rho$ follow from the same axioms for type zero equality. But there's still something missing. Even though our system does not have functions of arity greater than zero, it does have typed constants, which take their place. So it would seem reasonable to ask for some kind of extensionality axioms, for example:

$$E_{\rho,\tau} : \quad \forall\, f^{\rho\to\tau}, x^\rho, y^\rho\, (x =_\rho y \to fx =_\tau fy)$$

for all pairs of types (ρ, τ).

When accepting the axioms of extensionality, we obtain extensional Heyting arithmetic, E-HA$^\omega$. In this context it is possible to prove:

Proposition 2.1.3 (Equal terms are interchangeable in E-HA$^\omega$)**.** For every formula A, and variables x^ρ, y^ρ, E-HA$^\omega$ proves:

$$x =_\rho y \wedge A(x) \to A(y)$$

Proof. Induction on the logical structure of A. For the base case it is necessary to prove $x =_\rho y \to t[x/z] =_\tau t[y/z]$ by induction on the structure of the terms. ☐

However, the soundness theorem for the *dialectica* interpretation (Theorem 2.2.6) does not hold in the context of E-HA$^\omega$, precisely because of the extensionality axioms (see [24]). In order to weaken them, we use a quantifier-free extensionality rule:

$$\text{QF-ER}: \quad \frac{A_0 \to t =_\rho u}{A_0 \to r[t/x] =_\tau r[u/x]}$$

where A_0 is a quantifier-free formula, x^ρ is a variable and t^ρ, u^ρ and r^τ are terms.

When accepting QF-ER in lieu of the $E_{\rho,\tau}$ axioms, we obtain weakly extensional Heyting arithmetic, WE-HA$^\omega$. In this context we can only prove a weaker version of Proposition 2.1.3, namely:

Proposition 2.1.4. For any quantifier-free formula A_0 and formula B, WE-HA$^\omega$ proves:

$$\frac{A_0 \to t =_\rho u}{A_0 \to (B[t/x^\rho] \leftrightarrow B[u/x^\rho])}$$

There is another way to weaken E-HA$^\omega$ to a theory suitable to perform the soundness of the *dialectica* interpretation, namely to take equality as intensional. This is actually what Gödel does when describing his interpretation in [22]. For a discussion of the several possible treatments of equality, see Section 2.5 of [1].

Arithmetical axioms and rules

Successor axioms:

$$Sx \neq_0 0 \qquad\qquad Sx =_0 Sy \to x =_0 y$$

Π and Σ axioms:

$\Pi_{\rho,\tau} xy =_\rho x$, for any x^ρ, y^τ

$\Sigma_{\delta,\rho,\tau} xyz =_\tau xz(yz)$, for any $x^{\delta\to\rho\to\tau}, y^{\delta\to\rho}, z^\delta$

Recursor axioms: let $\boldsymbol{\rho} = \rho_1, \ldots, \rho_k$ be any tuple of types. Let x^0, $\boldsymbol{y} = y_1, \ldots, y_k$ with each y_i of type ρ_i, and $\boldsymbol{z} = z_1, \ldots, z_k$ with each z_i of type $0 \to \rho_1 \to \cdots \to \rho_k \to \rho_i$ be variables. The axioms are:

$$\left.\begin{array}{l} (R_i)_{\boldsymbol{\rho}} 0 \boldsymbol{y}\boldsymbol{z} =_{\rho_i} y_i \\ (R_i)_{\boldsymbol{\rho}} (Sx) \boldsymbol{y}\boldsymbol{z} =_{\rho_i} z_i x (\boldsymbol{R}_{\boldsymbol{\rho}} x \boldsymbol{y}\boldsymbol{z}) \end{array}\right. \quad \text{for } i \in \{1, \ldots, k\}$$

We changed the order of the arguments of the z_i's, compared to what is done in [29]. This is to simplify the construction of some terms ahead.

Induction schema:

$$A(0) \wedge \forall x^0 \, (A(x) \to A(Sx)) \to \forall x \, A(x)$$

We could have equivalently introduced the induction rule instead of the induction schema:

$$\frac{A(0) \quad A(x) \to A(Sx)}{A(x)}$$

Useful definitions and results

Our first goal is to show that quantifier-free formulas satisfy the law of excluded middle when in the presence of arithmetic. To achieve it, some preliminary results are necessary.

Lemma 2.1.5. WE-HA$^\omega \vdash \forall x^0 \, (x =_0 0 \vee x \neq_0 0)$.

Idea of Proof. Induction on x^0. □

Lemma 2.1.6. WE-HA$^\omega$ proves the following:

1. $x =_0 y \leftrightarrow |x - y| =_0 0$
2. $x =_0 0 \wedge y =_0 0 \leftrightarrow x + y =_0 0$
3. $x =_0 0 \vee y =_0 0 \leftrightarrow x \cdot y =_0 0$
4. $(x =_0 0 \to y =_0 0) \leftrightarrow \overline{\mathsf{sign}}(x) \cdot y =_0 0$

where $|x - y|$, $x + y$ and $x \cdot y$ have the usual meaning and $\overline{\mathsf{sign}}(x)$ is 0 when $x \neq_0 0$ and $S0$ when $x =_0 0$.

Idea of Proof. It should be clear that the functions used here are primitive recursive and hence that it is possible to define suitable terms for them in WE-HA$^\omega$, making use of recursion.

The proof of the equivalences goes by double induction on x, y and uses Lemma 2.1.5. □

Proposition 2.1.7. Let $A_0(\boldsymbol{x})$ be a quantifier-free formula of WE-HA$^\omega$, with free variables among $\boldsymbol{x}$. Then there exists a closed term t_{A_0} such that:

$$\text{WE-HA}^\omega \vdash \forall \boldsymbol{x} \, (t_{A_0}\boldsymbol{x} =_0 0 \leftrightarrow A_0(\boldsymbol{x}))$$

Idea of Proof. Induction on the logical structure of A_0.

Notice that the atomic formulas of WE-HA$^\omega$ are either $\perp$ or of the form $t =_0 u$ for terms t and u. Clearly WE-HA$^\omega \vdash 0 =_0 S0 \leftrightarrow \perp$, so even $\perp$ can be seen as an equality between terms.

The result then follows from Lemma 2.1.6: item 1 takes care of the base of induction for atomic formulas other than $\perp$, and the other three items of the steps for each logical symbol: $\wedge$, $\vee$, and $\to$, respectively. □

Corollary 2.1.8 (Law of excluded middle for quantifier-free formulas)**.** If A_0 is a quantifier-free formula of the language of WE-HA$^\omega$, then:

$$\text{WE-HA}^\omega \vdash A_0 \vee \neg A_0$$

Proof. Direct from Lemma 2.1.5 and Proposition 2.1.7. □

Corollary 2.1.9 (Elimination of $\vee$)**.** For every quantifier-free formula $A_0(\boldsymbol{x})$ of the language of WE-HA$^\omega$, there exists an equivalent quantifier-free formula $B_0(\boldsymbol{x})$ without $\vee$.

Proof. Simply take $B_0(\boldsymbol{x}) :\equiv (t_{A_0}\boldsymbol{x} =_0 0)$, as given by Proposition 2.1.7. This is an atomic formula, and clearly doesn't have any $\vee$. □

The next results focus on providing us with a very convenient way of expressing terms, namely λ-abstraction.

Definition 2.1.10 (λ-abstraction)**.**

- $(\lambda\, x^\rho\,.\, x)^{\rho\to\rho} := \Sigma_{\rho,\sigma\to\rho,\rho}\Pi_{\rho,\sigma\to\rho}\Pi_{\rho,\sigma}$
- $(\lambda\, x^\rho\,.\, t^\sigma)^{\rho\to\sigma} := \Pi_{\sigma,\rho}t$, if $x \notin$ var(t)
- $(\lambda\, x^\rho\,.\, t^{\sigma\to\tau}u^\sigma)^{\rho\to\tau} := \Sigma_{\rho,\sigma,\sigma\to\tau}(\lambda\, x\,.\, t)(\lambda\, x\,.\, u)$, if $x \in$ var(tu)

Remark 2.1.11. var$(\lambda\, x\,.\, t) =$ var$(t) \setminus \{x\}$, as can be clearly seen by induction on the construction of the lambda terms.

One easily shows by induction on the construction of the lambda terms that:

Proposition 2.1.12 (Combinatorial completeness)**.** WE-HA$^\omega \vdash (\lambda\, x^\rho\,.\, t^\tau)s^\rho =_\tau t[s/x]$.

We often write $\lambda\, x, y\,.\, t$ as shorthand for $\lambda\, x\,.\,(\lambda\, y\,.\, t)$. Furthermore, the notation $\lambda\, \boldsymbol{x}\,.\, \boldsymbol{t}$ should be interpreted as $(\lambda\, x_1, \ldots, x_k\,.\, t_1), \ldots, (\lambda\, x_1, \ldots, x_k\,.\, t_l)$.

Corollary 2.1.13. For every term t^τ and variable x^ρ, there exists a term T of type $\rho \to \tau$ and variables var$(T) =$ var$(t) \setminus \{x\}$ such that:

$$\text{WE-HA}^\omega \vdash Ts^\rho =_\tau t[s/x]$$

Proof. Taking $T := \lambda\, x\,.\, t$, this is a direct consequence of Remark 2.1.11 and Proposition 2.1.12. □

We close with the descriptions of some useful terms, namely a zero of arbitrary type, and definition by cases.

Definition 2.1.14 ($\mathcal{O}^\rho$)**.** For each type $\rho = \rho_1 \to \cdots \to \rho_k \to 0$, we distinguish a "zero of type ρ":

$$\mathcal{O}^\rho = \lambda\, x_1^{\rho_1}, \ldots, x_k^{\rho_k}\,.\, 0$$

This zero will be useful when we need a term of a specific type, but the term in itself is not relevant.

Proposition 2.1.15 (Definition by cases)**.** For every type ρ, there exists a closed term C such that:

$$\text{WE-HA}^\omega \vdash \forall\, x^0, y^\rho, z^\rho\, [(x = 0 \to Cxyz = y) \wedge (x \neq 0 \to Cxyz = z)]$$

Proof. Let

$$C := \lambda\, x^0, y^\rho, z^\rho\,.\, R_\rho xy(\lambda\, r^0, q^\rho\,.\, z)$$

Notice that C is well defined, for R_ρ has type $0 \to \rho \to (0 \to \rho \to \rho) \to \rho$, x has type 0, y has type ρ and $\lambda\, r^0, q^\rho\,.\, z$ has type $0 \to \rho \to \rho$. Furthermore:

$$\begin{aligned} C0yz &=_\rho R_\rho 0y(\lambda\, r^0, q^\rho\,.\, z) \\ &=_\rho y \\ C(Sx)yz &=_\rho R_\rho(Sx)y(\lambda\, r^0, q^\rho\,.\, z) \\ &=_\rho (\lambda\, r^0, q^\rho\,.\, z)x(R_\rho xy(\lambda\, r^0, q^\rho\,.\, z)) \\ &=_\rho z \end{aligned}$$

It only remains to notice that, as $x \neq_0 0 \to x = S(\mathsf{pred}\, x)$ (where $\mathsf{pred}^{0\to 0}$ is the predecessor term), then $x \neq_0 0 \to Cxyz =_\rho z$. □

Proposition 2.1.15 can be proven even without the use of R_ρ: the base type recursor R_0 suffices. This is shown in Proposition 3.19 of [29].

Remark 2.1.16. It should be clear that all of the results outlined above for WE-HA$^\omega$ also hold for E-HA$^\omega$, since they share the same language and one can derive the quantifier-free rule of extensionality from the extensionality axioms.

2.1.2 Extensional and weakly extensional Peano arithmetic, E-PA$^\omega$ and WE-PA$^\omega$

Peano arithmetic is obtained from Heyting arithmetic by using classical logic instead of intuitionistic logic. One way to do this is to maintain everything and add a new logical axiom, the law of excluded middle for all formulas:

$$\text{LEM}: \quad A \vee \neg A$$

Hence WE-PA$^\omega$ is WE-HA$^\omega$ + LEM, and E-PA$^\omega$ is E-HA$^\omega$ + LEM.

However, these definitions now have superfluous axioms and rules. Furthermore, one can obtain some logical connectives from others. This is a nuisance when one wants to prove something about a system of Peano arithmetic by induction on the logical structure of a formula, or by induction on the proof of a statement.

A minimal treatment of classical logic is, for example, to take $\vee$, $\neg$, and $\forall$ as primitive, and $\bot, \wedge$, $\to$, and $\exists$ as defined. Following Section 2.6 of [38], the logical axioms and rules can be reduced to:

LEM:

$$A \vee \neg A$$

Instantiation:

$$\forall x\, A \to A[t/x]$$

Expansion:

$$\frac{A}{A \vee B}$$

Contraction:

$$\frac{A \vee A}{A}$$

Association:

$$\frac{A \vee (B \vee C)}{(A \vee B) \vee C}$$

Cut:

$$\frac{A \vee B \quad \neg A \vee C}{B \vee C}$$

$\forall$-introduction:

$$\frac{A \vee B}{\forall x\, A \vee B}, x \notin \mathsf{fv}(B)$$

Remark 2.1.17. All the results from Section 2.1.1 proved in E-HA$^\omega$ or WE-HA$^\omega$ also hold for their classical counterparts.

The most obvious model of these systems is the set-theoretical: the objects of type 0 are the natural numbers, and the objects of type $\rho \to \tau$ are all the set-theoretical functionals from objects of type ρ to objects of type τ. There are other models; for example, it is possible to only allow sequentially continuous functionals, or only majorizable functionals. For a discussion of several models, see Section 3.6 of [29].

2.2 Gödel's *dialectica* interpretation

In this section, we give an overview of Gödel's *dialectica* interpretation for intuitionistic arithmetical formulas. It was first published as [22] in the *dialectica* journal, hence the name of the interpretation. Here we follow Chapter 8 of [29].

Definition 2.2.1 (D-interpretation)**.** The D-interpretation (also known as Gödel's or *dialectica* interpretation) associates to each formula A in the language of WE-HA$^\omega$ a formula A^D of the form

$$A^D \equiv \exists \boldsymbol{x}\, \forall \boldsymbol{y}\, A_D(\boldsymbol{x}, \boldsymbol{y})$$

with the same free variables and in the same language as A. The (possibly empty) variable tuples $\boldsymbol{x}$ and $\boldsymbol{y}$ and their types are uniquely determined by the logical structure of A. It is important that these variables do not appear free in A. The formula $A_D(\boldsymbol{x}, \boldsymbol{y})$ is quantifier-free and without $\vee$. The definition proper is given below. The sub-formulas inside square brackets are the formulas corresponding to the A_D above.

- If A is an atomic formula, $A^D :\equiv A_D :\equiv A$

Given the interpretations $A^D \equiv \exists \boldsymbol{x}\, \forall \boldsymbol{y}\, A_D(\boldsymbol{x}, \boldsymbol{y})$ and $B^D \equiv \exists \boldsymbol{u}\, \forall \boldsymbol{v}\, B_D(\boldsymbol{u}, \boldsymbol{v})$:

- $(A \wedge B)^D :\equiv \exists \boldsymbol{x}, \boldsymbol{u}\, \forall \boldsymbol{y}, \boldsymbol{v}\, [A_D(\boldsymbol{x}, \boldsymbol{y}) \wedge B_D(\boldsymbol{u}, \boldsymbol{v})]$
- $(A \vee B)^D :\equiv \exists z^0, \boldsymbol{x}, \boldsymbol{u}\, \forall \boldsymbol{y}, \boldsymbol{v}\, [(z = 0 \to A_D(\boldsymbol{x}, \boldsymbol{y})) \wedge (z \neq 0 \to B_D(\boldsymbol{u}, \boldsymbol{v}))]$
- $(A \to B)^D :\equiv \exists \boldsymbol{U}, \boldsymbol{Y}\, \forall \boldsymbol{x}, \boldsymbol{v}\, [A_D(\boldsymbol{x}, \boldsymbol{Y}\boldsymbol{x}\boldsymbol{v}) \to B_D(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v})]$
- $(\exists\, z\, A(z))^D :\equiv \exists\, z, \boldsymbol{x}\, \forall \boldsymbol{y}\, [A_D(\boldsymbol{x}, \boldsymbol{y}, z)]$
- $(\forall\, z\, A(z))^D :\equiv \exists\, \boldsymbol{X}\, \forall\, z, \boldsymbol{y}\, [A_D(\boldsymbol{X} z, \boldsymbol{y}, z)]$

Remark 2.2.2 (Trivial D-interpretations)**.**

1. $(A^D)^D \equiv A^D$, and consequently $(A \,\Box\, B)^D \equiv (A^D \,\Box\, B^D)^D$, for $\Box \in \{\wedge, \vee, \to\}$ and $(\triangle\, x\, A)^D \equiv (\triangle\, x\, A^D)^D$, for $\triangle \in \{\forall, \exists\}$.
2. If A_0 is a quantifier-free formula without $\vee$, then $A_0^D \equiv A$.
3. If $A \equiv \triangle\, \boldsymbol{x}\, A_0$ with A_0 a quantifier-free formula without $\vee$ and $\triangle \in \{\forall, \exists\}$, then $A^D \equiv A$.

From now on we assume that all quantifier-free formulas are without $\vee$, possible by Corollary 2.1.9, which permits the use of Remark 2.2.2 without problems.

The *dialectica* interpretation is almost, but not quite, a way of formalizing the ideas of the Brouwer-Heyting-Kolmogorov interpretation, described, for example, in Section 2.1 of [42]. Actual formalizations are Kleene's realizability, introduced in [27], and Kreisel's modified realizability, introduced in [33].

What distinguishes the D-interpretation is the treatment of implication. A motivation for its definition can be found in Section 8.1 of [29]. In the case of the realizability interpretations, realizers of an implication don't worry about occurrences of $\forall$ in the premise. Since negation is in fact an implication, this means that the weak existential quantifier $\neg\forall\neg$ is not given a proper witness.

Gödel's interpretation is much more demanding. All occurrences of some kind of existential quantification, be it in the form of disjunction, universal quantification in a premise or actual strong existential quantification must be catalogued and witnessed. A more thorough discussion of these topics can be found in [31].

We make use of three principles, which are characteristic principles of the *dialectica* interpretation (for reasons which will be made clear ahead), and we need their definitions:

Definition 2.2.3 (AC^ω)**.** The schema of choice, AC^ω, is the union for all finite types ρ and τ of:

$$\mathsf{AC}^{\rho,\tau} : \forall\, x^\rho\ \exists\, y^\tau\, A(x,y) \to \exists\, Y^{\rho\to\tau}\ \forall\, x^\rho\, A(x, Yx)$$

where A is any formula of the language of $\mathsf{WE\text{-}HA}^\omega$.

Definition 2.2.4 (M^ω)**.** Markov's principle, M^ω, is the union for all tuples of finite types $\boldsymbol{\rho}$ of:

$$\mathsf{M}^{\boldsymbol{\rho}} : \neg\,\forall\, \boldsymbol{x}^{\boldsymbol{\rho}}\, A_0(\boldsymbol{x}) \to \exists\, \boldsymbol{x}^{\boldsymbol{\rho}}\, \neg A_0(\boldsymbol{x})$$

where A_0 is a quantifier-free formula of the language of $\mathsf{WE\text{-}HA}^\omega$.

Definition 2.2.5 ($\mathsf{IP}^\omega_\forall$)**.** The independence of premise schema for purely universal premises, $\mathsf{IP}^\omega_\forall$, is the union for all finite types ρ of:

$$\mathsf{IP}^\rho_\forall : (\forall\, \boldsymbol{x}\, A_0(\boldsymbol{x}) \to \exists\, y^\rho\, B(y)) \to \exists\, y^\rho\, (\forall\, \boldsymbol{x}\, A_0(\boldsymbol{x}) \to B(y))$$

where A_0 is a quantifier-free formula of the language of $\mathsf{WE\text{-}HA}^\omega$, and y is not free in A_0.

Theorem 2.2.6 (Soundness of the D-interpretation)**.** Let $A(\boldsymbol{a})$ be an arbitrary formula of the language of $\mathsf{WE\text{-}HA}^\omega$, containing only $\boldsymbol{a}$ free, such that $A(\boldsymbol{a})^D \equiv \exists\, \boldsymbol{x}\ \forall\, \boldsymbol{y}\, A_D(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a})$. Furthermore, let $\Delta_\forall$ be a collection of purely universal sentences in the same language. Suppose:

$$\mathsf{WE\text{-}HA}^\omega + \mathsf{AC}^\omega + \mathsf{IP}^\omega_\forall + \mathsf{M}^\omega + \Delta_\forall \vdash A(\boldsymbol{a})$$

Then there are closed terms $\boldsymbol{t}$, which can be extracted from a proof of $A(\boldsymbol{a})$, such that:

$$\mathsf{WE\text{-}HA}^\omega + \Delta_\forall \vdash \forall\, \boldsymbol{y}\, A_D(\boldsymbol{t}\boldsymbol{a}, \boldsymbol{y}, \boldsymbol{a})$$

Idea of Proof. Induction on the proof of A. For each axiom possibly used in such a proof, we must provide witnessing closed terms $\boldsymbol{t}$ for the existentially quantified variables of its D-interpretation. Similarly, we must provide witnessing terms for the conclusions of each rule possibly used in the proof of A, assuming by induction hypothesis that we already have access to such terms for the premises of the rule.

This means that the full proof is very long, as there are a considerable number of axioms and rules to take into account. It is however not terribly complicated. Most terms needed are simple projections and the proof of $\forall \boldsymbol{y}\, A_D(\boldsymbol{ta}, \boldsymbol{y}, \boldsymbol{a})$ usually ends with the very same axiom or rule which is under consideration.

A notable exception is the axiom $A \to A \wedge A$. The first step to obtain $(A \to A \wedge A)^D$ is to find A^D. Each of the three instances of A^D should have different names for the bound variables, so that there is no confusion. So let's say that the free variables of A are in $\boldsymbol{a}$, that $A^D \equiv \exists \boldsymbol{x}\, \forall \boldsymbol{y}\, A_D(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a})$, and use the pairs $(\boldsymbol{u}, \boldsymbol{v})$ and $(\boldsymbol{z}, \boldsymbol{w})$ for the other two instances. Then:

$$(A \to A \wedge A)^D \equiv \exists \boldsymbol{U}, \boldsymbol{Z}, \boldsymbol{Y}\, \forall \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{w}\, (A_D(\boldsymbol{x}, \boldsymbol{Yxvw}, \boldsymbol{a}) \to A_D(\boldsymbol{Ux}, \boldsymbol{v}, \boldsymbol{a}) \wedge A_D(\boldsymbol{Zx}, \boldsymbol{w}, \boldsymbol{a})) \quad (2.2.1)$$

Take $\boldsymbol{t_U} := \boldsymbol{t_Z} := \lambda\, \boldsymbol{a}, \boldsymbol{x}\, .\, \boldsymbol{x}$, and:

$$\boldsymbol{t_Y} := \lambda\, \boldsymbol{a}, \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{w}\, .\, \begin{cases} \boldsymbol{v} & \text{if } \neg A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \\ \boldsymbol{w} & \text{if } A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \end{cases}$$

Notice that we can define $\boldsymbol{t_Y}$ as shown, because, as $A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a})$ is a quantifier-free formula, by Proposition 2.1.7 we know that there exists a closed term t such that:

$$\text{WE-HA}^\omega \vdash t\boldsymbol{xva} = 0 \leftrightarrow A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a})$$

which means that checking whether $A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a})$ holds is the same as checking whether $t\boldsymbol{xva} = 0$. The terms $\boldsymbol{t_Y}$ can then be defined with the help of Proposition 2.1.15.

We now need to show that replacing in (2.2.1) the existentially quantified $\boldsymbol{U}$, $\boldsymbol{Z}$ and $\boldsymbol{Y}$ by their respective terms followed by $\boldsymbol{a}$ produces a provable formula, namely:

$$\begin{cases} \forall \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{r}\, (A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \to A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \wedge A_D(\boldsymbol{x}, \boldsymbol{r}, \boldsymbol{a})) & \text{if } \neg A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \\ \forall \boldsymbol{x}, \boldsymbol{v}, \boldsymbol{r}\, (A_D(\boldsymbol{x}, \boldsymbol{r}, \boldsymbol{a}) \to A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \wedge A_D(\boldsymbol{x}, \boldsymbol{r}, \boldsymbol{a})) & \text{if } A_D(\boldsymbol{x}, \boldsymbol{v}, \boldsymbol{a}) \end{cases}$$

But it is clear that in both cases the formulas are provable in WE-HA$^\omega$, so we are done.

The full proof can be found in Section 8.2 of [29]. □

The aforementioned principles (AC$^\omega$, M$^\omega$ and IP$^\omega_\forall$) are said to be characteristic principles of the *dialectica* interpretation because they are not needed to prove their own D-interpretations. In fact, the interpretations of these three principles are instances of $A \to A$.

Theorem 2.2.7 (Program extraction by the D-interpretation)**.** Let $\forall x\, \exists y\, A_0(x, y)$ be a sentence in the language of WE-HA$^\omega$, with A_0 quantifier-free. Furthermore, let $\Delta_\forall$ be a collection of purely universal formulas. Suppose:

$$\text{WE-HA}^\omega + \text{AC}^\omega + \text{M}^\omega + \text{IP}^\omega_\forall + \Delta_\forall \vdash \forall x\, \exists y\, A_0(x, y)$$

Then there is a closed term t such that:

$$\text{WE-HA}^\omega + \Delta_\forall \vdash \forall x\, A_0(x, tx)$$

Proof. Direct consequence of Theorem 2.2.6, taking into account that:

$$(\forall x\ \exists y\ A_0(x,y))^D \equiv \exists Y\ \forall x\ A_0(x, Yx)$$

□

When in the presence of these characteristic principles, we can prove the equivalence between a formula and its *dialectica* interpretation:

Theorem 2.2.8 (Characterization theorem for the D-interpretation)**.** For all formulas A in the language of WE-HA$^\omega$:

$$\text{WE-HA}^\omega + \text{AC}^\omega + \text{M}^\omega + \text{IP}^\omega_\forall \vdash A \leftrightarrow A^D$$

Proof. Induction on the logical structure of A. In each case, the implication $A^D \to A$ is straightforward. For $A \to A^D$, the only cases which are not intuitionistic truths are those of universal quantification and implication. For the first, the axiom of choice is needed. For the second, we make use of all three principles. □

2.3 Shoenfield's interpretation

Shoenfield's interpretation, or S-interpretation, was proposed by Shoenfield in Section 8.3 of his book [38]. It is an interpretation of WE-PA$^\omega$ in WE-HA$^\omega$, as opposed to Gödel's, which only interprets WE-HA$^\omega$ in itself.

Definition 2.3.1 (The S-interpretation)**.** The S-interpretation (also known as Shoenfield's interpretation) associates to each formula A in the language of WE-PA$^\omega$ a formula A^S of the form

$$A^S \equiv \forall \boldsymbol{x}\ \exists \boldsymbol{y}\ A_S(\boldsymbol{x}, \boldsymbol{y})$$

with the same free variables and in the same language as A. The (possibly empty) variable tuples $\boldsymbol{x}$ and $\boldsymbol{y}$ and their types are uniquely determined by the logical structure of A. It is important that these variables do not appear free in A. The formula $A_S(\boldsymbol{x}, \boldsymbol{y})$ is quantifier-free. The definition proper is given below. The sub-formulas inside square brackets are the formulas corresponding to the A_S above.

- If A is an atomic formula, $A^S :\equiv A_S :\equiv A$

Given the interpretations $A^S \equiv \forall \boldsymbol{x}\ \exists \boldsymbol{y}\ A_S(\boldsymbol{x}, \boldsymbol{y})$ and $B^S \equiv \forall \boldsymbol{u}\ \exists \boldsymbol{v}\ B_S(\boldsymbol{u}, \boldsymbol{v})$:

- $(\neg A)^S :\equiv \forall \boldsymbol{Y}\ \exists \boldsymbol{x}\ [\neg A_S(\boldsymbol{x}, \boldsymbol{Y}\boldsymbol{x})]$
- $(A \vee B)^S :\equiv \forall \boldsymbol{x}, \boldsymbol{u}\ \exists \boldsymbol{y}, \boldsymbol{v}\ [A_S(\boldsymbol{x}, \boldsymbol{y}) \vee B_S(\boldsymbol{u}, \boldsymbol{v})]$
- $(\forall z\, A(z))^S :\equiv \forall z, \boldsymbol{x}\ \exists \boldsymbol{y}\ [A_S(\boldsymbol{x}, \boldsymbol{y}, z)]$

Lemma 2.3.2 (S-interpretation of some formulas)**.**
Let A_0 be a quantifier-free formula. Then:

- $A_0^S \equiv A_0$
- $(\forall \boldsymbol{z}\, A_0)^S \equiv \forall \boldsymbol{z}\, A_0$

Taking $A^S \equiv \forall \boldsymbol{x}\, \exists \boldsymbol{y}\, A_S(\boldsymbol{x}, \boldsymbol{y})$ and $B^S \equiv \forall \boldsymbol{u}\, \exists \boldsymbol{v}\, B_S(\boldsymbol{u}, \boldsymbol{v})$, WE-PA$^\omega$ proves:

- $(A \to B)^S \leftrightarrow \forall \boldsymbol{Y}, \boldsymbol{u}\, \exists \boldsymbol{x}, \boldsymbol{v}\, [A_S(\boldsymbol{x}, \boldsymbol{Yx}) \to B_S(\boldsymbol{u}, \boldsymbol{v})]$
- $(A \wedge B)^S \leftrightarrow \forall \boldsymbol{X}, \boldsymbol{U}\, \exists \boldsymbol{Y}, \boldsymbol{V}\, [A_S(\boldsymbol{XYV}, \boldsymbol{Y}(\boldsymbol{XYV})) \wedge B_S(\boldsymbol{UYV}, \boldsymbol{V}(\boldsymbol{UYV}))]$
- $(\exists z\, A(z))^S \leftrightarrow \forall \boldsymbol{X}\, \exists z, \boldsymbol{Y}\, [A_S(\boldsymbol{X}z\boldsymbol{Y}, \boldsymbol{Y}(\boldsymbol{X}z\boldsymbol{Y}), z)]$

The S-interpretation of $\wedge$ looks unnecessarily complicated. After all it seems like

$$(A \wedge B)^S \equiv \forall \boldsymbol{x}, \boldsymbol{u}\, \exists \boldsymbol{y}, \boldsymbol{v}\, [A_S(\boldsymbol{x}, \boldsymbol{y}) \wedge B_S(\boldsymbol{u}, \boldsymbol{v})]$$

would have been possible and certainly more straightforward. This is in fact true, as shown in Section 4 of [41].

Definition 2.3.3 (QF-AC$^\omega$)**.** The quantifier-free schema of choice, QF-AC$^\omega$, is the restriction of AC$^\omega$ to quantifier-free formulas. So it is the union for all finite types $\boldsymbol{\rho}, \boldsymbol{\tau}$ of:

$$\text{QF-AC}^{\boldsymbol{\rho},\boldsymbol{\tau}} : \forall \boldsymbol{x}^{\boldsymbol{\rho}}\, \exists \boldsymbol{y}^{\boldsymbol{\tau}}\, A_0(\boldsymbol{x}, \boldsymbol{y}) \to \exists Y^{\boldsymbol{\rho} \to \boldsymbol{\tau}}\, \forall \boldsymbol{x}^{\boldsymbol{\rho}}\, A_0(\boldsymbol{x}, \boldsymbol{Yx})$$

where $\boldsymbol{\rho} \to \boldsymbol{\tau}$ is an abuse of notation which should be read as

$$\rho_1 \to \cdots \to \rho_k \to \tau_1,\ \ldots,\ \rho_1 \to \cdots \to \rho_k \to \tau_l$$

for $\boldsymbol{\rho} = \rho_1, \ldots, \rho_k$ and $\boldsymbol{\tau} = \tau_1, \ldots, \tau_l$, and A_0 is a quantifier-free formula of the language of WE-PA$^\omega$.

We need to define QF-AC$^\omega$ for tuples of variables, because, contrary to what happens with AC$^\omega$, it would not possible to apply QF-AC$^\omega$ to a formula several times in a row to make up for the lack of tuples in the definition (not all the subformulas would be quantifier-free).

Theorem 2.3.4 (Soundness of the S-interpretation)**.** Let $A(\boldsymbol{a})$ be an arbitrary formula of the language of WE-PA$^\omega$, containing only $\boldsymbol{a}$ free, such that $A(\boldsymbol{a})^S \equiv \forall \boldsymbol{x}\, \exists \boldsymbol{y}\, A_S(\boldsymbol{x}, \boldsymbol{y}, \boldsymbol{a})$. Furthermore, let $\Delta_\forall$ be a collection of purely universal sentences in the same language. Suppose:

$$\text{WE-PA}^\omega + \text{QF-AC}^\omega + \Delta_\forall \vdash A(\boldsymbol{a})$$

Then there are closed terms $\boldsymbol{t}$, which can be extracted from a proof of $A(\boldsymbol{a})$, such that:

$$\text{WE-HA}^\omega + \Delta_\forall \vdash \forall \boldsymbol{x}\, A_S(\boldsymbol{x}, \boldsymbol{tax}, \boldsymbol{a})$$

Shoenfield gives a proof of a similar result to this in Section 8.3 of his book [38]. Gaspar presents a full direct proof in Section 5.1 of [17], but he uses intensional instead of weakly-extensional equality. The treatment of the quantifier-free extensionality rule (QF-ER) is therefore missing. But it uses the same idea as in the soundness of the *dialectica* interpretation. Finally, by observing that Shoenfield's interpretation can be factored in Krivine's negative translation and Gödel's *dialectica*, as shown in [41], the soundness of the S-interpretation is also a consequence of the soundness of these two interpretations. More details can be found in Chapter 10 of [29].

Corollary 2.3.5 (Relative consistency)**.** WE-PA$^\omega$ is consistent relative to WE-HA$^\omega$.

Proof. Consequence of Theorem 2.3.4 when applied to $\bot$, since $\bot^S \equiv \bot$. □

Theorem 2.3.6 (Program extraction by the S-interpretation)**.** Let $\forall x\, \exists y\, A_0(x, y)$ be a sentence in the language of WE-PA$^\omega$, with A_0 quantifier-free. Furthermore, let $\Delta_\forall$ be a collection of purely universal formulas. Suppose:

$$\text{WE-PA}^\omega + \text{QF-AC}^\omega + \Delta_\forall \vdash \forall x\, \exists y\, A_0(x, y)$$

Then there is a closed term t such that:

$$\text{WE-HA}^\omega + \Delta_\forall \vdash \forall x\, A_0(x, tx)$$

Proof. Consequence of Theorem 2.3.4, taking into account that:

$$(\forall x\, \exists y\, A_0(x, y))^S \equiv \forall x\, \exists y\, \neg\neg A_0(x, y)$$

and that quantifier-free formulas are decidable in WE-HA$^\omega$ (Corollary 2.1.8). □

Theorem 2.3.7 (Characterization theorem for the S-interpretation)**.** For all formulas A of the language of WE-PA$^\omega$:

$$\text{WE-PA}^\omega + \text{QF-AC}^\omega \vdash A \leftrightarrow A^S$$

Proof. By induction on the logical structure of A. The only step that needs QF-AC$^\omega$ is the step for the negation. The full proof can be found in Section 5.1 of [17]. □

3

Functional interpretations for nonstandard arithmetic

Contents

It can be seen by compactness arguments that there are nonstandard models of arithmetic, *i.e.*, models not isomorphic to the natural numbers (Section 12.3 of [4]; see also Chapter 25 of the same book). It turns out that working in a nonstandard setting is a great way of formalizing reasoning about infinitesimals and infinitely big numbers.

Nonstandard analysis was popularized by Abraham Robinson [37]. After him, several people presented nonstandardness in different ways. Particularly relevant for us is the work of Nelson, who developed an axiomatization of nonstandard analysis named Internal Set Theory (IST) [35]. IST is Zermelo-Fraenkel set theory with choice (ZFC), plus a new predicate "st" for "being standard", and three new axioms: idealization, standardization and transfer. Nelson proved that IST is conservative over ZFC: any *internal* formula (not containing "st") provable in IST is already provable in ZFC.

It was noted by Benno van den Berg, Eyvind Briseid and Pavol Safarik [2] that the algorithm presented by Nelson in [36] to show the conservativity of IST over ZFC is remarkably similar to Shoenfield's functional interpretation [38]. From this observation, Van den Berg et al. developed a functional interpretation for nonstandard Heyting arithmetic, named the D_{st}-interpretation [2]. They worked in E-$\text{HA}^{\omega*}_{\text{st}}$, a formal system with higher types.

Our goal for Section 3.1 is to present the system E-$\text{HA}^{\omega*}_{\text{st}}$, as well as its classical variant, E-$\text{PA}^{\omega*}_{\text{st}}$, introduced in [2]. These systems are E-HA^{ω} and E-PA^{ω}, extended with a new type construction, the star type, and with a new predicate, the "st" predicate, in the spirit of Nelson's IST.

In Section 3.2, we describe some principles common in nonstandard settings, which will appear in the ensuing sections in prominent ways.

Section 3.3 is dedicated to the H_{st}-interpretation, a functional interpretation for nonstandard Heyting arithmetic (E-$\text{HA}^{\omega*}_{\text{st}}$) similar to the D_{st}-interpretation of Van den Berg et al. In Section 3.4 we present the S_{st}-interpretation of Dinis and Ferreira [10], which can be thought of as the H_{st}-interpretation's classical counterpart.

Finally, Section 3.5 is dedicated to the principle of countable saturation. We present the work of Van den Berg et al. [3], who showed that this principle is weak intuitionistically, but when added to a classical setting, the ensemble interprets full second-order arithmetic.

3.1 Systems for Heyting and Peano nonstandard arithmetic

3.1.1 Heyting and Peano arithmetic with the star type, E-$\text{HA}^{\omega*}$ and E-$\text{PA}^{\omega*}$

We want to extend the terms of E-HA^{ω} and E-PA^{ω} to include finite sequences. Actually, what we really want is to represent finite sets, but finite sequences are simpler and enough for our purposes. There is a new type to represent these finite sequences: the star type. We extend Definition 2.1.1 (Finite types) to include it:

- If ρ is a finite type, then ρ^* is a finite type.

The type ρ^* should be thought of as representing the finite sequences of elements of type ρ. Note that we have been talking about tuples of terms, and they should not be confused with sequences. A tuple $\boldsymbol{t}$ is shorthand for the terms $t_1^{\rho_1}, \ldots, t_k^{\rho_k}$, each with possibly different types. A sequence x^{ρ^*} is a specific term which semantically represents an ordered finite list of terms of type ρ: $x_0^\rho, \ldots, x_{n-1}^\rho$.

E-HA$^{\omega*}$ and E-PA$^{\omega*}$ share the same terms. The only difference between them is that E-HA$^{\omega*}$ is based on intuitionistic logic, while E-PA$^{\omega*}$ is based on classical logic.

Terms

In addiction to the terms of E-HA$^{\omega}$ and E-PA$^{\omega}$ described in Section 2.1.1, E-HA$^{\omega*}$ and E-PA$^{\omega*}$ also have terms to deal with the sequences:

- $\{\}_\rho$ of type ρ^*, for all types ρ (empty sequence);
- prep$_\rho$ of type $\rho \to \rho^* \to \rho^*$, for all types ρ (prepend an element to a sequence);
- $(\boldsymbol{L})_{\sigma,\boldsymbol{\rho}} = (L_1)_{\sigma,\boldsymbol{\rho}}, \ldots, (L_k)_{\sigma,\boldsymbol{\rho}}$ where $\boldsymbol{\rho} = \rho_1, \ldots, \rho_k$ and each L_i has type:

  $$\sigma^* \to \rho_1 \to \cdots \to \rho_k \to (\sigma \to \rho_1 \to \cdots \to \rho_k \to \rho_1) \to \cdots \to (\sigma \to \rho_1 \to \cdots \to \rho_k \to \rho_k) \to \rho_i$$

  for all types $\sigma, \boldsymbol{\rho}$ (simultaneous list recursors);
- $\Pi_{\rho,\tau}$, $\Sigma_{\delta,\rho,\tau}$ and $(\boldsymbol{R})_{\boldsymbol{\sigma}}$ where the types ρ, τ, δ and $\boldsymbol{\sigma}$ vary over all the types, including the new ones.

We slightly changed the notation from what is presented in [2], in order to be more consistent with the notation of [10], which we adopt.

Formulas

The same formulas as E-HA$^{\omega}$ (respectively E-PA$^{\omega}$), with terms in the language of E-HA$^{\omega*}$ (respectively E-PA$^{\omega*}$).

Equality

Following [2], we take equality as primitive in all types. The axioms are the following:

Reflexivity: $x =_\rho x$

Symmetry: $x =_\rho y \to y =_\rho x$

Transitivity: $x =_\rho y \wedge y =_\rho z \to x =_\rho z$

Left congruence: $x =_\rho y \to fx =_\tau fy$

Right congruence: $f =_{\rho\to\tau} g \leftrightarrow \forall x\,(fx =_\tau gx)$

Axioms and rules

- Logical axioms and rules of E-HA$^\omega$ (respectively E-PA$^\omega$) accounting for the new formulas available (since there are new terms);
- Axioms for Π, Σ and $\boldsymbol{R}$ accounting for the new types available;
- List recursor axioms: let $\boldsymbol{\rho} = \rho_1, \dots, \rho_k$ be any tuple of types. Let x^σ, s^{σ^*}, $\boldsymbol{y} = y_1, \dots, y_k$ with each y_i of type ρ_i, and $\boldsymbol{z} := z_1, \dots, z_k$ with each z_i of type $\sigma \to \rho_1 \to \dots \to \rho_k \to \rho_i$ be variables. The axioms are:

$$\left.\begin{array}{l}(L_i)_{\sigma,\boldsymbol{\rho}}\{\}_\sigma \boldsymbol{y}\boldsymbol{z} =_{\rho_i} y_i \\ (L_i)_{\sigma,\boldsymbol{\rho}}(\mathsf{prep}_\sigma\, xs)\boldsymbol{y}\boldsymbol{z} =_{\rho_i} z_i x(\boldsymbol{L}_{\sigma,\boldsymbol{\rho}} s\boldsymbol{y}\boldsymbol{z})\end{array}\right. \quad \text{for } i \in \{1, \dots, k\}$$

  We changed the order of the arguments of the z_i's, compared to what is done in [2]. This is to simplify the construction of some terms ahead.

- Sequence axiom: $\forall s^{\rho^*} (s = \{\}_\rho \vee \exists x^\rho, w^{\rho^*} s = \mathsf{prep}_\rho\, xw)$.

Useful Definitions and Results

Proposition 3.1.1 (Equal terms are interchangeable in E-HA$^{\omega *}$)**.** For every formula A of its language, E-HA$^{\omega *}$ proves:

$$x = y \wedge A(x) \to A(y)$$

Proof. As in Proposition 2.1.3, this is done by induction on the logical structure of A. Even though now equality is a primitive notion, the same proof holds. □

Proposition 3.1.2. E-HA$^{\omega *}$ is a definitional extension of E-HA$^\omega$. Likewise, E-PA$^{\omega *}$ is a definitional extension of E-PA$^\omega$.

Proof. This follows from the fact that we can already talk about sequences in E-HA$^\omega$ (respectively in E-PA$^\omega$) by making use of the terms available to code them, like is done in pages 58–60 of [29]. □

Our goal now is to define some useful notions to deal with terms of star type, such as what we mean by the length of a sequence, an element of a sequence, and so on. We also define a preorder over these terms.

Definition 3.1.3 ($|\cdot|_\rho, (\cdot)_{\cdot,\rho}$)**.** We define length and projection terms:

- Length: $|\cdot|_\rho$ of type $\rho^* \to 0$, defined as:

$$|\cdot|_\rho := \lambda s^{\rho^*} . L_{\rho,0} s 0 (\lambda x^\rho, n^0 . Sn)$$

  We write $|t|$ instead of $|\cdot|t$. Using the axioms of L, we obtain:

$$|\{\}| = 0$$
$$|\mathsf{prep}\, xs| = S|s|$$

- Projection: $(\cdot)_{\cdot,\rho}$ of type $\rho^* \to 0 \to \rho$, defined in appendix A. We write $(s)_i$ instead of $(\cdot).si$. We need the following properties:

$$(\{\})_i = \mathcal{O}$$
$$(\mathsf{prep}\, xs)_0 = x$$
$$(\mathsf{prep}\, xs)_{(Si)} = (s)_i$$

Definition 3.1.4 ($\in$)**.** If x^ρ, s^{ρ^*} are terms:

$$x \in_\rho s :\equiv \exists\, i^0 <_0 |s| \;\; x = (s)_i$$

where $<_0$ is definable already in WE-HA$^\omega$ such that it has the expected properties (see Section 3.5 of [29] for more details).

Lemma 3.1.5 (Induction schema for sequences)**.**

$$\text{E-HA}^{\omega *} \vdash \varphi(\{\}_\rho) \wedge \forall\, x^\rho, s^{\rho^*} \, (\varphi(s) \to \varphi(\mathsf{prep}\, xs)) \to \forall\, s^{\rho^*} \, \varphi(s)$$

Proof. One first shows using the sequence axiom that:

1. $\forall\, s^{\rho^*} \, (|s| = 0 \leftrightarrow s = \{\})$
2. $\forall\, n^0, s^{\rho^*} \, (|s| = Sn \leftrightarrow \exists\, x^\rho, w^{\rho^*} \, (s = \mathsf{prep}_\rho\, xw \wedge |w| = n))$

This effectively reduces sequence induction to ordinary induction, by considering the statement:

$$\forall\, n^0, s^{\rho^*} \, (|s| = n \to \varphi(s))$$

□

Definition 3.1.6 ($\{\cdot\}_\rho, \cup_\rho, \bigcup_{\rho,\tau}$)**.** We define terms that construct a singleton sequence, the union of two sequences and the indexed union of several sequences.

- Singleton: $\{\cdot\}_\rho$ of type $\rho \to \rho^*$, defined as:

$$\{\cdot\}_\rho := \lambda\, x^\rho \,.\, \mathsf{prep}_\rho\, x\{\}_\rho$$

  We write $\{x\}$ instead of $\{\cdot\}x$.

- Union: $\cup_\rho$ of type $\rho^* \to \rho^* \to \rho^*$, defined as:

$$\cup_\rho := \lambda\, t^{\rho^*}, u^{\rho^*} \,.\, L_{\rho,\rho^*} tu\, \mathsf{prep}_\rho$$

  We write $t \cup u$ instead of $\cup tu$. Using the axioms of L, we obtain:

$$\{\} \cup u = u$$
$$(\mathsf{prep}\, xt') \cup u = \mathsf{prep}\, x(t' \cup u)$$

- Indexed union: $\bigcup_{\rho,\tau}$ of type $(\rho \to \tau^*) \to \rho^* \to \tau^*$, defined as:

$$\bigcup_{\rho,\tau} := \lambda\, f^{\rho\to\tau^*}, t^{\rho^*} .\, L_{\rho,\tau^*} t\{\}(\lambda\, a, s\, .\, ((fa) \cup s))$$

  We write $\bigcup_{x\in t} fx$ instead of $\bigcup ft$. Using the axioms of L, we obtain:

$$\bigcup_{x\in\{\}} fx = \{\}$$
$$\bigcup_{x\in(\mathsf{prep}\, at')} fx = (fa) \cup \bigcup_{x\in t'} fa$$

Following the spirit of [10], we use the following three properties of the just defined terms to prove the soundness theorems of the functional interpretations ahead (Theorems 3.3.9 and 3.4.5). Notice, however, that since we have access to full arithmetic, we define the terms differently, making use of the list recursors.

Lemma 3.1.7. E-HA$^{\omega*}$ proves:

1. $x \in \{x\}$
2. $x \in t \vee x \in u \to x \in t \cup u$
3. $y \in t \wedge z \in fy \to z \in \bigcup_{x\in t} fx$

Proof.

1. We need to show:

$$\exists\, i < |\{x\}|\;\; x = (\{x\})_i$$

   Since $|\{x\}| = 1$ and $(\{x\})_0 = x$, taking $i := 0$ suffices.

2. We need to show:

$$(\exists\, i < |t|\;\; x = (t)_i) \vee (\exists\, j < |u|\;\; x = (u)_j) \to (\exists\, k < |t \cup u|\;\; x = (t \cup u)_k)$$

   First notice that $|t \cup u| = |t| + |u|$, which can be shown using induction over the sequence t. Then check that $\forall\, i < |t|\;\; (t)_i = (t \cup u)_i$, and that $\forall\, j < |u|\;\; (u)_j = (t \cup u)_{|t|+j}$, both by sequence induction over t. With this in mind, the result is trivial.

3. We need to show:

$$(\exists\, i < |t|\;\; y = (t)_i) \wedge (\exists\, j < |fy|\;\; z = (fy)_j) \to (\exists\, k < |\bigcup_{x\in t} fx|\;\; z = (\bigcup_{x\in t} fx)_k)$$

   First notice that $|\bigcup_{x\in t} fx| = \sum_{i<|t|} |f(t)_i|$, which can be seen by sequence induction on t. Given i, j as in the assumption, take $k := \sum_{l<i} |f(t)_l| + j$. This works, which can be seen by sequence induction on t.

□

It is possible to define a preorder on the star typed terms:

Definition 3.1.8 ($\subseteq_\tau$)**.** Given terms t, u of type τ^*:

$$t \subseteq_\tau u :\equiv \forall\, i < |t| \;\; \exists\, j < |u| \;\; (t)_i = (u)_j$$

Remark 3.1.9. The relation $\subseteq$ is both reflexive and transitive.

Notice that the order of the elements in the sequences t and u is not relevant to determine whether $t \subseteq u$. This is in accordance with the fact that we are thinking of terms of star type as stand ins for sets.

We can obtain as corollary from Lemma 3.1.7 a similar result for sequences:

Lemma 3.1.10. E-HA$^{\omega*}$ proves:

1. $x \in t \wedge t \subseteq t' \to x \in t'$
2. $x \subseteq t \vee x \subseteq u \to x \subseteq t \cup u$
3. $y \in t \wedge z \subseteq fy \to z \subseteq \bigcup_{x \in t} fx$

This concludes everything we need for terms of star type. However, there is another class of terms in which we are interested: the terms of *end-star* type.

Remark 3.1.11. Remark 2.1.2 no longer holds with the star types. Now it is also possible for a term to have type $\rho = \rho_1 \to \cdots \to \rho_k \to \tau^*$.

Definition 3.1.12 (End-star types)**.** If $\rho = \rho_1 \to \cdots \to \rho_k \to \tau^*$, for some natural number $k \geq 0$ and types $\rho_1, \ldots, \rho_k, \tau$, we say that ρ is an end-star type.

We are going to need a preorder between terms of end-star type. The following is an idea due to Fernando Ferreira, which works for our interpretation in a very elegant way.

Definition 3.1.13 ($\sqsubseteq_{\boldsymbol{\rho},\tau}$)**.** Given a tuple of types $\boldsymbol{\rho} = \rho_1, \ldots, \rho_k$ and terms f, g of end-star type $\rho_1 \to \cdots \to \rho_k \to \tau^*$:

$$f \sqsubseteq_{\boldsymbol{\rho},\tau} g :\equiv \forall\, \boldsymbol{x}^{\boldsymbol{\rho}}\, f\boldsymbol{x} \subseteq_\tau g\boldsymbol{x}$$

In the case where $\boldsymbol{\rho}$ is the empty tuple, $\sqsubseteq_{\boldsymbol{\rho},\tau}$ reduces to $\subseteq_\tau$.

Remark 3.1.14. The relation $\sqsubseteq$ is both reflexive and transitive.

Since we are going to work a lot with terms of end-star type, we define constants akin to $\cup$ and $\bigcup$ (Definition 3.1.6) in order to concatenate terms of this special kind of type.

Definition 3.1.15 ($\sqcup_{\boldsymbol{\rho},\tau}, \bigsqcup_{\sigma,\boldsymbol{\rho},\tau}$)**.** We define terms akin to $\cup$ and $\bigcup$, only for inputs of end-star type instead of star type.

- $\sqcup_{\boldsymbol{\rho},\tau}$ of type $(\rho_1 \to \cdots \to \rho_k \to \tau^*) \to (\rho_1 \to \cdots \to \rho_k \to \tau^*) \to (\rho_1 \to \cdots \to \rho_k \to \tau^*)$, defined as:

  $$\sqcup_{\boldsymbol{\rho},\tau} := \lambda\, t^{\rho_1 \to \cdots \to \rho_k \to \tau^*}, u^{\rho_1 \to \cdots \to \rho_k \to \tau^*} \,.\, (\lambda\, \boldsymbol{x}^{\boldsymbol{\rho}} \,.\, (t\boldsymbol{x}) \cup (u\boldsymbol{x}))$$

  We write $t \sqcup u$ instead of $\sqcup tu$.

- $\bigsqcup_{\sigma,\boldsymbol{\rho},\tau}$ of type $(\sigma \to \rho_1 \to \dots \to \rho_k \to \tau^*) \to \sigma^* \to (\rho_1 \to \dots \to \rho_k \to \tau^*)$, defined as:

$$\bigsqcup_{\sigma,\boldsymbol{\rho},\tau} := \lambda\, f^{\sigma \to \rho_1 \to \dots \to \rho_k \to \tau^*}, t^{\sigma^*} . (\lambda\, \boldsymbol{y}^{\boldsymbol{\rho}} . \bigcup_{x \in t} fx\boldsymbol{y})$$

  We write $\bigsqcup_{x \in t} fx$ instead of $\bigsqcup ft$.

Using Lemma 3.1.10, it is easy to check that the following properties of $\sqcup$ and $\bigsqcup$ hold. We will use them in the proof of the soundness theorem of the H_{st}-interpretation, Theorem 3.3.9.

Lemma 3.1.16. E-HA$^{\omega*}$ proves:

1. $x \sqsubseteq t \vee x \sqsubseteq u \to x \sqsubseteq t \sqcup u$
2. $y \in t \wedge z \sqsubseteq fy \to z \sqsubseteq \bigsqcup_{x \in t} fx$

Definition 3.1.17 (Monotonicity)**.** If x is of end-star type, we say that the formula $A(x)$ is monotone (or upward closed) in x if:

$$\forall\, y\, (x \sqsubseteq y \wedge A(x) \to A(y))$$

Definition 3.1.18 ($\mathcal{O}^\rho$)**.** Taking into account Remark 3.1.11, we need to define (closed) terms $\mathcal{O}^\rho$ where ρ is an end-star type:

$$\mathcal{O}^{\rho_1 \to \dots \to \rho_k \to \tau^*} := \lambda\, x_1^{\rho_1}, \dots, x_k^{\rho_k} . \{\mathcal{O}^\tau\}$$

Remark 3.1.19. All the results proved in E-HA$^{\omega*}$ during this section also hold in E-PA$^{\omega*}$.

3.1.2 Heyting and Peano nonstandard arithmetic, E-HA$^{\omega*}_{\mathsf{st}}$ and E-PA$^{\omega*}_{\mathsf{st}}$

Now it is time to introduce systems to deal with nonstandardness: E-HA$^{\omega*}_{\mathsf{st}}$ and E-PA$^{\omega*}_{\mathsf{st}}$. These will be extensions of E-HA$^{\omega*}$ and E-PA$^{\omega*}$, respectively. The terms remain the same, but we add some new formulas.

Formulas

- All formulas of the language of E-HA$^{\omega*}$ are formulas of E-HA$^{\omega*}_{\mathsf{st}}$ and all the formulas of the language of E-PA$^{\omega*}$ are formulas of E-PA$^{\omega*}_{\mathsf{st}}$;
- If t^ρ is a term, $\mathsf{st}_\rho\, t$ is an atomic formula;
- For E-HA$^{\omega*}_{\mathsf{st}}$: if x^ρ is a variable and A is a formula, $\forall^{\mathsf{st}}\, x^\rho\, A$ and $\exists^{\mathsf{st}}\, x^\rho\, A$ are formulas.
- For E-PA$^{\omega*}_{\mathsf{st}}$: $\forall^{\mathsf{st}}\, x\, A(x)$ is defined as shorthand for $\forall\, x\, (\neg\, \mathsf{st}\, x \vee A(x))$, and $\exists^{\mathsf{st}}\, x\, A(x)$ as shorthand for $\neg\, \forall\, x\, (\neg\, \mathsf{st}\, x \vee \neg A(x))$.

Definition 3.1.20 (Internal and external formulas)**.** An internal formula is a formula in the language of E-HA$^{\omega*}$ (or E-PA$^{\omega*}$), *i.e.*, without any st, $\forall^{\mathsf{st}}\, x$ or $\exists^{\mathsf{st}}\, x$. An external formula is a formula which is not internal.

We adopt an important convention, the same as Nelson in [36] and Van den Berg et al. in [2, 3]: from now on internal formulas are represented by lower-case Greek letters. Upper-case Latin letters represent formulas which can be external.

We say that formulas of the form $\forall^{\mathsf{st}}\, \boldsymbol{x}\, \varphi(\boldsymbol{x})$, with φ an internal formula, are $\forall^{\mathsf{st}}$-formulas.

Axioms and Rules

- All axioms and rules of E-HA$^{\omega*}$ or E-PA$^{\omega*}$ except induction, if considering E-HA$^{\omega*}_{\mathsf{st}}$ or E-PA$^{\omega*}_{\mathsf{st}}$, respectively
- External quantifiers (only for E-HA$^{\omega*}_{\mathsf{st}}$):

$$\forall^{\mathsf{st}} z\, A(z) \leftrightarrow \forall z\, (\mathsf{st}\, z \to A(z))$$
$$\exists^{\mathsf{st}} z\, A(z) \leftrightarrow \exists z\, (\mathsf{st}\, z \wedge A(z))$$

- Standardness axioms:

$$\mathsf{st}\, z \wedge z = w \to \mathsf{st}\, w$$
$$\mathsf{st}\, t,\ t \text{ a closed term}$$
$$\mathsf{st}\, f \wedge \mathsf{st}\, z \to \mathsf{st}(fz)$$

- Internal induction:

$$(\alpha(0) \wedge \forall n^0\, (\alpha(n) \to \alpha(Sn))) \to \forall n^0\, \alpha(n)$$

- External induction:

$$(A(0) \wedge \forall^{\mathsf{st}} n^0\, (A(n) \to A(Sn))) \to \forall^{\mathsf{st}} n^0\, A(n)$$

Useful definitions and results

Proposition 3.1.21 (Equal terms are interchangeable in E-HA$^{\omega*}_{\mathsf{st}}$)**.** For every formula A of its language, E-HA$^{\omega*}_{\mathsf{st}}$ proves:

$$x = y \wedge A(x) \to A(y)$$

Proof. By induction on the logical structure of A. The only steps needed to extend the proof of Proposition 3.1.1 to E-HA$^{\omega*}_{\mathsf{st}}$ are the cases where A is $\mathsf{st}\, t$, $\forall^{\mathsf{st}} x\, B(x)$, or $\exists^{\mathsf{st}} x\, B(x)$. In the case of $\mathsf{st}\, t$, we use the standardness axioms and the fact that $x = y \to r[x/z] = r[y/z]$ (proven by induction on r). The steps for the external quantifiers are corollaries of the steps for the internal quantifiers and $\mathsf{st}\, t$. ☐

Remark 3.1.22. The terms $|\cdot|$, $(\cdot)_\cdot$, $\{\cdot\}$, $\cup$, $\bigcup$, $\sqcup$ and $\bigsqcup$ are all standard, because they are all closed. This means that if they are applied to standard terms, the result is still standard. Both these statements are direct consequences of the standardness axioms.

Lemma 3.1.23. If n^0 is standard, all terms below n are also standard:

$$\text{E-HA}^{\omega*}_{\mathsf{st}} \vdash \mathsf{st}_0\, n \wedge m \leq n \to \mathsf{st}_0\, m$$

Proof. External induction on the formula $A(n^0) \equiv \forall m^0\, (m \leq n \to \mathsf{st}\, m)$. ☐

Lemma 3.1.24. Elements of a standard sequence are themselves standard:

$$\text{E-HA}^{\omega*}_{\mathsf{st}} \vdash \mathsf{st}\, s \wedge x \in s \to \mathsf{st}\, x$$

Proof. We recall that $x \in s \equiv \exists\, i < |s|\ \ x = (s)_i$. Assuming st s, it is clear by Remark 3.1.22 that $|s|$ is standard. Then by Lemma 3.1.23 we get st i. This is enough to show st $(s)_i$, which means, by the first standardness axiom, that st x. □

Remark 3.1.25. All the results proved in E-HA$^{\omega*}_{\mathsf{st}}$ during this section also hold in E-PA$^{\omega*}_{\mathsf{st}}$.

Definition 3.1.26 (Internalization)**.** For every formula A of the language of E-HA$^{\omega*}_{\mathsf{st}}$, we define its internalization, A^{int}, in the language of E-HA$^{\omega*}$, as the formula obtained from A by replacing every occurrence of st z by $z = z$, and every occurrence of $\forall^{\mathsf{st}}\, z$ and $\exists^{\mathsf{st}}\, z$ by $\forall\, z$ and $\exists\, z$, respectively.

Similarly, the internalization of a formula in the language of E-PA$^{\omega*}_{\mathsf{st}}$ is obtained by replacing all instances of st z by $z = z$.

Proposition 3.1.27. E-HA$^{\omega*}_{\mathsf{st}}$ is a conservative extension of E-HA$^{\omega*}$. Likewise, E-PA$^{\omega*}_{\mathsf{st}}$ is a conservative extension of E-PA$^{\omega*}$.

Proof. It suffices to show that the internalization of the external axioms of E-HA$^{\omega*}_{\mathsf{st}}$ (respectively E-PA$^{\omega*}_{\mathsf{st}}$) are provable in E-HA$^{\omega*}$ (respectively in E-PA$^{\omega*}$), which is straightforward. □

3.2 Nonstandard principles

We have defined a system, E-HA$^{\omega*}_{\mathsf{st}}$, which, despite being intended for nonstandard arithmetic, still admits the standard model (Proposition 3.1.27). In other words, the existence of nonstandard elements is never imposed. This changes when we start to consider nonstandard principles.

Definition 3.2.1 (OS$^{\omega}$)**.** The overspill principle, OS$^{\omega}$, is the union for all finite types ρ of:

$$\mathsf{OS}^{\rho} : \forall^{\mathsf{st}}\, x^{\rho}\, \varphi(x) \to \exists\, x^{\rho}\, (\neg\, \mathsf{st}\, x \wedge \varphi(x))$$

where φ is any internal formula in the language of E-HA$^{\omega*}_{\mathsf{st}}$.

Proposition 3.2.2. The overspill principle implies the existence of nonstandard objects of any type.

Proof. Just take $\varphi(x) \equiv (x = x)$ in OS$^{\omega}$, obtaining:

$$\exists\, x^{\rho}\, (\neg\, \mathsf{st}\, x)$$

□

Definition 3.2.3 (I$^{\omega}$)**.** The idealization principle, I$^{\omega}$, is the union for all finite types $\boldsymbol{\rho}$ and τ of:

$$\mathsf{I}^{\boldsymbol{\rho},\tau} : \forall^{\mathsf{st}}\, \boldsymbol{z'}^{\boldsymbol{\rho}^*}\ \exists\, w^{\tau}\ \forall\, \boldsymbol{z} \in \boldsymbol{z'}\ \ \varphi(\boldsymbol{z}, w) \to \exists\, w^{\tau}\ \forall^{\mathsf{st}}\, \boldsymbol{z}^{\boldsymbol{\rho}}\, \varphi(\boldsymbol{z}, w)$$

where if $\boldsymbol{\rho} = \rho_1, \ldots, \rho_k$, then $\boldsymbol{\rho}^* = \rho_1^*, \ldots, \rho_k^*$, the expression $\forall\, \boldsymbol{z} \in \boldsymbol{z'}$ is $\forall\, z_1 \in z_1' \cdots \forall\, z_k \in z_k'$, and φ is an internal formula in the language of E-HA$^{\omega*}$.

We wish to show that idealization implies overspill. For that, we need an auxiliary lemma:

Lemma 3.2.4. There is no standard sequence of type ρ^* which holds every standard term of type ρ:

$$\text{E-HA}^{\omega*}_{\mathsf{st}} \vdash \forall^{\mathsf{st}}\, z'^{\rho^*}\ \exists^{\mathsf{st}}\, w^{\rho}\ \forall\, z \in z'\ \ (z \neq w)$$

Proof. By induction on the type ρ.

$\boxed{\rho = 0}$

Let z'^{0^*} be an arbitrary standard sequence. Since z' is standard, it has a standard finite number of elements. Simply take w as the successor of the maximum element of z'. Then clearly w is standard and $w \notin z'$.

$\boxed{\rho = \sigma \to \tau}$

Let $z'^{(\sigma\to\tau)^*}$ be an arbitrary standard sequence. By induction hypothesis:

$$\forall^{\mathsf{st}}\, u'^{\tau^*}\ \exists^{\mathsf{st}}\, v^{\tau}\ \forall\, u \in u' \ \ (u \neq v) \tag{3.2.1}$$

Take in (3.2.1) $u' := \bigcup_{z\in z'}\{z\mathcal{O}^{\sigma}\}$, and let v be a standard term such that $\forall\, u \in u' \ \ (u \neq v)$. We now prove that $w := \lambda\, x^{\sigma} \,.\, v$ is such that $\forall\, z \in z' \ \ (z \neq w)$. Let $z \in z'$ be arbitrary and suppose $z =_{\sigma\to\tau} w$. Then in particular $z\mathcal{O}^{\sigma} =_{\tau} w\mathcal{O}^{\sigma}$. By definition of w, this means that $z\mathcal{O}^{\sigma} =_{\tau} v$. But this is a contradiction, since $z\mathcal{O}^{\sigma} \in u'$ and v was chosen precisely so that $v \notin u'$.

$\boxed{\rho = \sigma^*}$

Let $z'^{\sigma^{**}}$ be an arbitrary standard sequence. By induction hypothesis:

$$\forall^{\mathsf{st}}\, u'^{\sigma^*}\ \exists^{\mathsf{st}}\, v^{\sigma}\ \forall\, u \in u' \ \ (u \neq v) \tag{3.2.2}$$

Take in (3.2.2) $u' := \bigcup_{z\in z'} z$ and let v be a standard term such that $\forall\, u \in u' \ \ (u \neq v)$. We now prove that $w := \{v\}$ is such that $\forall\, z \in z' \ \ (z \neq w)$. Let $z \in z'$ be arbitrary and suppose $z =_{\sigma^*} w$. By definition of w, this means $z =_{\sigma^*} \{v\}$. But this is a contradiction, since on the one hand $z \subseteq u'$ by definition of u', and on the other hand $v \notin u'$ by hypothesis on v, which means that $\{v\} \nsubseteq u'$.

□

Proposition 3.2.5. $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} \vdash \mathsf{OS}^{\omega}$.

Proof. We prove that we get OS^{ω} for a general internal formula ψ from I^{ω} for the formula $\varphi(z, w) \equiv z \neq w \wedge \psi(w)$. That is, given:

$$\forall^{\mathsf{st}}\, z'^{\rho^*}\ \exists\, w^{\rho}\ \forall\, z \in z' \ \ (z \neq w \wedge \psi(w)) \to \exists\, w^{\rho}\ \forall^{\mathsf{st}}\, z^{\rho}\, (z \neq w \wedge \psi(w)) \tag{3.2.3}$$

we show:

$$\forall^{\mathsf{st}}\, x^{\rho}\, \psi(x) \to \exists\, x^{\rho}\, (\neg\, \mathsf{st}\, x \wedge \psi(x)) \tag{3.2.4}$$

Notice that the consequent of (3.2.3) implies the consequent of (3.2.4), so the only thing it remains to guarantee is:

$$\forall^{\mathsf{st}}\, x^{\rho}\, \psi(x) \to \forall^{\mathsf{st}}\, z'^{\rho^*}\ \exists\, w^{\rho}\ \forall\, z \in z' \ \ (z \neq w \wedge \psi(w)) \tag{3.2.5}$$

Assume the antecedent of (3.2.5) and let z' be standard but otherwise arbitrary. Take w standard such that $\forall\, z \in z' \ \ (z \neq w)$, as given by Lemma 3.2.4. Then it only remains to show $\psi(w)$. But since w is standard, this is a consequence of the assumption $\forall^{\mathsf{st}}\, x\, \psi(x)$.

□

The realization principle, so named in [2], is the contrapositive of the idealization principle. Hence in a classical context it is a consequence of I^ω.

Definition 3.2.6 (R^ω)**.** The realization principle, R^ω, is the union for all finite types ρ and τ of:

$$\mathsf{R}^{\rho,\tau} : \forall\, w^\rho\ \exists^{\mathsf{st}}\, z^\tau\ \varphi(z,w) \to \exists^{\mathsf{st}}\, z'^{\tau^*}\ \forall\, w^\rho\ \exists\, z \in z'\ \ \varphi(z,w)$$

where φ is any internal formula in the language of E-PA$^{\omega*}$.

Lemma 3.2.7 (R^ω for monotone formulas)**.** If z is of end-star type and $\varphi(z,w)$ is monotone on z (for any w), then R^ω reduces to:

$$\forall\, w^\rho\ \exists^{\mathsf{st}}\, z^\tau\ \varphi(z,w) \to \exists^{\mathsf{st}}\, z^\tau\ \forall\, w^\rho\ \varphi(z,w)$$

Proof. It is enough to show:

$$\exists^{\mathsf{st}}\, z'^{\tau^*}\ \forall\, w^\rho\ \exists\, z \in z'\ \ \varphi(z,w) \to \exists^{\mathsf{st}}\, z^\tau\ \forall\, w^\rho\ \varphi(z,w)$$

After fixing z' in the antecedent, take in the consequent $z := \bigsqcup_{v\in z'} v$. This is a standard term (remembering that z' is standard itself) and does the job. □

When we are in an intuitionistic context, the realization principle R^ω will not be enough. We need a version for external formulas:

Definition 3.2.8 (NCR^ω)**.** The non-classical realization principle, NCR^ω, is the union for all finite types ρ,τ of:

$$\mathsf{NCR}^{\rho,\tau} : \forall\, w^\rho\ \exists^{\mathsf{st}}\, z^\tau\ A(z,w) \to \exists^{\mathsf{st}}\, z'^{\tau^*}\ \forall\, w^\rho\ \exists\, z \in z'\ \ A(z,w)$$

where A is any formula in the language of E-HA$^{\omega*}_{\mathsf{st}}$.

Lemma 3.2.9 (NCR^ω for monotone formulas)**.** If z if of end-star type and $A(z,w)$ is monotone on z (for any w), then NCR^ω reduces to:

$$\forall\, w^\rho\ \exists^{\mathsf{st}}\, z^\tau\ A(z,w) \to \exists^{\mathsf{st}}\, z^\tau\ \forall\, w^\rho\ A(z,w)$$

Proof. Same as Lemma 3.2.7. □

The non-classical realization principle owes its name to the fact that, when added to E-HA$^{\omega*}_{\mathsf{st}}$, it entails the undecidability of the standardness predicate, as is shown by Proposition 3.5 of [2].

Definition 3.2.10 ($\mathsf{LLPO}^\omega_{\mathsf{st}}$)**.** The lesser limited principle of omniscience, $\mathsf{LLPO}^\omega_{\mathsf{st}}$, is the union for all types ρ and τ of:

$$\mathsf{LLPO}^{\rho;\tau}_{\mathsf{st}} : \forall^{\mathsf{st}}\, z'^{\rho^*}, w'^{\tau^*}\ (\forall\, z \in z'\ \ \varphi(z) \vee \forall\, w \in w'\ \ \psi(w)) \to \forall^{\mathsf{st}}\, z^\rho\, \varphi(z) \vee \forall^{\mathsf{st}}\, w^\tau\, \psi(w)$$

where φ,ψ are any internal formulas of the language of E-HA$^{\omega*}$.

The proof of the next lemma uses an idea of [11] (where $\mathsf{LLPO}^\omega_{\mathsf{st}}$ in their context is the bounded universal disjunction principle, BUD^ω).

Lemma 3.2.11. E-HA$^{\omega*}_{\mathsf{st}}$ $+ \mathsf{I}^{\omega} \vdash$ LLPO$^{\omega}_{\mathsf{st}}$.

Proof. Assume:

$$\forall^{\mathsf{st}} z'^{\rho^*}, w'^{\tau^*} \, (\forall z \in z' \;\; \varphi(z) \vee \forall w \in w' \;\; \psi(w)) \tag{3.2.6}$$

It is easy to check that (3.2.6) implies:

$$\forall^{\mathsf{st}} z', w' \; \exists n^0 \, ((n =_0 0 \to \forall z \in z' \;\; \varphi(z)) \wedge (n \neq_0 0 \to \forall w \in w' \;\; \psi(w)))$$

which, after some manipulation, gives us:

$$\forall^{\mathsf{st}} z', w' \; \exists n \; \forall z \in z' \;\; \forall w \in w' \;\; ((n =_0 0 \to \varphi(z)) \wedge (n \neq_0 0 \to \psi(w))) \tag{3.2.7}$$

Applying $\mathsf{I}^{\rho,\tau,0}$ to (3.2.7), we obtain:

$$\exists n \; \forall^{\mathsf{st}} z, w \, ((n =_0 0 \to \varphi(z)) \wedge (n \neq_0 0 \to \psi(w)))$$

Since, as seen in Lemma 2.1.5, equality of type 0 is decidable in WE-HA$^{\omega}$, and by extension in E-HA$^{\omega*}_{\mathsf{st}}$, we know that either $n =_0 0$ or $n \neq_0 0$. In the first case we are able to prove $\forall^{\mathsf{st}} z \, \varphi(z)$, and in the second we prove $\forall^{\mathsf{st}} w \, \psi(w)$. Hence we obtain the desired formula:

$$\forall^{\mathsf{st}} z \, \varphi(z) \vee \forall^{\mathsf{st}} w \, \psi(w)$$

□

These are all the nonstandard principles we will need. For more details about them, and for other nonstandard principles in the context of E-HA$^{\omega*}_{\mathsf{st}}$, see Section 3 of [2].

3.3 The H_{st}-interpretation

We now introduce an adaptation of Gödel's *dialectica* interpretation for nonstandard arithmetic: the herbrandised functional interpretation, or H_{st}-interpretation. It is heavily based in Van den Berg, Briseid, and Safarik's D_{st}-interpretation introduced in Section 5 of [2]. The main difference is in the treatment of the existentially quantified variables: while Van den Berg et al. require these variables to be of star type, here we just require them to be of end-star type.

There is another small difference: here we change the clause for the disjunction. We do this in order to be able to do the verification of the soundness theorem intuitionistically. In [2], the step for the expansion rule of the proof of the soundness theorem for the D_{st}-interpretation, though classically true, does not seem to hold in E-HA$^{\omega*}$. This was pointed out by Fernando Ferreira, after wondering about the dissimilarity between the disjunction clauses of the D_{st}-interpretation, and the bounded functional interpretation [16], when the other clauses are so strikingly similar.

Definition 3.3.1 (H_{st}-interpretation)**.** The H_{st}-interpretation associates to each formula A of E-HA$^{\omega*}_{\mathsf{st}}$ a formula $A^{H_{\mathsf{st}}}$ of the form:

$$A^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} \boldsymbol{x} \; \forall^{\mathsf{st}} \boldsymbol{y} \; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y})$$

with the same free variables and in the same language as A. The (possibly empty) variable tuples $\boldsymbol{x}$ and $\boldsymbol{y}$ and their types are uniquely determined by the logical structure of A. It is important that these variables do not appear free in A. Furthermore, $\alpha_{H_{\text{st}}}$ is an internal formula.

The definition proper is given below. The sub-formulas inside square brackets are the internal formulas corresponding to the $\alpha_{H_{\text{st}}}$ above.

- $\alpha^{H_{\text{st}}} :\equiv \alpha_{H_{\text{st}}} :\equiv \alpha$ for internal atomic formulas α
- $(\text{st}\, z)^{H_{\text{st}}} :\equiv \exists^{\text{st}} z' \, [z \in z']$

Given the interpretations $A^{H_{\text{st}}} \equiv \exists^{\text{st}} \boldsymbol{x} \, \forall^{\text{st}} \boldsymbol{y} \, \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y})$ and $B^{H_{\text{st}}} \equiv \exists^{\text{st}} \boldsymbol{u} \, \forall^{\text{st}} \boldsymbol{v} \, \beta_{H_{\text{st}}}(\boldsymbol{u}, \boldsymbol{v})$:

- $(A \wedge B)^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{x}, \boldsymbol{u} \, \forall^{\text{st}} \boldsymbol{y}, \boldsymbol{v} \, [\alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}) \wedge \beta_{H_{\text{st}}}(\boldsymbol{u}, \boldsymbol{v})]$
- $(A \vee B)^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{x}, \boldsymbol{u} \, \forall^{\text{st}} \boldsymbol{y'}, \boldsymbol{v'} \, [\forall \boldsymbol{y} \in \boldsymbol{y'} \;\; \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}) \vee \forall \boldsymbol{v} \in \boldsymbol{v'} \;\; \beta_{H_{\text{st}}}(\boldsymbol{u}, \boldsymbol{v})]$
- $(A \to B)^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{U}, \boldsymbol{Y} \, \forall^{\text{st}} \boldsymbol{x}, \boldsymbol{v} \, [\forall \boldsymbol{y} \in \boldsymbol{Y}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}) \to \beta_{H_{\text{st}}}(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v})]$
- $(\forall z \, A(z))^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{x} \, \forall^{\text{st}} \boldsymbol{y} \, [\forall z \, \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}, z)]$
- $(\exists z \, A(z))^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{x} \, \forall^{\text{st}} \boldsymbol{y'} \, [\exists z \; \forall \boldsymbol{y} \in \boldsymbol{y'} \;\; \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}, z)]$
- $(\forall^{\text{st}} z \, A(z))^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{X} \, \forall^{\text{st}} z, \boldsymbol{y} \, [\alpha_{H_{\text{st}}}(\boldsymbol{X}z, \boldsymbol{y}, z)]$
- $(\exists^{\text{st}} z \, A(z))^{H_{\text{st}}} :\equiv \exists^{\text{st}} \boldsymbol{x}, z' \, \forall^{\text{st}} \boldsymbol{y'} \, [\exists z \in z' \;\; \forall \boldsymbol{y} \in \boldsymbol{y'} \;\; \alpha_{H_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}, z)]$

It is worth mentioning that, since it is possible to write $\forall^{\text{st}}$ and $\exists^{\text{st}}$ from other logical constants, it would not have been necessary to define their interpretations. However, the definitions here presented are equivalent to the interpretations of the equivalent formulas.

Notice that, due to the clause for $\exists^{\text{st}}$, the H_{st}-interpretation is not idempotent.

Lemma 3.3.2 (Trivial H_{st}-interpretations)**.**
For an internal formula φ:

1. $\varphi^{H_{\text{st}}} \equiv \varphi$
2. $(\forall^{\text{st}} \boldsymbol{z} \, \varphi(\boldsymbol{z}))^{H_{\text{st}}} \equiv \forall^{\text{st}} \boldsymbol{z} \, \varphi(\boldsymbol{z})$

Proof. Just use the definition of the H_{st}-interpretation. □

Proposition 3.3.3. Let A be an arbitrary formula of the language of E-HA$^{\omega *}_{\text{st}}$ such that

$$A^{H_{\text{st}}} \equiv \exists^{\text{st}} \boldsymbol{x} \, \forall^{\text{st}} \boldsymbol{y} \, \alpha(\boldsymbol{x}, \boldsymbol{y})$$

Then $\boldsymbol{x}$ are of end-star type and, for fixed arbitrary $\boldsymbol{y}$, $\alpha(\boldsymbol{x}, \boldsymbol{y})$ is monotone on $\boldsymbol{x}$.

Proof. It is simple to check that $\boldsymbol{x}$ is of end-star type, by induction on the logical structure of A and the definition of the H_{st}-interpretation.

To see that α is monotone on $\boldsymbol{x}$, another simple induction is needed. We only show two steps, to illustrate.

$\boxed{\mathsf{st}\, z}$

$$(\mathsf{st}\, z)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} z' \, (z \in z')$$

Let $z'' \sqsupseteq z'$ be arbitrary. We need to show:

$$z \in z' \wedge z' \subseteq z'' \to z \in z''$$

which is clear from Lemma 3.1.10.

$\boxed{A \to B}$

$$(A \to B)^{H_{\mathsf{st}}} :\equiv \exists^{\mathsf{st}} \boldsymbol{U}, \boldsymbol{Y} \, \forall^{\mathsf{st}} \boldsymbol{x}, \boldsymbol{v} \, (\forall \boldsymbol{y} \in \boldsymbol{Y}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}) \to \beta_{H_{\mathsf{st}}}(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v}))$$

Let $\boldsymbol{U'} \sqsupseteq \boldsymbol{U}, \boldsymbol{Y'} \sqsupseteq \boldsymbol{Y}$ be arbitrary. Fix $\boldsymbol{x}$ and $\boldsymbol{v}$. We need to show:

$$(\forall \boldsymbol{y} \in \boldsymbol{Y}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}) \to \beta_{H_{\mathsf{st}}}(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v})) \to (\forall \boldsymbol{y} \in \boldsymbol{Y'}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}) \to \beta_{H_{\mathsf{st}}}(\boldsymbol{U'}\boldsymbol{x}, \boldsymbol{v}))$$

So assume both antecedents:

$$\forall \boldsymbol{y} \in \boldsymbol{Y}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}) \to \beta_{H_{\mathsf{st}}}(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v}) \tag{3.3.1}$$

$$\forall \boldsymbol{y} \in \boldsymbol{Y'}\boldsymbol{x}\boldsymbol{v} \;\; \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}) \tag{3.3.2}$$

From the definition of $\boldsymbol{Y} \sqsubseteq \boldsymbol{Y'}$, it is clear that $\boldsymbol{Y}\boldsymbol{x}\boldsymbol{v} \subseteq \boldsymbol{Y'}\boldsymbol{x}\boldsymbol{v}$. Then (3.3.2) is stronger than the antecedent of (3.3.1), which gives us $\beta_{H_{\mathsf{st}}}(\boldsymbol{U}\boldsymbol{x}, \boldsymbol{v})$. Since $\boldsymbol{U}\boldsymbol{x} \sqsubseteq \boldsymbol{U'}\boldsymbol{x}$, we can conclude $\beta_{H_{\mathsf{st}}}(\boldsymbol{U'}\boldsymbol{x}, \boldsymbol{v})$ by monotonicity of $\beta_{H_{\mathsf{st}}}$.

□

We proceed to describe some of the characteristic principles of the H_{st}-interpretation.

Definition 3.3.4 (HAC$^\omega$)**.** The herbrandised schema of choice, HAC$^\omega$, is the union for all types ρ, τ of:

$$\mathsf{HAC}^{\rho,\tau} : \forall^{\mathsf{st}} z^\rho \, \exists^{\mathsf{st}} w^\tau \, A(z, w) \to \exists^{\mathsf{st}} W'^{\rho \to \tau^*} \, \forall^{\mathsf{st}} z^\rho \, \exists w \in W'z \;\; A(z, w)$$

where A is any formula in the language of E-HA$^{\omega *}_{\mathsf{st}}$.

Lemma 3.3.5 (HAC$^\omega$ for monotone formulas)**.** If w is of end-star type and $A(z, w)$ is monotone on w (for any z), then HAC$^\omega$ reduces to:

$$\forall^{\mathsf{st}} z^\rho \, \exists^{\mathsf{st}} w^\tau \, A(z, w) \to \exists^{\mathsf{st}} W^{\rho \to \tau} \, \forall^{\mathsf{st}} z^\rho \, A(z, Wz)$$

Proof. It is enough to show:

$$\exists^{\mathsf{st}} W'^{\rho \to \tau^*} \, \forall^{\mathsf{st}} z^\rho \, \exists w \in W'z \;\; A(z, w) \to \exists^{\mathsf{st}} W^{\rho \to \tau} \, \forall^{\mathsf{st}} z^\rho \, A(z, Wz)$$

After fixing W' as in the antecedent, take in the consequent $W := \lambda z \, . \bigsqcup_{v \in W'z} v$. This is a standard term (remembering that W' is standard itself) and does the job. □

Definition 3.3.6 (HGMP$^\omega_{\mathsf{st}}$)**.** The herbrandised generalized Markov's principle, HGMP$^\omega_{\mathsf{st}}$, is the union for all types ρ of:

$$\mathsf{HGMP}^\rho_{\mathsf{st}} : (\forall^{\mathsf{st}} z^\rho \, \varphi(z) \to \psi) \to \exists^{\mathsf{st}} z'^{\rho^*} \, (\forall z \in z' \;\; \varphi(z) \to \psi)$$

where φ, ψ are internal formulas in the language of E-HA$^{\omega *}$.

Definition 3.3.7 ($\mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$)**.** The herbrandised independence of premise for $\forall^{\mathsf{st}}$-formulas schema, $\mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$, is the union for all types ρ, τ of:

$$\mathsf{HIP}^{\rho,\tau}_{\forall^{\mathsf{st}}} : (\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to \exists^{\mathsf{st}} w^{\tau}\, A(w)) \to \exists^{\mathsf{st}} w'^{\tau^*}\, (\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to \exists\, w \in w' \;\; A(w))$$

Lemma 3.3.8 ($\mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ for monotone formulas)**.** If w is of end-star type and $A(w)$ is monotone on w, then $\mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ reduces to:

$$(\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to \exists^{\mathsf{st}} w^{\tau}\, A(w)) \to \exists^{\mathsf{st}} w^{\tau}\, (\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to A(w))$$

Proof. It is enough to show:

$$\exists^{\mathsf{st}} w'^{\tau^*}\, (\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to \exists\, w \in w' \;\; A(w)) \to \exists^{\mathsf{st}} w^{\tau}\, (\forall^{\mathsf{st}} z^{\rho}\, \varphi(z) \to A(w))$$

After fixing w' as in the antecedent, take in the consequent $w := \bigsqcup_{v \in w'} v$. This is a standard term (remembering that w' is standard itself) and does the job. □

Theorem 3.3.9 (Soundness of the H_{st}-interpretation)**.** Let A be an arbitrary formula of the language of $\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}}$, possibly with free variables, such that $A^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} \boldsymbol{x}\, \forall^{\mathsf{st}} \boldsymbol{y}\, \alpha(\boldsymbol{x}, \boldsymbol{y})$. Furthermore, let $\Delta_{\forall^{\mathsf{st}}}$ be a collection of $\forall^{\mathsf{st}}$-formulas, and $\Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}}$ their internalizations. Suppose:

$$\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}} + \Delta_{\forall^{\mathsf{st}}} \vdash A$$

Then there are closed terms $\boldsymbol{t}$, which can be extracted from a proof of A, such that:

$$\mathsf{E\text{-}HA}^{\omega *} + \Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}} \vdash \forall\, \boldsymbol{y}\, \alpha(\boldsymbol{t}, \boldsymbol{y})$$

Proof. The proof follows by induction on the derivation of A. Throughout this proof we discard the tuple notation for simplicity. It should be evident from the context which variables are in fact tuples of variables. Furthermore, we fix:

$$A^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} x\; \forall^{\mathsf{st}} y\, \alpha(x, y)$$
$$B^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} p\; \forall^{\mathsf{st}} q\, \beta(p, q)$$
$$C^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} z\; \forall^{\mathsf{st}} w\, \gamma(z, w)$$

and also use these formulas with different names for the bound variables.

$\boxed{A \to A \wedge A}$

$$(A \to A \wedge A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} U, R, Y\; \forall^{\mathsf{st}} x, v, s\, (\forall\, y \in Yxvs \;\; \alpha(x, y) \to \alpha(Ux, v) \wedge \alpha(Rx, s))$$

Take $t_U := t_R := \lambda\, x\, .\, x$ and $t_Y := \lambda\, x, v, s\, .\, \{v\} \cup \{s\}$. We need to show:

$$\forall\, x, v, s\, (\forall\, y \in \{v, s\} \;\; \alpha(x, y) \to \alpha(x, v) \wedge \alpha(x, s))$$

which is clear.

Note that this step was much more straightforward here than in the case of the D-interpretation (Theorem 2.2.6). This is a clear example of how useful it is to allow finite sequences of possible witnesses, instead of requiring a single, specific witness.

$\boxed{A \vee A \to A}$

$$(A \vee A \to A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} R, Y', V' \; \forall^{\mathsf{st}} x, u, s \, (\\ \forall y' \in Y'xus \;\; \forall v' \in V'xus \;\; (\forall y \in y' \;\; \alpha(x,y) \vee \forall v \in v' \;\; \alpha(u,v)) \to \alpha(Rxu, s))$$

Take $t_{Y'} := t_{V'} := \lambda \, x, u, s \, . \, \{\{s\}\}$ and $t_R := \lambda \, x, u \, . \, x \sqcup u$. We need to show:

$$\forall \, x, u, s \, (\forall \, y' \in \{\{s\}\} \;\; \forall \, v' \in \{\{s\}\} \;\; (\forall \, y \in y' \;\; \alpha(x,y) \vee \forall \, v \in v' \;\; \alpha(u,v)) \to \alpha(x \sqcup u, s))$$

Let x, u, s be arbitrary, and assume the antecedent. With $y' := v' := \{s\}$ we get:

$$\forall \, y \in \{s\} \;\; \alpha(x,y) \vee \forall \, v \in \{s\} \;\; \alpha(u,v)$$

Assume the first case (the second one is analogous). Then clearly $\alpha(x,s)$, and so by Lemma 3.1.16.1 and the monotonicity of α, we get $\alpha(x \sqcup u, s)$, as desired.

$\boxed{A \wedge B \to A}$

$$(A \wedge B \to A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} R, Y, Q \; \forall^{\mathsf{st}} x, p, s \, (\forall \, y \in Yxps \;\; \forall \, q \in Qxps \;\; (\alpha(x,y) \wedge \beta(p,q)) \to \alpha(Rxp, s))$$

Take $t_Y := \lambda \, x, p, s \, . \, \{s\}$, $t_Q := \lambda \, x, p, s \, . \, \mathcal{O}$ and $t_R := \lambda \, x, p \, . \, x$. We need to show:

$$\forall \, x, p, s \, (\forall \, y \in \{s\} \;\; \forall \, q \in \mathcal{O} \;\; (\alpha(x,y) \wedge \beta(p,q)) \to \alpha(x,s))$$

Let x, p, s be arbitrary and assume the antecedent. Take $y := s$ and any $q \in \mathcal{O}$. We conclude $\alpha(x,s)$, as was our goal.

$\boxed{A \to A \vee B}$

$$(A \to A \vee B)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} U, P, Y \; \forall^{\mathsf{st}} x, v', q' \, (\\ \forall \, y \in Yxv'q' \;\; \alpha(x,y) \to \forall \, v \in v' \;\; \alpha(Ux, v) \vee \forall \, q \in q' \;\; \beta(Px, q))$$

Take $t_Y := \lambda \, x, v', q' \, . \, v'$, $t_U := \lambda \, x \, . \, x$ and $t_P := \lambda \, x \, . \, \mathcal{O}$. We need to show:

$$\forall \, x, v', q' \, (\forall \, y \in v' \;\; \alpha(x,y) \to \forall \, v \in v' \;\; \alpha(x,v) \vee \forall \, q \in q' \;\; \beta(\mathcal{O}, q))$$

Let x, v', q' be arbitrary. The resulting formula is an instance of the weakening axiom under consideration, after renaming of a bound variable.

$\boxed{A \wedge B \to B \wedge A}$

$$(A \wedge B \to B \wedge A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} R, U, Y, Q \; \forall^{\mathsf{st}} x, p, s, v \, (\\ \forall \, y \in Yxpsv \;\; \forall \, q \in Qxpsv \;\; (\alpha(x,y) \wedge \beta(p,q)) \to (\beta(Rxp, s) \wedge \alpha(Uxp, v)))$$

Take:

$$t_R := \lambda \, x, p \, . \, p \qquad t_Y := \lambda \, x, p, s, v \, . \, \{v\}$$
$$t_U := \lambda \, x, p \, . \, x \qquad t_Q := \lambda \, x, p, s, v \, . \, \{s\}$$

We need to show:

$$\forall\, x, p, s, v\, (\forall\, y \in \{v\} \;\; \forall\, q \in \{s\} \;\; (\alpha(x, y) \wedge \beta(p, q)) \to (\beta(p, s) \wedge \alpha(x, v))) \tag{3.3.3}$$

Let x, p, s, v be arbitrary, and assume the antecedent of (3.3.3). Let $y := v$ and $q := s$. Then we have $\alpha(x, v) \wedge \beta(p, s)$. The result follows from the symmetry axiom for $\wedge$.

$\boxed{A \vee B \to B \vee A}$

$$\begin{aligned}(A \vee B \to B \vee A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, R, U, Y', Q' \; \forall^{\mathsf{st}}\, x, p, s', v'\, (\\ \forall\, y' \in Y'xps'v' \;\; \forall\, q' \in Q'xps'v' \;\; (\forall\, y \in y' \;\; \alpha(x, y) \vee \forall\, q \in q' \;\; \beta(p, q)) \\ \to \\ (\forall\, s \in s' \;\; \beta(Rxp, s) \vee \forall\, v \in v' \;\; \alpha(Uxp, v)))\end{aligned}$$

Take:

$$t_R := \lambda\, x, p\, .\, p \qquad t_{Y'} := \lambda\, x, p, s', v'\, .\, \{v'\}$$
$$t_U := \lambda\, x, p\, .\, x \qquad t_{Q'} := \lambda\, x, p, s', v'\, .\, \{s'\}$$

We need to show:

$$\begin{aligned}\forall\, x, p, s', v'\, (\, \forall\, y' \in \{v'\} \;\; \forall\, q' \in \{s'\} \;\; (\forall\, y \in y' \;\; \alpha(x, y) \vee \forall\, q \in q' \;\; \beta(p, q)) \\ \to \\ (\forall\, s \in s' \;\; \beta(p, s) \vee \forall\, v \in v' \;\; \alpha(x, v)))\end{aligned}$$

Let x, p, s', v' be arbitrary, and assume the antecedent. Take $y' := v'$ and $q' := s'$ in said antecedent. Now we have as assumption:

$$\forall\, y \in v' \;\; \alpha(x, y) \vee \forall\, q \in s' \;\; \beta(p, q)$$

and need to show:

$$\forall\, s \in s' \;\; \beta(p, s) \vee \forall\, v \in v' \;\; \alpha(x, v)$$

The result follows by the symmetry axiom for $\vee$, after renaming of bound variables.

$\boxed{\bot \to A}$

$$(\bot \to A)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, x \; \forall^{\mathsf{st}}\, y\, (\bot \to \alpha(x, y))$$

Take $t_x := \mathcal{O}$. We need to show:

$$\forall\, y\, (\bot \to \alpha(\mathcal{O}, y))$$

Let y be arbitrary. The resulting formula is an instance of the *ex falso quodlibet* axiom.

$\boxed{\forall\, z\, A \to A[t/z]}$

For simplicity we write $A(z)$ and $A(t)$ instead of A and $A[t/z]$, respectively. This does not mean, however, that z is necessarily a free variable of A, even though the case where it isn't is of little interest.

$$(\forall\, z\, A(z) \to A(t))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, U, Y \; \forall^{\mathsf{st}}\, x, v\, (\forall\, y \in Yxv \;\; \forall\, z\, \alpha(x, y, z) \to \alpha(Ux, v, t))$$

Take $t_U := \lambda\, x \,.\, x$ and $t_Y := \lambda\, x, v \,.\, \{v\}$. We need to show:

$$\forall\, x, v\, (\forall\, y \in \{v\} \ \ \forall\, z\, \alpha(x, y, z) \to \alpha(x, v, t))$$

Let x, v be arbitrary and assume the antecedent. Then, with $y := v$, we get $\forall\, z\, \alpha(x, v, z)$. Instantiating $z := t$, we obtain $\alpha(x, v, t)$, as desired.

$A[t/z] \to \exists\, z\, A$

As before, we write $A(t)$ and $A(z)$ for $A[t/z]$ and A, respectively.

$$(A[t/z] \to \exists\, z\, A)^{H_{\text{st}}} \equiv \exists^{\text{st}}\, U, Y\ \forall^{\text{st}}\, x, v'\, (\forall\, y \in Yxv' \ \ \alpha(x, y, t) \to \exists\, z\ \forall\, v \in v' \ \ \alpha(Ux, v, z))$$

Take $t_U := \lambda\, x \,.\, x$ and $t_Y := \lambda\, x, v' \,.\, v'$. We need to show:

$$\forall\, x, v'\, (\forall\, y \in v' \ \ \alpha(x, y, t) \to \exists\, z\ \forall\, v \in v' \ \ \alpha(x, v, z))$$

Let x, v' be arbitrary. The resulting formula is an instance of the axiom under consideration, after renaming of a bound variable.

Modus ponens

$$\frac{A \quad A \to B}{B}$$

By induction hypothesis, we know closed terms t_1, t_2 and t_3 such that:

$$\forall\, y\, \alpha(t_1, y) \tag{3.3.4}$$

$$\forall\, x, q\, (\forall\, y \in t_2xq \ \ \alpha(x, y) \to \beta(t_3x, q)) \tag{3.3.5}$$

We want to find a closed term t_p such that $\forall\, q\, \beta(t_p, q)$. Take $t_p := t_3t_1$.

Let q be arbitrary. Instantiate $x := t_1$ and $q := q$ in (3.3.5). The goal is now to show the premise of (3.3.5), which allows us to conclude $\beta(t_3t_1, q)$, as desired. Let $y \in t_2t_1q$ be arbitrary. Then by (3.3.4) with $y := y$, we get $\alpha(t_1, y)$, which is precisely what was needed.

Syllogism

$$\frac{A \to B \quad B \to C}{A \to C}$$

By induction hypothesis, we know closed terms t_1, t_2, t_3 and t_4 such that:

$$\forall\, x, q\, (\forall\, y \in t_1xq \ \ \alpha(x, y) \to \beta(t_2x, q)) \tag{3.3.6}$$

$$\forall\, p, w\, (\forall\, q \in t_3pw \ \ \beta(p, q) \to \gamma(t_4p, w)) \tag{3.3.7}$$

We want to find closed terms t_Y and t_Z such that:

$$\forall\, x, w\, (\forall\, y \in t_Yxw \ \ \alpha(x, y) \to \gamma(t_Zx, w)) \tag{3.3.8}$$

Take $t_Y := \lambda\, x, w \,.\, \bigcup_{v \in t_3(t_2x)w} t_1xv$ and $t_Z := \lambda\, x \,.\, t_4(t_2x)$. Let x, w be arbitrary and assume the antecedent of (3.3.8):

$$\forall\, y \in \bigcup_{v \in t_3(t_2x)w} t_1xv \ \ \alpha(x, y) \tag{3.3.9}$$

From (3.3.7) with $p := t_2x$ and $w := w$ we get:

$$\forall\, q \in t_3(t_2x)w \;\; \beta(t_2x, q) \to \gamma(t_4(t_2x), w) \tag{3.3.10}$$

which means that the goal now is to show the antecedent of (3.3.10). Let $q_0 \in t_3(t_2x)w$ be arbitrary. From (3.3.6) with $x := x$ and $q := q_0$ we get:

$$\forall\, y \in t_1xq_0 \;\; \alpha(x, y) \to \beta(t_2x, q_0) \tag{3.3.11}$$

so it only remains to show the antecedent of (3.3.11). Let $y_0 \in t_1xq_0$ be arbitrary. Since $q_0 \in t_3(t_2x)w$, it follows from Lemma 3.1.7.3 that $y_0 \in \bigcup_{v \in t_3(t_2x)w} t_1xv$. Then by (3.3.9) we are done.

Exportation and Importation

$$\frac{A \wedge B \to C}{A \to B \to C} \qquad \frac{A \to B \to C}{A \wedge B \to C}$$

$$\begin{aligned}
&(A \wedge B \to C)^{H_{\mathsf{st}}} \leftrightarrow \\
&\quad \leftrightarrow \exists^{\mathsf{st}}\, Z, Y, Q\; \forall^{\mathsf{st}}\, x, p, w\, (\forall\, y \in Yxpw \;\; \alpha(x, y) \wedge \forall\, q \in Qxpw \;\; \beta(p, q) \to \gamma(Zxp, w)) \\
&(A \to B \to C)^{H_{\mathsf{st}}} \equiv \\
&\quad \equiv \exists^{\mathsf{st}}\, Z, Q, Y\; \forall^{\mathsf{st}}\, x, p, w\, (\forall\, y \in Yxpw \;\; \alpha(x, y) \to \forall\, q \in Qxpw \;\; \beta(p, q) \to \gamma(Zxp, w))
\end{aligned}$$

For both the exportation and importation rules, the terms we get as induction hypothesis from the premise of the rule do the job for the conclusion.

Expansion

$$\frac{A \to B}{C \vee A \to C \vee B}$$

By induction hypothesis, we have closed terms t_1 and t_2 such that:

$$\forall\, x, q\, (\forall\, y \in t_1xq \;\; \alpha(x, y) \to \beta(t_2x, q)) \tag{3.3.12}$$

We want to find closed terms t_F, t_P, $t_{W'}$ and $t_{Y'}$ such that:

$$\begin{aligned}
&\forall\, z, x, g', q'\, (\forall\, w' \in t_{W'}zxg'q' \;\; \forall\, y' \in t_{Y'}zxg'q' \;\; (\forall\, w \in w' \;\; \gamma(z, w) \vee \forall\, y \in y' \;\; \alpha(x, y)) \\
&\qquad \to \\
&\qquad \forall\, g \in g' \;\; \gamma(t_Fzx, g) \vee \forall\, q \in q' \;\; \beta(t_Pzx, q))
\end{aligned}$$

Take:

$$\begin{aligned}
t_F &:= \lambda\, z, x \,.\, z & t_{W'} &:= \lambda\, z, x, g', q' \,.\, \{g'\} \\
t_P &:= \lambda\, z, x \,.\, t_2x & t_{Y'} &:= \lambda\, z, x, g', q' \,.\, \{\textstyle\bigcup_{q \in q'} t_1xq\}
\end{aligned}$$

We need to show:

$$\begin{aligned}
&\forall\, z, x, g', q'\, (\forall\, w' \in \{g'\} \;\; \forall\, y' \in \{\textstyle\bigcup_{q \in q'} t_1xq\} \;\; (\forall\, w \in w' \;\; \gamma(z, w) \vee \forall\, y \in y' \;\; \alpha(x, y)) \\
&\qquad \to \\
&\qquad \forall\, g \in g' \;\; \gamma(z, g) \vee \forall\, q \in q' \;\; \beta(t_2x, q))
\end{aligned}$$

Let z, x, g', q' be arbitrary and assume the antecedent. Take $w' := g'$ and $y' := \bigcup_{q \in q'} t_1 xq$. There are two cases to consider: if $\forall\, w \in g' \;\; \gamma(z, w)$, we are done; if not, then we get:

$$\forall\, y \in \textstyle\bigcup_{q \in q'} t_1 xq \;\; \alpha(x, y) \tag{3.3.13}$$

Let $q \in q'$ be arbitrary. We need to show $\beta(t_2 x, q)$. This is a consequence of (3.3.12) with $x := x$ and $q := q$, as long as we can prove the antecedent $\forall\, y \in t_1 xq \;\; \alpha(x, y)$. But by Lemma 3.1.10.3, $t_1 xq \subseteq \bigcup_{q \in q'} t_1 xq$, so by (3.3.13) we are done.

∀-introduction

$$\frac{B \to A}{B \to \forall\, z\, A}, z \notin \mathsf{fv}(B)$$

By induction hypothesis, we have closed terms t_1 and t_2 such that:

$$\forall\, p, y\, (\forall\, q \in t_1 py \;\; \beta(p, q) \to \alpha(t_2 p, y)) \tag{3.3.14}$$

We want to find closed terms t_Q and t_X such that:

$$\forall\, p, y\, (\forall\, q \in t_Q py \;\; \beta(p, q) \to \forall\, z\, \alpha(t_X p, y))$$

Take $t_Q := t_1$ and $t_X := t_2$. By (3.3.14) and the ∀-introduction rule (z cannot be free in $\forall\, q \in t_1 py \;\; \beta(p, q)$ because t_1 is closed) it is easy to see that these work.

∃-introduction

$$\frac{A \to B}{\exists\, z\, A \to B}, z \notin \mathsf{fv}(B)$$

By induction hypothesis there are closed terms t_1 and t_2 such that:

$$\forall\, x, q\, (\forall\, y \in t_1 xq \;\; \alpha(x, y, z) \to \beta(t_2 x, q)) \tag{3.3.15}$$

We want to find closed terms $t_{Y'}$ and t_P such that:

$$\forall\, x, q\, (\forall\, y' \in t_{Y'} xq \;\; \exists\, z\; \forall\, y \in y' \;\; \alpha(x, y, z) \to \beta(t_P x, q))$$

Take $t_{Y'} := \lambda\, x, q\, .\, \{t_1 xq\}$ and $t_P := t_2$. We need to show:

$$\forall\, x, q\, (\forall\, y' \in \{t_1 xq\} \;\; \exists\, z\; \forall\, y \in y' \;\; \alpha(x, y, z) \to \beta(t_2 x, q)) \tag{3.3.16}$$

Notice that to show (3.3.16), it is enough to show:

$$\forall\, x, q\; \exists\, y' \in \{t_1 xq\} \;\; (\exists\, z\; \forall\, y \in y' \;\; \alpha(x, y, z) \to \beta(t_2 x, q)) \tag{3.3.17}$$

Taking x, q as arbitrary and $y' := t_1 xq$ in (3.3.17), it remains to show:

$$\exists\, z\; \forall\, y \in t_1 xq \;\; \alpha(x, y, z) \to \beta(t_2 x, q)$$

which is a consequence of (3.3.15) with $x := x$ and $q := q$ by the ∃-introduction rule.

External induction

Instead of interpreting the external induction axiom, we interpret the equivalent rule, which is simpler:

$$\frac{A(0) \quad \forall^{\mathsf{st}} n^0 \, (A(n) \to A(Sn))}{\forall^{\mathsf{st}} n^0 \, A(n)}$$

By induction hypothesis, there are closed terms t_1, t_2 and t_3 such that:

$$\forall y \, \alpha(t_1, y, 0) \tag{3.3.18}$$

$$\forall n^0, x, v \, (\forall y \in t_2 nxv \;\; \alpha(x, y, n) \to \alpha(t_3 nx, v, Sn)) \tag{3.3.19}$$

We want a closed term t_4 such that:

$$\forall n^0, w \, \alpha(t_4 n, w, n) \tag{3.3.20}$$

Take $t_4 := \lambda n \, . \, Rnt_1 t_3$. Then, by the defining equations for R, we get:

$$t_4 0 = t_1$$

$$t_4(Sn) = t_3 n(t_4 n)$$

We prove (3.3.20) by induction on n. The base case with $n = 0$ is a direct consequence of (3.3.18). For the step, assume:

$$\forall y \, \alpha(t_4 n, y, n) \tag{3.3.21}$$

We need to show:

$$\forall w \, \alpha(t_3 n(t_4 n), w, Sn)$$

Let w be arbitrary. From (3.3.19) with $n := n$, $x := t_4 n$ and $v := w$, we obtain:

$$\forall y \in t_2 n(t_4 n)w \;\; \alpha(t_4 n, y, n) \to \alpha(t_3 n(t_4 n), w, Sn) \tag{3.3.22}$$

Observing that (3.3.21) is stronger than the antecedent of (3.3.22), we can conclude the consequent of (3.3.22), which is what we wanted to show in the first place.

$\forall^{\mathsf{st}} z \, A(z) \to \forall w \, (\mathsf{st} \, w \to A(w))$

$$(\forall^{\mathsf{st}} z \, A(z) \to \forall w \, (\mathsf{st} \, w \to A(w)))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} U, Z, Y \;\, \forall^{\mathsf{st}} X, w', v \, ($$
$$\forall z \in ZXw'v \;\; \forall y \in YXw'v \;\; \alpha(Xz, y, z) \to \forall w \in w' \;\; \alpha(UXw', v, w))$$

Take:

$$t_U := \lambda X, w' \, . \, \textstyle\bigsqcup_{z \in w'} Xz$$

$$t_Z := \lambda X, w', v \, . \, w'$$

$$t_Y := \lambda X, w', v \, . \, \{v\}$$

We need to show:

$$\forall X, w', v \, (\forall z \in w' \;\; \forall y \in \{v\} \;\; \alpha(Xz, y, z) \to \forall w \in w' \;\; \alpha(\textstyle\bigsqcup_{z \in w'} Xz, v, w))$$

Let X, w', v be arbitrary, and assume the antecedent:

$$\forall\, z \in w' \;\; \forall\, y \in \{v\} \;\; \alpha(Xz, y, z) \tag{3.3.23}$$

We need to show:

$$\forall\, w \in w' \;\; \alpha(\textstyle\bigsqcup_{z \in w'} Xz, v, w)$$

Let $w \in w'$ be arbitrary. From (3.3.23) with $z := w$ and $y := v$, we get $\alpha(Xw, v, w)$. Since $w \in w'$, we know $Xw \sqsubseteq \bigsqcup_{z \in w'} Xz$ by Lemma 3.1.16.2. Then by monotonicity of α, we are done.

$\boxed{\forall\, w\, (\mathsf{st}\, w \to A(w)) \to \forall^{\mathsf{st}}\, z\, A(z)}$

$$\begin{aligned}(\forall\, w\, (\mathsf{st}\, w \to A(w)) \to \forall^{\mathsf{st}}\, z\, A(z))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, X, W', V \;\forall^{\mathsf{st}}\, U, z, y\, (\\ \forall\, w' \in W'Uzy \;\; \forall\, v \in VUzy \;\; \forall\, w \in w' \;\; \alpha(Uw', v, w) \to \alpha(XUz, y, z))\end{aligned}$$

Take:

$$\begin{aligned} t_X &:= \lambda\, U, z\, .\, U\{z\} \\ t_{W'} &:= \lambda\, U, z, y\, .\, \{\{z\}\} \\ t_V &:= \lambda\, U, z, y\, .\, \{y\} \end{aligned}$$

We want to show:

$$\forall\, U, z, y\, (\forall\, w' \in \{\{z\}\} \;\; \forall\, v \in \{y\} \;\; \forall\, w \in w' \;\; \alpha(Uw', v, w) \to \alpha(U\{z\}, y, z))$$

Let U, z, y be arbitrary, and assume the antecedent. From the antecedent, take $w' := \{z\}$ and $v := y$. So we get $\forall\, w \in \{z\} \;\; \alpha(U\{z\}, y, w)$ as assumption. Letting $w := z$, we obtain the desired result.

$\boxed{\exists^{\mathsf{st}}\, z\, A(z) \to \exists\, w\, (\mathsf{st}\, w \wedge A(w))}$

$$\begin{aligned}(\exists^{\mathsf{st}}\, z\, A(z) \to \exists\, w\, (\mathsf{st}\, w \wedge A(w)))^{H_{\mathsf{st}}} \leftrightarrow \exists^{\mathsf{st}}\, W', U, Y' \;\forall^{\mathsf{st}}\, x, z', v'\, (\\ \forall\, y' \in Y'xz'v' \;\; \exists\, z \in z' \;\; \forall\, y \in y' \;\; \alpha(x, y, z) \to \exists\, w \in W'xz' \;\; \forall\, v \in v' \;\; \alpha(Uxz', v, w))\end{aligned}$$

Take $t_{W'} := \lambda\, x, z'\, .\, z'$, $t_U := \lambda\, x, z'\, .\, x$ and $t_{Y'} := \lambda\, x, z', v'\, .\, \{v'\}$. We want to show:

$$\forall\, x, z', v'\, (\forall\, y' \in \{v'\} \;\; \exists\, z \in z' \;\; \forall\, y \in y' \;\; \alpha(x, y, z) \to \exists\, w \in z' \;\; \forall\, v \in v' \;\; \alpha(x, v, w))$$

Let x, z', v' be arbitrary, and assume the antecedent. Take $y' := v'$. Then we have as assumption:

$$\exists\, z \in z' \;\; \forall\, y \in v' \;\; \alpha(x, y, z)$$

and need to show:

$$\exists\, w \in z' \;\; \forall\, v \in v' \;\; \alpha(x, v, w)$$

which are the same formulas, after renaming of bound variables.

$\boxed{\exists\, w\, (\mathsf{st}\, w \wedge A(w)) \to \exists^{\mathsf{st}}\, z\, A(z)}$

$$(\exists\, w\, (\mathsf{st}\, w \wedge A(w)) \to \exists^{\mathsf{st}}\, z\, A(z))^{H_{\mathsf{st}}} \leftrightarrow \exists^{\mathsf{st}}\, X, Z', V'\ \forall^{\mathsf{st}}\, w', u, y'\, (\\ \forall\, v' \in V'w'uy'\ \ \exists\, w \in w'\ \ \forall\, v \in v'\ \ \alpha(u, v, w) \to \exists\, z \in Z'w'u\ \ \forall\, y \in y'\ \ \alpha(Xw'u, y, z))$$

Take $t_X := \lambda\, w', u \,.\, u$, $t_{Z'} := \lambda\, w', u \,.\, w'$ and $t_{V'} := \lambda\, w', u, y' \,.\, \{y'\}$. We want to show:

$$\forall\, w', u, y'\, (\forall\, v' \in \{y'\}\ \ \exists\, w \in w'\ \ \forall\, v \in v'\ \ \alpha(u, v, w) \to \exists\, z \in w'\ \ \forall\, y \in y'\ \ \alpha(u, y, z))$$

Let w', u, y' be arbitrary and take $v' := y'$. We have as assumption:

$$\exists\, w \in w'\ \ \forall\, v \in y'\ \ \alpha(u, v, w)$$

and need to show:

$$\exists\, z \in w'\ \ \forall\, y \in y'\ \ \alpha(u, y, z)$$

which are the same formulas, after renaming of bound variables.

$\boxed{\mathsf{st}\, z \wedge z = w \to \mathsf{st}\, w}$

$$(\mathsf{st}\, z \wedge z = w \to \mathsf{st}\, w)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, W'\ \forall^{\mathsf{st}}\, z'\, (z \in z' \wedge z = w \to w \in W'z')$$

Take $t_{W'} := \lambda\, z' \,.\, z'$. We need to show:

$$\forall\, z'\, (z \in z' \wedge z = w \to w \in z')$$

which is a consequence of Proposition 3.1.1.

$\boxed{\mathsf{st}\, u,\ u \text{ closed}}$

$$(\mathsf{st}\, u)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, u'\, (u \in u')$$

Take $t_{u'} := \{u\}$. We need to show $u \in \{u\}$, which is an instance of Lemma 3.1.7.1. Note that u needs to be closed, otherwise $\{u\}$ wouldn't be closed, as is required.

$\boxed{\mathsf{st}\, f \wedge \mathsf{st}\, z \to \mathsf{st}(fz)}$

$$(\mathsf{st}\, f \wedge \mathsf{st}\, z \to \mathsf{st}(fz))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, U\ \forall^{\mathsf{st}}\, f', z'\, (f \in f' \wedge z \in z' \to fz \in Uf'z')$$

Take:

$$t_U := \lambda\, f', z' \,.\, \textstyle\bigcup_{f \in f'} \bigcup_{z \in z'} fz$$

We need to show:

$$\forall\, f', z'\, (f \in f' \wedge z \in z' \to fz \in \textstyle\bigcup_{f \in f'} \bigcup_{z \in z'} fz)$$

which is a consequence of Lemma 3.1.7.3.

$\boxed{\text{Internal axioms}}$

From Lemma 3.3.2, the H_{st}-interpretation of internal formulas leaves them unchanged. This means that this result follows trivially for such formulas. The internal axioms that get taken care of are the equality axioms, the axiom of internal induction, the axioms for S, for Π, for Σ, for $\boldsymbol{R}$, and for $\boldsymbol{L}$, and the sequence axiom.

$\boxed{\mathsf{I}^\omega}$

Recall that, in the following, z' and u represent tuples of variables. Using Lemma 3.3.2.2:

$$(\forall^{\mathsf{st}} z' \, \exists w \, \forall z \in z' \;\; \varphi(z,w) \to \exists v \, \forall^{\mathsf{st}} u \, \varphi(u,v))^{H_{\mathsf{st}}} \leftrightarrow$$
$$\leftrightarrow \exists^{\mathsf{st}} Z' \, \forall^{\mathsf{st}} u' \, (\forall z' \in Z'u' \;\; \exists w \, \forall z \in z' \;\; \varphi(z,w) \to \exists v \, \forall u \in u' \;\; \varphi(u,v))$$

Take $t_{Z'} := \lambda u' . \{u'\}$. We need to show:

$$\forall u' \, (\forall z' \in \{u'\} \;\; \exists w \, \forall z \in z' \;\; \varphi(z,w) \to \exists v \, \forall u \in u' \;\; \varphi(u,v)) \tag{3.3.24}$$

Let u' be arbitrary, and assume the antecedent of (3.3.24). Let $z' := u'$. We get

$$\exists w \, \forall z \in u' \;\; \varphi(z,w)$$

which is precisely what we wanted, after renaming of bound variables.

$\boxed{\mathsf{NCR}^\omega}$

$$(\forall w \, \exists^{\mathsf{st}} z \, A(z,w) \to \exists^{\mathsf{st}} s' \, \forall r \, \exists s \in s' \;\; A(s,r))^{H_{\mathsf{st}}} \leftrightarrow$$
$$\leftrightarrow \exists^{\mathsf{st}} U, S'', Y' \, \forall^{\mathsf{st}} x, z', v'' \, ($$
$$\forall y' \in Y'xz'v'' \;\; \forall w \, \exists z \in z' \;\; \forall y \in y' \;\; \alpha(x,y,z,w)$$
$$\to$$
$$\exists s' \in S''xz' \;\; \forall v' \in v'' \;\; \forall r \, \exists s \in s' \;\; \forall v \in v' \;\; \alpha(Uxz', v, s, r))$$

Take $t_U := \lambda x, z' . x$, $t_{S''} := \lambda x, z' . \{z'\}$ and $t_{Y'} := \lambda x, z', v'' . v''$. We need to show:

$$\forall x, z', v'' \, (\forall y' \in v'' \;\; \forall w \, \exists z \in z' \;\; \forall y \in y' \;\; \alpha(x,y,z,w)$$
$$\to$$
$$\exists s' \in \{z'\} \;\; \forall v' \in v'' \;\; \forall r \, \exists s \in s' \;\; \forall v \in v' \;\; \alpha(x,v,s,r))$$

Let z, x', v'' be arbitrary, and assume the antecedent. Take $s' := z'$ in the consequent. So currently we have as assumption:

$$\forall y' \in v'' \;\; \forall w \, \exists z \in z' \;\; \forall y \in y' \;\; \alpha(x,y,z,w)$$

and need to show:

$$\forall v' \in v'' \;\; \forall r \, \exists s \in z' \;\; \forall v \in v' \;\; \alpha(x,v,s,r)$$

These are the same formulas, after renaming of bound variables.

$\boxed{\mathsf{HAC}^\omega}$

$$(\forall^{\mathsf{st}} z \, \exists^{\mathsf{st}} w \, A(z,w) \to \exists^{\mathsf{st}} S \, \forall^{\mathsf{st}} r \, \exists s \in Sr \;\; A(r,s))^{H_{\mathsf{st}}} \leftrightarrow$$
$$\leftrightarrow \exists^{\mathsf{st}} \mathcal{U}, \mathcal{S}', Z, Y' \, \forall^{\mathsf{st}} X, W', r', v'' \, ($$
$$\forall z \in ZXW'r'v'' \;\; \forall y' \in Y'XW'r'v'' \;\; \exists w \in W'z \;\; \forall y \in y' \;\; \alpha(Xz, y, z, w)$$
$$\to$$
$$\exists S \in \mathcal{S}'XW' \;\; \forall r \in r' \;\; \forall v' \in v'' \;\; \exists s \in Sr \;\; \forall v \in v' \;\; \alpha(\mathcal{U}XW'r, v, r, s))$$

Take:

$$t_{\mathcal{U}} := \lambda\, X, W' \,.\, X \qquad\qquad t_Z := \lambda\, X, W', r', v'' \,.\, r'$$
$$t_{\mathcal{S}'} := \lambda\, X, W' \,.\, \{W'\} \qquad t_{Y'} := \lambda\, X, W', r', v'' \,.\, v''$$

We need to show:

$$\begin{aligned}\forall\, X, W', r', v''\, (\forall\, z \in r' \ \ \forall\, y' \in v'' \ \ \exists\, w \in W'z \ \ \forall\, y \in y' \ \ \alpha(Xz, y, z, w)\\ \to\\ \exists\, S \in \{W'\} \ \ \forall\, r \in r' \ \ \forall\, v' \in v'' \ \ \exists\, s \in Sr \ \ \forall\, v \in v' \ \ \alpha(Xr, v, r, s))\end{aligned}$$

Let X, W', r', v'' be arbitrary, and assume the antecedent. Take $S := W'$ in the consequent. So we currently have as assumption:

$$\forall\, z \in r' \ \ \forall\, y' \in v'' \ \ \exists\, w \in W'z \ \ \forall\, y \in y' \ \ \alpha(Xz, y, z, w)$$

and need to show:

$$\forall\, r \in r' \ \ \forall\, v' \in v'' \ \ \exists\, s \in W'r \ \ \forall\, v \in v' \ \ \alpha(Xr, v, r, s)$$

These are exactly the same formulas, after renaming of bound variables.

$\boxed{\mathsf{HGMP}^{\omega}_{\mathsf{st}}}$

$$\begin{aligned}&((\forall^{\mathsf{st}}\, z\, \varphi(z) \to \psi) \to \exists^{\mathsf{st}}\, w'\, (\forall\, w \in w' \ \ \varphi(w) \to \psi))^{H_{\mathsf{st}}} \equiv\\ &\quad \equiv \exists^{\mathsf{st}}\, W''\, \forall^{\mathsf{st}}\, z'\, ((\forall\, z \in z' \ \ \varphi(z) \to \psi) \to \exists\, w' \in W''z' \ \ (\forall\, w \in w' \ \ \varphi(w) \to \psi))\end{aligned}$$

Take $t_{W''} := \lambda\, z' \,.\, \{z'\}$. We need to show:

$$\forall\, z'\, ((\forall\, z \in z' \ \ \varphi(z) \to \psi) \to \exists\, w' \in \{z'\} \ \ (\forall\, w \in w' \ \ \varphi(w) \to \psi))$$

Let z' be arbitrary, and assume the antecedent. Take $w' := z'$ in the consequent. We currently have as assumption:

$$\forall\, z \in z' \ \ \varphi(z) \to \psi$$

and want to show:

$$\forall\, w \in z' \ \ \varphi(w) \to \psi$$

These are the same formulas, after renaming of a bound variable.

$\boxed{\mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}}$

$$\begin{aligned}&((\forall^{\mathsf{st}}\, z\, \varphi(z) \to \exists^{\mathsf{st}}\, w\, A(w)) \to \exists^{\mathsf{st}}\, s'\, (\forall^{\mathsf{st}}\, r\, \varphi(r) \to \exists\, s \in s' \ \ A(s)))^{H_{\mathsf{st}}} \leftrightarrow\\ &\quad \leftrightarrow \exists^{\mathsf{st}}\, U, \mathcal{R}, S'', Y'\ \forall^{\mathsf{st}}\, x, w', Z, v''\, (\\ &\qquad \forall\, y' \in Y'xw'Zv'' \ \ (\forall\, z \in Zy' \ \ \varphi(z) \to \exists\, w \in w' \ \ \forall\, y \in y' \ \ \alpha(x, y, w))\\ &\qquad \to\\ &\qquad \exists\, s' \in S''xw'Z \ \ \forall\, v' \in v'' \ \ (\forall\, r \in \mathcal{R}xw'Zv' \ \ \varphi(r) \to \exists\, s \in s' \ \ \forall\, v \in v' \ \ \alpha(Uxw'Z, v, s)))\end{aligned}$$

Take:

$$t_U := \lambda\, x, w', Z\,.\, x \qquad t_{S''} := \lambda\, x, w', Z\,.\, \{w'\}$$
$$t_{\mathcal{R}} := \lambda\, x, w', Z\,.\, Z \qquad t_{Y'} := \lambda\, x, w', Z, v''\,.\, v''$$

We need to show:

$$\begin{aligned}\forall\, x, w', Z, v''\, (&\forall\, y' \in v'' \;\; (\forall\, z \in Zy' \;\; \varphi(z) \to \exists\, w \in w' \;\; \forall\, y \in y' \;\; \alpha(x, y, w))\\ &\to\\ &\exists\, s' \in \{w'\} \;\; \forall\, v' \in v'' \;\; (\forall\, r \in Zv' \;\; \varphi(r) \to \exists\, s \in s' \;\; \forall\, v \in v' \;\; \alpha(x, v, s)))\end{aligned}$$

Let x, w', Z, v'' be arbitrary and assume the antecedent. Take $s' := w'$ in the consequent. We have as assumption:

$$\forall\, y' \in v'' \;\; (\forall\, z \in Zy' \;\; \varphi(z) \to \exists\, w \in w' \;\; \forall\, y \in y' \;\; \alpha(x, y, w))$$

and we need to show:

$$\forall\, v' \in v'' \;\; (\forall\, r \in Zv' \;\; \varphi(r) \to \exists\, s \in w' \;\; \forall\, v \in v' \;\; \alpha(x, v, s))$$

These are the same formulas, after renaming of bound variables.

$\boxed{\Delta_{\forall^{\mathsf{st}}}}$

Let $A \equiv \forall^{\mathsf{st}}\, z\, \varphi(z)$ be a $\forall^{\mathsf{st}}$-formula in $\Delta_{\forall^{\mathsf{st}}}$. Then by Lemma 3.3.2.2, $A^{H_{\mathsf{st}}} \equiv A$. This means that we need to show

$$\forall\, z\, \varphi(z) \tag{3.3.25}$$

in $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}}$. But (3.3.25) is the internalization of A, so it is a formula of $\Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}}$.

□

Corollary 3.3.10 (Conservativity)**.** $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ is a conservative extension of $\mathsf{E\text{-}HA}^{\omega*}$. In other words, if $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ proves a formula in the language of $\mathsf{E\text{-}HA}^{\omega*}$, then $\mathsf{E\text{-}HA}^{\omega*}$ already proves it.

Proof. Follows from Theorem 3.3.9, since internal formulas (formulas of the language of $\mathsf{E\text{-}HA}^{\omega*}$) don't get changed by the H_{st}-interpretation. □

Corollary 3.3.10 further implies that $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ is consistent relative to $\mathsf{E\text{-}HA}^{\omega*}$, since $\bot$ is an internal formula.

Theorem 3.3.11 (Program extraction by the H_{st}-interpretation)**.** Let $\forall^{\mathsf{st}}\, x\, \exists^{\mathsf{st}}\, y\, \varphi(x, y)$ be a sentence in the language of $\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}}$, with φ internal. Furthermore, let $\Delta_{\forall^{\mathsf{st}}}$ be a collection of $\forall^{\mathsf{st}}$-formulas, and $\Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}}$ their internalizations. Suppose:

$$\mathsf{E\text{-}HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}} + \Delta_{\forall^{\mathsf{st}}} \vdash \forall^{\mathsf{st}}\, x\, \exists^{\mathsf{st}}\, y\, \varphi(x, y)$$

Then there is a closed term t such that:

$$\mathsf{E\text{-}HA}^{\omega*} + \Delta^{\mathsf{int}}_{\forall^{\mathsf{st}}} \vdash \forall\, x\, \exists\, y \in tx \;\; \varphi(x, y)$$

Proof. Direct consequence of Theorem 3.3.9, taking into account that:

$$(\forall^{\mathsf{st}} x\ \exists^{\mathsf{st}} y\, \varphi(x,y))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} Y\ \forall^{\mathsf{st}} x\ \exists y \in Yx\ \ \varphi(x,y)$$

□

Theorem 3.3.12 (Characterization theorem for the H_{st}-interpretation)**.** For all formulas of the language of E-HA$^{\omega*}_{\mathsf{st}}$:

$$\text{E-HA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}} \vdash A \leftrightarrow A^{H_{\mathsf{st}}}$$

Proof. By induction on the logical structure of A.

Internal atomic formulas

For an internal atomic formula φ, $\varphi^{H_{\mathsf{st}}} \equiv \varphi$, so there is nothing to show.

st z

$$(\mathsf{st}\, z)^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} z'\, (z \in z')$$

To show $\mathsf{st}\, z \to \exists^{\mathsf{st}} z'\, (z \in z')$, take $z' := \{z\}$. This is a standard term by Remark 3.1.22, and does the job by Lemma 3.1.7.1.

The other implication, $\exists^{\mathsf{st}} z'\, (z \in z') \to \mathsf{st}\, z$, results from Lemma 3.1.24.

For the steps of the induction dealing with formulas A and B, we follow [2] in noting that, according to Section 1.6.17 of [43], there are internal formulas α and β such that E-HA$^{\omega*}_{\mathsf{st}}$ proves:

$$A^{H_{\mathsf{st}}} \leftrightarrow \exists^{\mathsf{st}} x\ \forall^{\mathsf{st}} y\, \alpha(x,y) \quad \text{and} \quad B^{H_{\mathsf{st}}} \leftrightarrow \exists^{\mathsf{st}} u\ \forall^{\mathsf{st}} v\, \beta(u,v)$$

These formulas are obtained through embeddings of tuples of types in a higher type, and coding with inverses of tuples of terms in a single term.

With this in mind, we can take the tuples of variables given by the H_{st}-interpretation as a single variable.

$A \wedge B$

Since by induction hypothesis $A \leftrightarrow A^{H_{\mathsf{st}}}$ and $B \leftrightarrow B^{H_{\mathsf{st}}}$, it suffices to show

$$A^{H_{\mathsf{st}}} \wedge B^{H_{\mathsf{st}}} \leftrightarrow (A \wedge B)^{H_{\mathsf{st}}}$$

$$\exists^{\mathsf{st}} x\ \forall^{\mathsf{st}} y\, \alpha(x,y) \wedge \exists^{\mathsf{st}} u\ \forall^{\mathsf{st}} v\, \beta(u,v) \leftrightarrow \exists^{\mathsf{st}} x, u\ \forall^{\mathsf{st}} y, v\, (\alpha(x,y) \wedge \beta(u,v))$$

which is straightforward in E-HA$^{\omega*}_{\mathsf{st}}$.

$A \vee B$

We need to show:

$$A^{H_{\mathsf{st}}} \vee B^{H_{\mathsf{st}}} \leftrightarrow (A \vee B)^{H_{\mathsf{st}}}$$

$$\exists^{\mathsf{st}} x\ \forall^{\mathsf{st}} y\, \alpha(x,y) \vee \exists^{\mathsf{st}} u\ \forall^{\mathsf{st}} v\, \beta(u,v) \leftrightarrow \exists^{\mathsf{st}} x, u\ \forall^{\mathsf{st}} y', v'\, (\forall y \in y'\ \ \alpha(x,y) \vee \forall v \in v'\ \ \beta(u,v))$$

From left to right, this is straightforward in E-HA$^{\omega*}_{\mathsf{st}}$. From right to left we need LLPO$^{\omega}_{\mathsf{st}}$, which is a consequence of I^{ω} by Lemma 3.2.11.

$\boxed{A \to B}$

Consider the following enumeration of formulas, starting with $A^{H_{\text{st}}} \to B^{H_{\text{st}}}$ and ending with $(A \to B)^{H_{\text{st}}}$:

(1) $\exists^{\text{st}} x\, \forall^{\text{st}} y\, \alpha(x,y) \to \exists^{\text{st}} u\, \forall^{\text{st}} v\, \beta(u,v)$

(2) $\forall^{\text{st}} x\, (\forall^{\text{st}} y\, \alpha(x,y) \to \exists^{\text{st}} u\, \forall^{\text{st}} v\, \beta(u,v))$

(3) $\forall^{\text{st}} x\, \exists^{\text{st}} u\, (\forall^{\text{st}} y\, \alpha(x,y) \to \forall^{\text{st}} v\, \beta(u,v))$

(4) $\forall^{\text{st}} x\, \exists^{\text{st}} u\, \forall^{\text{st}} v\, (\forall^{\text{st}} y\, \alpha(x,y) \to \beta(u,v))$

(5) $\forall^{\text{st}} x\, \exists^{\text{st}} u\, \forall^{\text{st}} v\, \exists^{\text{st}} y'\, (\forall\, y \in y' \;\; \alpha(x,y) \to \beta(u,v))$

(6) $\exists^{\text{st}} U\, \forall^{\text{st}} x\, \forall^{\text{st}} v\, \exists^{\text{st}} y'\, (\forall\, y \in y' \;\; \alpha(x,y) \to \beta(Ux,v))$

(7) $\exists^{\text{st}} U\, \forall^{\text{st}} x\, \exists^{\text{st}} Y\, \forall^{\text{st}} v\, (\forall\, y \in Yv \;\; \alpha(x,y) \to \beta(Ux,v))$

(8) $\exists^{\text{st}} U, \mathcal{Y}\, \forall^{\text{st}} x, v\, (\forall\, y \in \mathcal{Y}xv \;\; \alpha(x,y) \to \beta(Ux,v))$

Clearly it suffices to show $(i) \leftrightarrow (i+1)$ for all $i \in \{1,\ldots,7\}$. All the implications $(i+1) \to (i)$ are easily provable in E-HA$^{\omega*}_{\text{st}}$ (Lemma 3.1.24 is necessary for $(5) \to (4)$). The implications $(1) \to (2)$ and $(3) \to (4)$ only need E-HA$^{\omega*}_{\text{st}}$ as well.

The implication $(2) \to (3)$ is a consequence of HIP$^{\omega}_{\forall^{\text{st}}}$, taking into account that $\forall^{\text{st}} v\, \beta(u,v)$ is monotone on u, and $(4) \to (5)$ comes form HGMP$^{\omega}_{\text{st}}$.

As for the remaining three implications, they all follow from HAC$^{\omega}$ and the following facts:

- $\forall^{\text{st}} v\, \exists^{\text{st}} y'\, (\forall\, y \in y' \;\; \alpha(x,y) \to \beta(u,v))$ is monotone on u, consequence of the monotonicity of β on the same variable;
- $\forall\, y \in y' \;\; \alpha(x,y) \to \beta(Ux,v)$ is monotone on y', consequence of Lemma 3.1.10.1;
- $\forall^{\text{st}} v\, (\forall\, y \in Yv \;\; \alpha(x,y) \to \beta(Ux,v))$ is monotone on Y, again consequence of Lemma 3.1.10.1.

$\boxed{\forall\, z\, A(z)}$

We need to show:

$$\forall\, z\, A(z)^{H_{\text{st}}} \leftrightarrow (\forall\, z\, A(z))^{H_{\text{st}}}$$

$$\forall\, z\, \exists^{\text{st}} x\, \forall^{\text{st}} x\, \alpha(x,y,z) \leftrightarrow \exists^{\text{st}} x\, \forall^{\text{st}} y\, \forall\, z\, \alpha(x,y,z)$$

From left to right we use the monotonicity of $\forall^{\text{st}} y\, \alpha(x,y,z)$ on x, and NCR$^{\omega}$. From right to left E-HA$^{\omega*}_{\text{st}}$ suffices.

$\boxed{\exists\, z\, A(z)}$

We need to show:

$$\exists\, z\, A(z)^{H_{\text{st}}} \leftrightarrow (\exists\, z\, A(z))^{H_{\text{st}}}$$

$$\exists\, z\, \exists^{\text{st}} x\, \forall^{\text{st}} y\, \alpha(x,y,z) \leftrightarrow \exists^{\text{st}} x\, \forall^{\text{st}} y'\, \exists\, z\, \forall\, y \in y' \;\; \alpha(x,y,z)$$

From left to right E-HA$^{\omega*}_{\text{st}}$ is enough (using Lemma 3.1.24). From right to left we use I$^{\omega}$.

$\boxed{\forall^{\mathsf{st}} z\, A(z)}$

We need to show:

$$\forall^{\mathsf{st}} z\, A(z)^{H_{\mathsf{st}}} \leftrightarrow (\forall^{\mathsf{st}} z\, A(z))^{H_{\mathsf{st}}}$$
$$\forall^{\mathsf{st}} z\; \exists^{\mathsf{st}} x\; \forall^{\mathsf{st}} y\, \alpha(x,y,z) \leftrightarrow \exists^{\mathsf{st}} X\; \forall^{\mathsf{st}} z, y\, \alpha(Xz,y,z)$$

From left to right we use HAC^{ω} and the monotonicity of $\forall^{\mathsf{st}} y\, \alpha(x,y,z)$ on x. From right to left $\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}}$ suffices.

$\boxed{\exists^{\mathsf{st}} z\, A(z)}$

We need to show:

$$\exists^{\mathsf{st}} z\, A(z)^{H_{\mathsf{st}}} \leftrightarrow (\exists^{\mathsf{st}} z\, A(z))^{H_{\mathsf{st}}}$$
$$\exists^{\mathsf{st}} z\; \exists^{\mathsf{st}} x\; \forall^{\mathsf{st}} y\, \alpha(x,y,z) \leftrightarrow \exists^{\mathsf{st}} z', x\; \forall^{\mathsf{st}} y'\; \exists\, z \in z' \quad \forall\, y \in y' \quad \alpha(x,y,z)$$

Similarly to the case for $\exists$, from left to right we use Lemma 3.1.24 and from right to left, the principle I^{ω}.

□

We can prove that there are no characteristic principles missing in the statement of the soundness theorem for the H_{st}-interpretation (Theorem 3.3.9):

Theorem 3.3.13. Let $\mathsf{H} := \mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}}$ be the theory in which we can prove the characterization theorem for the H_{st}-interpretation.

Let C be a characteristic principle, *i.e.*, a principle for which it is possible to prove that for every formula A of the language of $\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}}$ with $A^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}} \boldsymbol{x}\; \forall^{\mathsf{st}} \boldsymbol{y}\, \alpha_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y})$, if

$$\mathsf{H} + C \vdash A$$

then there are closed terms of the language of $\mathsf{E\text{-}HA}^{\omega *}$ such that

$$\mathsf{E\text{-}HA}^{\omega *} \vdash \forall\, \boldsymbol{y}\, \alpha_{H_{\mathsf{st}}}(\boldsymbol{t}, \boldsymbol{y}).$$

Then C is already provable from H.

Proof. Let $C^{H_{\mathsf{st}}} \equiv \forall^{\mathsf{st}} \boldsymbol{x}\; \exists^{\mathsf{st}} \boldsymbol{y}\, \gamma_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y})$. Since clearly $\mathsf{H} + C \vdash C$, we know that there are closed terms $\boldsymbol{t}$ such that:

$$\mathsf{E\text{-}HA}^{\omega *} \vdash \forall\, \boldsymbol{y}\, \gamma_{H_{\mathsf{st}}}(\boldsymbol{t}, \boldsymbol{y})$$

Then $\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}}$ proves the same statement. In particular:

$$\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}} \vdash \forall^{\mathsf{st}} \boldsymbol{y}\, \gamma_{H_{\mathsf{st}}}(\boldsymbol{t}, \boldsymbol{y})$$

which is a weaker statement. Now, since the terms $\boldsymbol{t}$ are closed, they are standard, and hence:

$$\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}} \vdash \exists^{\mathsf{st}} \boldsymbol{x}\; \forall^{\mathsf{st}} \boldsymbol{y}\, \gamma_{H_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y})$$
$$\mathsf{E\text{-}HA}^{\omega *}_{\mathsf{st}} \vdash C^{H_{\mathsf{st}}}$$

Now since the theory H is stronger than E-HA$^{\omega*}_{\text{st}}$, we also have H $\vdash C^{H_{\text{st}}}$, and by the characterization theorem (Theorem 3.3.12), H $\vdash C$ as we wanted to show. $\square$

From the characterization theorems of the H_{st}-interpretation (Theorem 3.3.12) and of the D_{st}-interpretation presented in [2], we get:

Theorem 3.3.14. For all formulas A of the language of E-HA$^{\omega*}_{\text{st}}$:

$$\text{E-HA}^{\omega*}_{\text{st}} + \text{I}^{\omega} + \text{NCR}^{\omega} + \text{HAC}^{\omega} + \text{HGMP}^{\omega}_{\text{st}} + \text{HIP}^{\omega}_{\forall^{\text{st}}} \vdash A^{H_{\text{st}}} \leftrightarrow A^{D_{\text{st}}}$$

As is the case with the *dialectica* interpretation of Gödel, the H_{st}-interpretation could be coupled with a negative translation to bring it into the realm of Peano arithmetic. It would be interesting to find the negative translation that would produce the S_{st}-interpretation defined below, in the same spirit as Streicher and Kohlenbach did for the interpretations of Gödel and Shoenfield in [41]. In fact, in [2] Van den Berg et al. did just that for their interpretations, which are very similar to the ones described in this thesis.

3.4 The S_{st}-interpretation

The S_{st}-interpretation is an interpretation of classical nonstandard arithmetic based on Shoenfield's interpretation (outlined in Section 2.3). It was proposed in its first version by Van den Berg, Briseid, and Safarik in [2]. Later it was slightly modified by Dinis and Ferreira in [10]. In their paper, Dinis and Ferreira consider a weaker theory than E-PA$^{\omega*}_{\text{st}}$, namely E-PRA$^{\omega*}_{\text{st}}$, where the only recursor available is of type 0, and induction is restricted to quantifier-free formulas. The version here presented is an extension of Dinis and Ferreira's S_{st}-interpretation to E-PA$^{\omega*}_{\text{st}}$.

Definition 3.4.1 (S_{st}-interpretation)**.** The S_{st}-interpretation associates to each formula A of the language of E-PA$^{\omega*}_{\text{st}}$ a formula $A^{S_{\text{st}}}$ of the form:

$$A^{S_{\text{st}}} \equiv \forall^{\text{st}} \boldsymbol{x}\, \exists^{\text{st}} \boldsymbol{y}\, \alpha_{S_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y})$$

with the same free variables and in the same language as A. The (possibly empty) variable tuples $\boldsymbol{x}$ and $\boldsymbol{y}$ and their types are uniquely determined by the logical structure of A. It is important that these variables do not appear free in A. Furthermore, $\alpha_{S_{\text{st}}}$ is an internal formula.

The definition proper is given below. The sub-formulas inside square brackets are the internal formulas corresponding to $\alpha_{S_{\text{st}}}$ above.

- $\alpha^{S_{\text{st}}} :\equiv \alpha_{S_{\text{st}}} :\equiv \alpha$ for internal atomic formulas α
- $(\text{st}\, z)^{S_{\text{st}}} :\equiv \exists^{\text{st}} z' \, [z \in z']$

Given the interpretations $A^{S_{\text{st}}} \equiv \forall^{\text{st}} \boldsymbol{x}\, \exists^{\text{st}} \boldsymbol{y}\, \alpha_{S_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y})$ and $B^{S_{\text{st}}} \equiv \forall^{\text{st}} \boldsymbol{u}\, \exists^{\text{st}} \boldsymbol{v}\, \beta_{S_{\text{st}}}(\boldsymbol{u}, \boldsymbol{v})$:

- $(A \vee B)^{S_{\text{st}}} :\equiv \forall^{\text{st}} \boldsymbol{x}, \boldsymbol{u}\, \exists^{\text{st}} \boldsymbol{y}, \boldsymbol{v}\, [\alpha_{S_{\text{st}}}(\boldsymbol{x}, \boldsymbol{y}) \vee \beta_{S_{\text{st}}}(\boldsymbol{u}, \boldsymbol{v})]$
- $(\neg A)^{S_{\text{st}}} :\equiv \forall^{\text{st}} \boldsymbol{Y}\, \exists^{\text{st}} \boldsymbol{x'}\, [\exists \boldsymbol{x} \in \boldsymbol{x'}\;\, \neg \alpha_{S_{\text{st}}}(\boldsymbol{x}, \boldsymbol{Y}\boldsymbol{x})]$

- $(\forall z\, A(z))^{S_{\mathsf{st}}} :\equiv \forall^{\mathsf{st}} \boldsymbol{x}\, \exists^{\mathsf{st}} \boldsymbol{y}\, [\forall z\, \alpha_{S_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y}, z)]$

Remark 3.4.2. Internal formulas remain unchanged after the S_{st}-interpretation.

Lemma 3.4.3. Let A be an arbitrary formula of the language of E-PA$^{\omega *}_{\mathsf{st}}$ such that

$$A^{S_{\mathsf{st}}} \equiv \forall^{\mathsf{st}} \boldsymbol{x}\, \exists^{\mathsf{st}} \boldsymbol{y}\, \alpha(\boldsymbol{x}, \boldsymbol{y})$$

Then $\boldsymbol{y}$ is of star type and α is monotone on $\boldsymbol{y}$.

Idea of Proof. Straightforward induction on the logical structure of A. □

Definition 3.4.4 (HAC$^{\omega}_{\mathsf{int}}$)**.** The herbrandised schema of choice for internal formulas, HAC$^{\omega}_{\mathsf{int}}$, is the union for all types ρ, τ of:

$$\mathsf{HAC}^{\rho,\tau}_{\mathsf{int}} : \forall^{\mathsf{st}} z^{\rho}\, \exists^{\mathsf{st}} w^{\tau}\, \varphi(z, w) \to \exists^{\mathsf{st}} W'^{\rho \to \tau^*}\, \forall^{\mathsf{st}} z^{\rho}\, \exists w \in W'z\;\; \varphi(z, w)$$

where φ is any internal formula in the language of E-PA$^{\omega *}$.

Theorem 3.4.5 (Soundness of the S_{st}-interpretation)**.** Let A be an arbitrary formula in the language of E-PA$^{\omega *}_{\mathsf{st}}$, possibly with free variables, such that $A^{S_{\mathsf{st}}} \equiv \forall^{\mathsf{st}} \boldsymbol{x}\, \exists^{\mathsf{st}} \boldsymbol{y}\, \alpha_{S_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{y})$. Furthermore, let Δ_{int} be a collection of internal formulas. Suppose:

$$\mathsf{E\text{-}PA}^{\omega *}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{HAC}^{\omega}_{\mathsf{int}} + \Delta_{\mathsf{int}} \vdash A$$

Then there are closed terms $\boldsymbol{t}$, which can be extracted from a proof of A, such that:

$$\mathsf{E\text{-}PA}^{\omega *} + \Delta_{\mathsf{int}} \vdash \forall \boldsymbol{x}\, \alpha_{S_{\mathsf{st}}}(\boldsymbol{x}, \boldsymbol{t}\boldsymbol{x})$$

Idea of Proof. By induction on the derivation of A, using the monotonicity of $\alpha_{S_{\mathsf{st}}}$ in the existentially quantified variables. The proof can be found in [10]. Even though they use a weaker theory than here, namely E-PRA$^{\omega *}_{\mathsf{st}}$, the only extra axiom that we need to check is external induction. For convenience, we interpret the external induction rule, which is equivalent, instead:

$$\frac{A(0) \quad \forall^{\mathsf{st}} n^0\, (A(n) \to A(Sn))}{\forall^{\mathsf{st}} n^0\, A(n)}$$

Suppose $A(n)^{S_{\mathsf{st}}} \equiv \forall^{\mathsf{st}} x^{\rho}\, \exists^{\mathsf{st}} y^{\tau^*}\, \alpha(x, y, n)$ where x, y should be thought of as tuples of variables. During the rest of this proof we discard the tuple notation for simplicity.

By induction hypothesis, there are closed terms Y of type $\rho \to \tau^*$, U' of type $0^* \to (\rho \to \tau^*) \to \rho \to \rho^*$, and Q of type $0^* \to (\rho \to \tau^*) \to \rho \to \tau^*$ such that:

$$\forall x\, \alpha(x, Yx, 0) \tag{3.4.1}$$

$$\forall n'^{0^*}, V, p\, \forall n \in n'\;\; (\forall u \in U'n'Vp\;\; \alpha(u, Vu, n) \to \alpha(p, Qn'Vp, Sn)) \tag{3.4.2}$$

And we need to find a closed term W of type $0^* \to \rho \to \tau^*$ such that:

$$\forall m'^{0^*}, z\, \forall m \in m'\;\; \alpha(z, Wm'z, m) \tag{3.4.3}$$

We want to use a similar strategy to the one used in the treatment of the external induction rule during the proof of Theorem 3.3.9. Ideally, the formulas (3.4.2) and (3.4.3) would start with a universal quantification over a term of type 0, instead of with a universal quantification over a term of type 0^*. In fact, we can massage said formulas until they are in our preferred form:

1. Let $\mathcal{U}' := \lambda\, n^0 \,.\, U'\{n\}$ and $\mathcal{Q} := \lambda\, n^0 \,.\, Q\{n\}$. Then from (3.4.2) it is possible to prove:

$$\forall\, n^0, V, p\, (\forall\, u \in \mathcal{U}'nVp\ \ \alpha(u, Vu, n) \to \alpha(p, \mathcal{Q}nVp, Sn)) \tag{3.4.4}$$

2. Given $\mathcal{W}$ of type $0 \to \rho \to \tau^*$ such that

$$\forall\, m^0, z\, \alpha(z, \mathcal{W}mz, m) \tag{3.4.5}$$

it is possible to prove (3.4.3) with $W := \lambda\, m'^{0^*} \,.\, \bigsqcup_{k \in m'} \mathcal{W}k$.

Now we can reformulate the goal: from (3.4.1) and (3.4.4), find $\mathcal{W}$ of type $0 \to \rho \to \tau^*$ such that (3.4.5).

Take $\mathcal{W} := \lambda\, m^0 \,.\, RmY\mathcal{Q}$. From the recursor axioms we get:

$$\mathcal{W}0 = Y$$
$$\mathcal{W}(Sm) = \mathcal{Q}m(\mathcal{W}m)$$

We show (3.4.5) by induction on m. The base case with $m = 0$ is a direct consequence of (3.4.1). For the step, assume:

$$\forall\, z\, \alpha(z, \mathcal{W}mz, m) \tag{3.4.6}$$

and let z_0 be arbitrary. We need to show $\alpha(z_0, \mathcal{W}(Sm)z_0, Sm)$. Unravelling $\mathcal{W}(Sm)$, we see that the goal is really $\alpha(z_0, \mathcal{Q}m(\mathcal{W}m)z_0, Sm)$. So take in (3.4.4) $n := m$, $V := \mathcal{W}m$, $p := z_0$, obtaining:

$$\forall\, u \in \mathcal{U}'m(\mathcal{W}m)z_0\ \ \alpha(u, \mathcal{W}mu, m) \to \alpha(z_0, \mathcal{Q}m(\mathcal{W}m)z_0, Sm) \tag{3.4.7}$$

Noticing that (3.4.6) is stronger than the antecedent of (3.4.7), we can conclude its consequent, which was precisely our goal. □

Theorem 3.4.6 (Characterization theorem for the S_{st}-interpretation)**.** For all formulas of the language of $\text{E-PA}^{\omega*}_{\mathsf{st}}$:

$$\text{E-PA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{HAC}^{\omega}_{\mathsf{int}} \vdash A \leftrightarrow A^{S_{\mathsf{st}}}$$

Proof. By induction on the logical structure of A.

Internal atomic formulas

For an internal atomic formula φ, $\varphi^{S_{\mathsf{st}}} \equiv \varphi$, so there is nothing to show.

$\mathsf{st}\, z$

Since $(\mathsf{st}\, z)^{S_{\mathsf{st}}} \equiv (\mathsf{st}\, z)^{H_{\mathsf{st}}}$, this step coincides with the step for $\mathsf{st}\, z$ in the proof of Theorem 3.3.12 (where only $\text{E-HA}^{\omega*}_{\mathsf{st}}$ is used).

As in the proof of Theorem 3.3.12, we note that there are internal formulas α and β such that $\text{E-PA}^{\omega*}_{\mathsf{st}}$ proves:

$$A^{S_{\mathsf{st}}} \leftrightarrow \forall^{\mathsf{st}}\, x\, \exists^{\mathsf{st}}\, y\, \alpha(x, y) \quad \text{and} \quad B^{S_{\mathsf{st}}} \leftrightarrow \forall^{\mathsf{st}}\, u\, \exists^{\mathsf{st}}\, v\, \beta(u, v)$$

and hence can take the tuples of variable given by the S_{st}-interpretation as a single variable.

$\boxed{A \vee B}$

By induction hypothesis, $A \leftrightarrow A^{S_{\text{st}}}$ and $B \leftrightarrow B^{S_{\text{st}}}$, so it suffices to show:

$$A^{S_{\text{st}}} \vee B^{S_{\text{st}}} \leftrightarrow (A \vee B)^{S_{\text{st}}}$$
$$\forall^{\text{st}} x\, \exists^{\text{st}} y\, \alpha(x,y) \vee \forall^{\text{st}} u\, \exists^{\text{st}} v\, \beta(u,v) \leftrightarrow \forall^{\text{st}} x, u\, \exists^{\text{st}} y, v\, (\alpha(x,y) \vee \beta(u,v))$$

which easy to check in E-$\text{PA}^{\omega *}_{\text{st}}$.

$\boxed{\neg A}$

We need to show:

$$\neg A^{S_{\text{st}}} \leftrightarrow (\neg A)^{S_{\text{st}}}$$
$$\neg \forall^{\text{st}} x\, \exists^{\text{st}} y\, \alpha(x,y) \leftrightarrow \forall^{\text{st}} Y\, \exists^{\text{st}} x'\, \exists x \in x' \;\; \neg\alpha(x,y)$$

Consider the following:

(1) $\neg \forall^{\text{st}} x\, \exists^{\text{st}} y\, \alpha(x,y)$

(2) $\neg \exists^{\text{st}} Y'\, \forall^{\text{st}} x\, \exists y \in Y'x \;\; \alpha(x,y)$

(3) $\forall^{\text{st}} Y'\, \exists^{\text{st}} x\, \forall y \in Y'x \;\; \neg\alpha(x,y)$

(4) $\forall^{\text{st}} Y\, \exists^{\text{st}} x'\, \exists^{\text{st}} x \in x' \;\; \neg\alpha(x, Yx)$

It is enough to show $(i) \leftrightarrow (i+1)$ for all $i \in \{1,2,3\}$. The implication $(1) \to (2)$ is a consequence of $\text{HAC}^{\omega}_{\text{int}}$. The other side is easy to check in E-$\text{PA}^{\omega *}_{\text{st}}$, using Lemma 3.1.24. The equivalence $(2) \leftrightarrow (3)$ follows from the rules for negation. For $(3) \to (4)$, take $Y' := \lambda x\,.\,\{Yx\}$ and $x' := \{x\}$; the result uses Lemma 3.1.7.1. For $(4) \to (3)$, take $Y := \lambda x\,.\, \bigcup_{v \in Y'x} v$; the result follows by Lemmas 3.1.24, 3.1.10.3, and the monotonicity of $\alpha(x,y)$ on y.

$\boxed{\forall z\, A(z)}$

We need to show:

$$\forall z\, A(z)^{S_{\text{st}}} \leftrightarrow (\forall z\, A(z))^{S_{\text{st}}}$$
$$\forall z\, \forall^{\text{st}} x\, \exists^{\text{st}} y\, \alpha(x,y,z) \leftrightarrow \forall^{\text{st}} x\, \exists^{\text{st}} y\, \forall z\, \alpha(x,y,z)$$

From right to left this is straightforward in E-$\text{PA}^{\omega *}_{\text{st}}$. From left to right we use the realization principle R^{ω}, taking into account that $\alpha(x,y,z)$ is monotone on y. This is possible because R^{ω} is the contrapositive of I^{ω}.

□

From the characterization theorems for the S_{st}-interpretation presented here and Van den Berg et al.'s interpretation for classical logic presented in [2] (here written as S_{st}(Berg)) we get:

Theorem 3.4.7. For all formulas A of the language of E-$\text{PA}^{\omega *}_{\text{st}}$:

$$\text{E-PA}^{\omega *}_{\text{st}} + \text{I}^{\omega} + \text{HAC}^{\omega}_{\text{int}} \vdash A^{S_{\text{st}}} \leftrightarrow A^{S_{\text{st}}(\text{Berg})}$$

3.5 Countable saturation

In this section, we briefly study the CSAT^ω principle, which is common among nonstandard arguments.

Definition 3.5.1 (CSAT^ω)**.** The countable saturation principle, CSAT^ω, is the union for all types ρ of:

$$\mathsf{CSAT}^\rho : \forall^{\mathsf{st}} n^0\ \exists z^\rho\, A(n,z) \to \exists Z^{0\to\rho}\ \forall^{\mathsf{st}} n^0\, A(n, Zn)$$

for all formulas A of the language of $\mathsf{E\text{-}HA}^{\omega*}$.

We will see that CSAT^ω is interpreted by the H_{st}-interpretation, or in other words, that it is a characteristic principle.

The story is different when we're in a classical setting. Here adding even CSAT^0 to $\mathsf{E\text{-}PA}^{\omega*}_{\mathsf{st}} + \mathsf{OS}^0$ already gives us the strength of full second-order Peano arithmetic. In fact, as is seen by Van den Berg et al. in [3], $\mathsf{E\text{-}PA}^{\omega*}_{\mathsf{st}} + \mathsf{I}^\omega + \mathsf{HAC}^\omega_{\mathsf{int}} + \mathsf{CSAT}^\omega$ has exactly the strength of second-order arithmetic.

3.5.1 Countable saturation in an intuitionistic setting

We will see that it is possible to reduce the H_{st}-interpretation of CSAT^ω to the finite axiom of choice, $\mathsf{AC}^\omega_{\mathsf{fin}}$. Hence a first goal is to show that $\mathsf{E\text{-}HA}^{\omega*}$ proves $\mathsf{AC}^\omega_{\mathsf{fin}}$.

Definition 3.5.2 ($\mathsf{AC}^\omega_{\mathsf{fin}}$)**.** The finite axiom of choice, $\mathsf{AC}^\omega_{\mathsf{fin}}$, is the union for all types ρ of:

$$\mathsf{AC}^\rho_{\mathsf{fin}} : \forall s^{0^*}\ (\forall n \in s\ \ \exists x^\rho\, \varphi(n,x) \to \exists X^{0\to\rho}\ \forall n \in s\ \ \varphi(n, Xn))$$

for all formulas φ of the language of $\mathsf{E\text{-}HA}^{\omega*}$.

Lemma 3.5.3. $\mathsf{E\text{-}HA}^{\omega*} \vdash \mathsf{AC}^\omega_{\mathsf{fin}}$.

Proof. By sequence induction on s^{0^*}. If $s = \{\}$, then the consequent of $\mathsf{AC}^\omega_{\mathsf{fin}}$ is trivially true, since $\forall n\, \neg(n \in \{\})$.

For a sequence $s' = \mathsf{prep}\, ks$ we know by induction hypothesis:

$$\forall n \in s\ \ \exists x\, \varphi(n,x) \to \exists X\ \forall n \in s\ \ \varphi(n, Xn) \tag{3.5.1}$$

Assume:

$$\forall n \in \mathsf{prep}\, ks\ \ \exists x\, \varphi(n,x) \tag{3.5.2}$$

From (3.5.2), take $n := k$ and x_k such that

$$\varphi(k, x_k) \tag{3.5.3}$$

Since (3.5.2) is stronger than the antecedent of (3.5.1), we also get its consequent:

$$\exists X\ \forall n \in s\ \ \varphi(n, Xn) \tag{3.5.4}$$

Take X_0 such that

$$\forall n \in s\ \ \varphi(x, X_0 n) \tag{3.5.5}$$

as in (3.5.4). We are now able to build:

$$X := \lambda\, n\,.\, \begin{cases} x_k & \text{if } n =_0 k \\ X_0 n & \text{otherwise} \end{cases}$$

by taking into account the decidability of type 0 equality. It remains to show:

$$\forall\, n \in \mathsf{prep}\, ks \;\; \varphi(n, Xn)$$

If $n = k$, this follows from (3.5.3). If $n \in s$, it follows from (3.5.5) instead. There are no more cases to consider. □

Theorem 3.5.4. $\text{E-HA}^{\omega *}_{\mathsf{st}} + \mathsf{I}^{\omega} + \mathsf{NCR}^{\omega} + \mathsf{HAC}^{\omega} + \mathsf{HGMP}^{\omega}_{\mathsf{st}} + \mathsf{HIP}^{\omega}_{\forall^{\mathsf{st}}} \vdash \mathsf{CSAT}^{\omega}$.

Proof. By Theorem 3.3.13, it suffices to extend the proof of the soundness theorem for the H_{st}-interpretation (Theorem 3.3.9) for the CSAT^{ω} principle.

Let $A^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, x\; \forall^{\mathsf{st}}\, y\, \alpha(x, y)$ where x, y should be regarded as tuples, even though we don't write $\boldsymbol{x}$ or $\boldsymbol{y}$ for simplicity.

$$\begin{aligned}(\forall^{\mathsf{st}}\, n\; \exists\, z\, A(n, z) \to \exists\, W\; \forall^{\mathsf{st}}\, m\, A(m, Wm))^{H_{\mathsf{st}}} \equiv \exists^{\mathsf{st}}\, U, N, Y'\; \forall^{\mathsf{st}}\, X, m', v'\; (&\\ \forall\, n \in NXm'v' \;\; \forall\, y' \in Y'Xm'v' \;\; \exists\, z\; \forall\, y \in y' \;\; \alpha(Xn, y, n, z)&\\ \to&\\ \exists\, W\; \forall\, m \in m' \;\; \forall\, v \in v' \;\; \alpha(UXm, v, m, Wm))&\end{aligned}$$

Take $t_U := \lambda\, X\,.\, X$, $t_N := \lambda\, X, m', v'\,.\, m'$ and $t_{Y'} := \lambda\, X, m', v'\,.\, \{v'\}$. We need to show in $\text{E-HA}^{\omega *}$:

$$\begin{aligned}\forall\, X, m', v'\; (\forall\, n \in m' \;\; \forall\, y' \in \{v'\} \;\; \exists\, z\; \forall\, y \in y' \;\; \alpha(Xn, y, n, z)&\\ \to&\\ \exists\, W\; \forall\, m \in m' \;\; \forall\, v \in v' \;\; \alpha(Xm, v, m, Wm))&\end{aligned}$$

or simply:

$$\begin{aligned}\forall\, X, m', v'\; (\forall\, n \in m' \;\; \exists\, z\; \forall\, y \in v' \;\; \alpha(Xn, y, n, z)&\\ \to&\\ \exists\, W\; \forall\, m \in m' \;\; \forall\, v \in v' \;\; \alpha(Xm, v, m, Wm))&\end{aligned}$$

which is an instance of $\mathsf{AC}^{\omega}_{\mathsf{fin}}$. So, by Lemma 3.5.3, we are done. □

3.5.2 Countable saturation in a classical setting

We now want to show that $\text{E-PA}^{\omega *}_{\mathsf{st}} + \mathsf{OS}^0 + \mathsf{CSAT}^0$ interprets full second-order arithmetic.

Definition 3.5.5 (PA_2)**.** Full second-order Peano arithmetic, PA_2, is a two-sorted classical system, that distinguishes between number terms (denoted as lower-case Latin letters) and set terms (denoted as upper-case Latin letters). The language includes the usual logic connectives (with distinguished quantifiers for numbers and for sets), the constants 0 and 1 with their usual meanings, the functions

$+$ (sum of two numbers) and $\cdot$ (product of two numbers), and three predicates: $=$ (equality between numbers), $<$ (inequality between numbers), and $\in$ (membership of a number in a set).

Other than logical axioms for classical logic and basic axioms governing the behaviour of the various terms available, it also includes an induction axiom:

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

and the full comprehension schema:

$$\exists X\ \forall n\,(n \in X \leftrightarrow \varphi(n))$$

where φ is any formula without X as free variable.

The full definition can be found, for example, in Section I.2 of [39] (where PA_2 is named Z_2).

Lemma 3.5.6. $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}} \vdash \forall^{\mathsf{st}} n^0\ \forall m^0\,(\neg\,\mathsf{st}\,m \to n < m)$.

Proof. It suffices to notice that: either $n < m$ and we are done, or $m \le n$, and then m must be standard by Lemma 3.1.23. □

Theorem 3.5.7. $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}} + \mathsf{OS}^0 + \mathsf{CSAT}^0$ interprets PA_2.

Proof. The idea is to see number terms of PA_2 as standard terms of type 0 in $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$ and set terms of PA_2 as terms of type 0^* in $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$. Then all the terms of PA_2 have ready interpretations in $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$.

Classical logic is interpreted, as well as the basic axioms defining the behaviour of the terms of PA_2. For the induction axiom of PA_2 we can use the external induction axiom of $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$. Now it only remains to show that full comprehension is verified. Translating this schema into the language of $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$, we need to show that for every formula A:

$$\exists\, s^{0^*}\ \forall^{\mathsf{st}} n^0\,(n \in s \leftrightarrow A(n))$$

We start by noticing that, with the help of the law of excluded middle, $\mathsf{E}\text{-}\mathsf{PA}^{\omega*}_{\mathsf{st}}$ shows:

$$\forall^{\mathsf{st}} n^0\ \exists\, k^0\,(k = 0 \leftrightarrow A(n)) \tag{3.5.6}$$

Then from CSAT^0 and (3.5.6) we obtain:

$$\exists\, K^{0\to 0}\ \forall^{\mathsf{st}} n^0\,(Kn = 0 \leftrightarrow A(n)) \tag{3.5.7}$$

Take K in the circumstances of (3.5.7). It is easy to show by external induction that:

$$\forall^{\mathsf{st}} m^0\ \exists\, s_m^{0^*}\,(\forall\, n \le m\ \ (n \in s_m \leftrightarrow Kn = 0) \wedge \forall\, l \in s_m\ \ (l \le m)) \tag{3.5.8}$$

It suffices to take:

$$s_0 := \begin{cases} \{0\} & \text{if } K0 =_0 0 \\ \{\} & \text{otherwise} \end{cases} \qquad\qquad s_{(Sm)} := \begin{cases} \mathsf{prep}_0(Sm)s_m & \text{if } K(Sm) =_0 0 \\ s_m & \text{otherwise} \end{cases}$$

We don't really care that $\forall\, l \in s_m\ \ (l \le m)$, but it is easy to show and makes the induction hypotheses strong enough to prove the other part.

From (3.5.8) and OS^0 we get:

$$\exists\, m^0\,(\neg\,\mathsf{st}\, m \wedge \exists\, s^{0^*}\,(\forall\, n \leq m\;\;(n \in s \leftrightarrow Kn = 0) \wedge \forall\, l \in s\;\;(l \leq m)))$$

So there is a nonstandard number m for which:

$$\exists\, s^{0^*}\;\forall\, n \leq m\;\;(n \in s \leftrightarrow Kn = 0) \tag{3.5.9}$$

Since by Lemma 3.5.6 all standard numbers are less than any nonstandard number, from (3.5.9) in particular:

$$\exists\, s^{0^*}\;\forall^{\mathsf{st}}\, n\,(n \in s \leftrightarrow Kn = 0) \tag{3.5.10}$$

From (3.5.6) and (3.5.10) we are done:

$$\exists\, s^{0^*}\;\forall^{\mathsf{st}}\, n\,(n \in s \leftrightarrow A(n))$$

□

4

A note on Weihrauch reducibility

Contents

In this chapter we study something different from the focus of the previous chapters: Weihrauch reducibility. This is a relation between multi-valued operations first introduced by Klaus Weihrauch in [45, 46], and later modified to its modern definition by Gherardi and Marcone [20]. It has been used in logic to find which theorems can be computationally (or continuously) reduced into which other (see, for example, [5, 6, 7, 20]). To deal with this, we look at theorems as multi-valued operations. In particular, theorems of the form:

$$\forall x\ \exists y\, A_0(x, y)$$

are good candidates to see as multi-valued operations, since to every x it's possible to assign at least a y such that $A_0(x, y)$.

In Section 4.1, we shortly describe the notion of Weihrauch reducibility between multi-valued operations. Then, in Section 4.2, we outline a connection between this new notion and the *dialectica* interpretation.

4.1 Weihrauch reducibility

Definition 4.1.1 (Multi-valued operation)**.** A multi-valued operation is a relation $f : X \rightrightarrows Y$ between spaces X and Y such that each element of X is in relation to at least one element from Y.

Note that a multi-valued operation is generally not a function in the usual sense of the word, but functions can be seen as multi-valued operations.

Definition 4.1.2 (Representation, represented space)**.** A representation δ_X of a set X is a surjective function $\delta_X : \mathbb{N}^{\mathbb{N}} \to X$. If such a function exists, we say that X is a represented space.

It can be useful to know specific representatives of the output of a given multi-valued operation. In other words, to have a function which realizes a given multi-valued operation.

Definition 4.1.3 (Realizer, $\vdash$)**.** Let $f : X \rightrightarrows Y$ be a multi-valued operation on represented spaces (X, δ_X) and (Y, δ_Y). We say that $F : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ is a realizer of f with respect to (δ_X, δ_Y), in symbols $F \vdash_{\delta_X,\delta_Y} f$, if for all $p \in \mathbb{N}^{\mathbb{N}}$ we have:

$$\delta_Y(F(p)) \in f(\delta_X(p))$$

Functions from $\mathbb{N}^{\mathbb{N}}$ to $\mathbb{N}^{\mathbb{N}}$ can be though of as functions from $\mathbb{R}$ to $\mathbb{R}$. We say that such a function is computable if there exists a type-two Turing machine which computes it. A function is continuous if every finite portion of the output is determined by a finite portion of the input (the notion induced by the Cantor topology on $\mathbb{N}^{\mathbb{N}}$). For more details about computable analysis, see [47].

Definition 4.1.4 (Computable and continuous multi-valued operations on represented spaces)**.** A multi-valued operation $f : X \rightrightarrows Y$ on represented spaces (X, δ_X) and (Y, δ_Y) is (δ_X, δ_Y)-computable (respect. (δ_X, δ_X)-continuous) if there is a computable (respect. continuous) realizer of f with respect to (δ_X, δ_Y).

We can now define what is meant by Weihrauch reducibility.

Definition 4.1.5 (Weihrauch reducibility, $\leq_W$)**.** Let $f, g : X \rightrightarrows Y$ be multi-valued operations on represented spaces (X, δ_X) and (Y, δ_Y). We say that f Weihrauch-reduces to g, in symbols $f \leq_W g$, if there exist computable functions $H : \mathbb{N}^\mathbb{N} \times \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ and $K : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that for all realizers G of g, the function $F : \mathbb{N}^\mathbb{N} \to \mathbb{N}^\mathbb{N}$ such that:

$$F(p) = H(p, G(K(p))) \quad \text{for all } p \in \mathbb{N}^\mathbb{N}$$

is a realizer of f.

It is interesting to ask which multi-valued operations Weihrauch-reduce to which others. If we see $\forall\exists$-formulas as multi-valued operations, we can also ask which formulas Weihrauch-reduce to which other. What follows is a technique to do just that, with the help of the *dialectica* interpretation.

4.2 A comment on how to use the *dialectica* interpretation to prove reductions

Recall that type 1 is the same as $0 \to 0$, and hence it's the appropriate type to represent $\mathbb{N}^\mathbb{N}$, assuming we are in the set theoretical model (see section 3.6 of [29] for details on models). Let $A_0(x^1, y^1)$ and $B_0(z^1, w^1)$ be quantifier-free formulas without $\vee$ in the language of WE-HA$^\omega$. Recall from Corollary 2.1.9 that in WE-HA$^\omega$ quantifier-free formulas can be rewritten to loose any $\vee$ they might have, so this is not really a restriction. Recall also from Remark 2.2.2 that these formulas remain unchanged after the *dialectica* interpretation. Consider the formulas $\forall x\, \exists y\, A_0(x, y)$ and $\forall z\, \exists w\, B_0(z, w)$ and the multi-valued operations associated with them:

$$\mathcal{A} : \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}^\mathbb{N} \quad \text{such that} \quad \mathcal{A}(x) = \{y \mid A_0(x, y)\}$$
$$\mathcal{B} : \mathbb{N}^\mathbb{N} \rightrightarrows \mathbb{N}^\mathbb{N} \quad \text{such that} \quad \mathcal{B}(z) = \{w \mid B_0(z, w)\}$$

Now suppose:

$$\forall x\, \exists y\, A_0(x, y) \to \forall z\, \exists w\, B_0(z, w) \tag{4.2.1}$$

Then it would be reasonable to wonder whether $\mathcal{B} \leq_W \mathcal{A}$. After all, this is just a precise way of stating that we are able to "solve" $\mathcal{B}$ after already knowing how to "solve" $\mathcal{A}$. That's what it means to reduce one problem to another. We will see that it is indeed possible to prove a result to this effect, albeit in a roundabout way.

We start by taking the (classically equivalent) contrapositive of (4.2.1):

$$\exists z\, \forall w\, \neg B_0(z, w) \to \exists x\, \forall y\, \neg A_0(x, y) \tag{4.2.2}$$

and assuming that we can prove its *dialectica* interpretation:

$$\exists X^{1\to 1}, W^{1\to 1\to 1}\, \forall z^1, y^1\, (\neg B_0(z, Wzy) \to \neg A_0(Xz, y)) \tag{4.2.3}$$

giving witnesses X, W in the language of WE-HA$^\omega$ to that effect.

To prove (4.2.3) constructively, we can proceed either directly, or via Theorem 2.2.6 (Soundness of the *dialectica* interpretation). The latter option seems better, but then we would need a constructive proof of (4.2.2), which in general we don't have. For now we proceed on the assumption that such a proof exists and that we have access to the witnesses X and W. We want to use these witnesses to prove $\mathcal{B} \leq_{\mathsf{W}} \mathcal{A}$.

As $\mathcal{A}, \mathcal{B} : \mathbb{N}^{\mathbb{N}} \rightrightarrows \mathbb{N}^{\mathbb{N}}$, we take $\delta_{\mathbb{N}^{\mathbb{N}}}$ as the identity representation. Let $\mathbb{A} : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ be a realizer of $\mathcal{A}$. In other words, such that for all $x \in \mathbb{N}^{\mathbb{N}}$ we have:

$$\mathbb{A}(x) \in \mathcal{A}(x)$$

or equivalently

$$A_0(x, \mathbb{A}(x))$$

Consider $\mathbb{B}$ such that for every $z \in \mathbb{N}^{\mathbb{N}}$ we have $\mathbb{B}(z) = t_W(z, \mathbb{A}(t_X(z)))$, where $t_X : \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ is the semantic interpretation of X and $t_W : \mathbb{N}^{\mathbb{N}} \times \mathbb{N}^{\mathbb{N}} \to \mathbb{N}^{\mathbb{N}}$ is the semantic interpretation of W. Both t_X and t_W are computable functions. We claim that $\mathbb{B}$ is a realizer of $\mathcal{B}$, hence showing $\mathcal{B} \leq_{\mathsf{W}} \mathcal{A}$ as desired. To see this, choose arbitrary $z \in \mathbb{N}^{\mathbb{N}}$. From (4.2.3), after having fixed X and W, we can instantiate $z := z$ and $y := \mathbb{A}(Xz)$, thus obtaining:

$$\neg B_0(z, Wz(\mathbb{A}(Xz))) \to \neg A_0(Xz, \mathbb{A}(Xz))$$

Taking the (classical) contrapositive of the previous expression, and from the fact that $\mathbb{A}$ is a realizer of $\mathcal{A}$, we are able to conclude:

$$B_0(z, Wz(\mathbb{A}(Xz)))$$

which is precisely the statement that $\mathbb{B}$ realizes $\mathcal{B}$, seeing as we took z to be arbitrary.

5

Conclusions and future work

Contents

5.1 Results

We saw that it is possible to modify the D_{st}-interpretation of Van den Berg et al. [2] to allow for a different notion of monotonicity, thus obtaining the H_{st}-interpretation. We proved soundness, characterization and term extraction theorems for this new interpretation. As corollary from the characterization theorem, we saw that the D_{st}-interpretation and the H_{st}-interpretation are equivalent in the presence of the principles I^{ω}, NCR^{ω}, HAC^{ω}, $\mathsf{HGMP}^{\omega}_{\text{st}}$ and $\mathsf{HIP}^{\omega}_{\forall^{\text{st}}}$.

We saw that the S_{st}-interpretation of Dinis and Ferreira [10] can be extended to $\text{E-PA}^{\omega*}_{\text{st}}$, and proved a characterization theorem for it. From this theorem, we obtained as corollary that the S_{st}-interpretations of Van den Berg et al. [2], and of Dinis and Ferreira [10] are equivalent in the presence of the principles I^{ω} and $\mathsf{HAC}^{\omega}_{\text{int}}$.

We used the H_{st}-interpretation as a way to investigate the strength of the countable saturation principle CSAT^{ω} in an intuitionistic context. We saw that this principle is weak intuitionistically. We also observed that when CSAT^{ω} is added to $\text{E-PA}^{\omega*}_{\text{st}} + \mathsf{I}^{\omega}$, the ensemble interprets full second-order arithmetic. These results had already been obtain by Van den Berg et al. in [3], albeit using the D_{st}-interpretation and not the H_{st}-interpretation as the main tool.

Finally, we observed that in certain circumstances it is possible to use the *dialectica* interpretation to show that a a $\forall\exists$-theorem Weihrauch-reduces to another one.

5.2 Future work

There are some interesting questions left to answer regarding to the topics of this thesis. For example, is there a negative translation which composed with the H_{st}-interpretation gives the S_{st}-interpretation? The work of Van den Berg et al. in Section 6 of [2] might be useful to answer this question. However, it is not possible to use their negative translation directly, because we need to preserve the monotonicity of the interpretation.

We would also like to know whether the soundness of the D_{st}-interpretation is in fact not intuitionistically verifiable with the clause of the disjunction as it is, or if it is possible to make use of the finiteness of the sequences to prove the soundness theorem in $\text{E-HA}^{\omega*}$.

From a more practical point of view, one should be able to use the functional interpretations described here to unwind nonstandard arguments.

Escardó and Oliva [12] have shown that it is possible to witness the D_{st}-interpretation of the double-negation-shift with the help of Spector's bar recursion [40]. Using this result in a fundamental way, Van den Berg et al. [3] show that $\text{E-PA}^{\omega*}_{\text{st}} + \mathsf{I}^{\omega} + \mathsf{HAC}^{\omega}$ coupled with the countable saturation principle CSAT^{ω} have the strength of full second-order arithmetic. It might be possible to interpret a nonstandard version of the dependent choices principle using bar recursion, and it would be interesting to see if such a result gives rise to new nonstandard proof principles.

Finally, we would like to improve the result on the Weihrauch-reducibility of $\forall\exists$-theorems by way of the *dialectica* interpretation. One direction would be to weaken the conditions of the result; we could also extend it to allow formulas with free variables of types other than $0 \to 0$.

Bibliography

[1] Avigad, J., and Feferman, S. Gödel's functional ("Dialectica") interpretation. In *Handbook of Proof Theory*, S. R. Buss, Ed. Elsevier, New York, 1998, pp. 337–406.

[2] van den Berg, B., Briseid, E., and Safarik, P. A functional interpretation for nonstandard arithmetic. *Annals of Pure and Applied Logic 163* (2012), 1962–1994.

[3] van den Berg, B., Briseid, E., and Safarik, P. The strength of countable saturation. arXiv:1605.02534, 2016.

[4] Boolos, G. S., Burgess, J. P., and Jeffrey, R. C. *Computability and Logic*, 4th ed. Cambridge University Press, New York, 2002.

[5] Brattka, V., and Gherardi, G. Effective choice and boundedness principles in computable analysis. *Bulletin of Symbolic Logic 17*, 1 (2011), 73–117.

[6] Brattka, V., and Gherardi, G. Weihrauch degrees, omniscience principles and weak computability. *Journal of Symbolic Logic 76*, 1 (2011), 143–176.

[7] Brattka, V., Gherardi, G., and Marcone, A. The Bolzano-Weierstrass theorem is the jump of weak König's lemma. *Annals of Pure and Applied Logic 163* (2012), 623–655.

[8] Diller, J. Logical problems of functional interpretations. *Annals of Pure and Applied Logic 114* (2002), 27–42.

[9] Diller, J., and Nahm, W. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik endlicher Typen. *Archiv für mathematische Logik und Grundlagenforschung 16* (1974), 49–66.

[10] Dinis, B., and Ferreira, F. Interpreting weak König's lemma in theories of nonstandard arithmetic. To appear in Mathematical Logic Quarterly, 2016.

[11] Dinis, B., and Gaspar, J. Intuitionistic nonstandard bounded modified realisability and functional interpretation. Manuscript, September 2016.

[12] Escardó, M., and Oliva, P. The Herbrand functional interpretation of the double negation shift. arXiv:1410.4353v2, October 2015.

[13] Feferman, S. Gödel's Dialectica interpretation and its two-way stretch. In *Computational Logic and Proof Theory*, G. Gottlob, A. Leitsch, and D. Mundici, Eds., vol. 713 of *Lecture Notes in Computer Science*. Springer-Verlag, Berlin, 1993, pp. 23–40.

[14] FERREIRA, F. Injecting uniformities into Peano arithmetic. *Annals of Pure and Applied Logic 157* (2009), 122–129.

[15] FERREIRA, F., AND GASPAR, J. Nonstandardness and the bounded functional interpretation. *Annals of Pure and Applied Logic 166* (2015), 701–712.

[16] FERREIRA, F., AND OLIVA, P. Bounded functional interpretation. *Annals of Pure and Applied Logic 135* (2005), 73–112.

[17] GASPAR, J. Em torno das interpretações funcionais da aritmética. Master's thesis, Universidade de Lisboa, 2007.

[18] GENTZEN, G. Über das Verhältnis zwischen intuitionisctischer und klassischer Arithmetik. *Mathematische Annalen* (1933). Galley proof.

[19] GENTZEN, G. Die Widerspruchsfreiheit der reinen Zahlentheorie. *Mathematische Annalen 112* (1936), 493–565.

[20] GHERARDI, G., AND MARCONE, A. How incomputable is the separable Hahn-Banach theorem? *Notre Dame Journal of Formal Logic 50* (2009), 393–425.

[21] GÖDEL, K. Zur intuitionistischen Arithmetik und Zahlentheorie. *Ergebnisse eines Mathematischen Kolloquiums 4* (1933), 34–38.

[22] GÖDEL, K. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *dialectica 12*, 3-4 (1958), 280–287.

[23] GÖDEL, K. On an extension of finitary mathematics which has not yet been used. In *Kurt Gödel Collected Works*, S. Feferman, J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, and J. van Heijenoort, Eds., vol. II. Oxford University Press, New York, 1972, pp. 271–280.

[24] HOWARD, W. A. Hereditarily majorizable functionals of finite type. In *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, A. S. Troelstra, Ed., vol. 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973, pp. 454–461.

[25] JØRGENSEN, K. F. Finite type arithmetic. Computable existence analysed by modified realisability and functional interpretation. Master's thesis, University of Roskilde, Roskilde, 2001.

[26] KAHLE, R. Gentzen's consistency proof in context. In *Gentzen's Centenary: The Quest for Consistency*, R. Kahle and M. Rathjen, Eds. Springer, 2015, pp. 3–24.

[27] KLEENE, S. C. On the interpretation of intuitionistic number theory. *Journal of Symbolic Logic 10*, 4 (1946), 109–124.

[28] KOHLENBACH, U. Analysing proofs in analysis. In *Logic: From Foundations to Applications*, W. Hodges, M. Hyland, C. Steinhorn, and J. Truss, Eds. Oxford University Press, Oxford, 1996, pp. 225–260.

[29] KOHLENBACH, U. *Applied Proof Theory: Proof Interpretations and their Use in Mathematics*. Springer, Berlin, 2008.

[30] KOHLENBACH, U. Effective uniform bounds from proofs in abstract functional analysis. In *New Computational Paradigms: Changing Conceptions of What is Computable*, S. B. Cooper, B. Löwe, and A. Sorbi, Eds. Springer, New York, 2008, pp. 223–258.

[31] KOHLENBACH, U. Gödel's functional interpretation and its use in current mathematics. In *Kurt Gödel and the Foundations of Mathematics: Horizons of Truth*, M. Baaz, C. H. Papadimitriou, H. W. Putnam, D. S. Scott, and C. L. Harper, Jr, Eds. Cambridge University Press, 2011, pp. 361–398.

[32] KOHLENBACH, U. Recent progress in proof mining in nonlinear analysis. To appear in a forthcoming book, July 2016.

[33] KREISEL, G. On weak completeness of intuitionistic predicate logic. *Journal of Symbolic Logic 27*, 2 (1962), 139–158.

[34] KRIVINE, J.-L. Opérateurs de mise en mémoire et traduction de Gödel. *Archive for Mathematical Logic 30*, 4 (1990), 241–267.

[35] NELSON, E. Internal set theory: A new approach to nonstandard analysis. *Bulletin of the American Mathematical Society 83*, 6 (1977), 1165–1198.

[36] NELSON, E. The syntax of nonstandard analysis. *Annals of Pure and Applied Logic 38*, 2 (1988), 123–134.

[37] ROBINSON, A. Non-standard analysis. *Proceedings of the Koninklijke Nederlandse Akademie Van Wetenschappen* (1961).

[38] SHOENFIELD, J. R. *Mathematical Logic*. Addison-Wesley, Reading, MA, 1967.

[39] SIMPSON, S. G. *Subsystems of Second Order Arithmetic*. Cambridge University Press, Cambridge, 2009.

[40] SPECTOR, C. Provably recursive functionals of analysis: a consistency proof of analysis by an extension of principles formulated in current intuitionistic mathematics. In *Recursive Function Theory: Proceedings Symposia in Pure Mathematics* (Providence, Rhode Island, 1962), J. C. E. Dekker, Ed., vol. V, American Mathematical Society, pp. 1–27.

[41] STREICHER, T., AND KOHLENBACH, U. Shoenfield is Gödel after Krivine. *Mathematical Logic Quarterly 53*, 2 (2007), 176–179.

[42] SØRENSEN, M. H. B., AND URZYCZYN, P. *Lectures on the Curry-Howard Isomorphism*. Elsevier, Amsterdam, 2006.

[43] TROELSTRA, A. S. *Metamathematical Investigation of Intuitionistic Arithmetic and Analysis*, vol. 344 of *Lecture Notes in Mathematics*. Springer, Berlin, 1973.

[44] TROELSTRA, A. S. Introductory note to [22] and [23]. In *Kurt Gödel Collected Works*, S. Feferman, J. W. Dawson, Jr., S. C. Kleene, G. H. Moore, R. M. Solovay, and J. van Heijenoort, Eds., vol. II. Oxford University Press, New York, 1990, pp. 214–241.

[45] WEIHRAUCH, K. The degrees of discontinuity of some translators between representations of the real numbers. Technical report TR-92-050, International Computer Science Institute, July 1992.

[46] WEIHRAUCH, K. The TTE-interpretation of three hierarchies of omniscience principles. Informatik Berichte 130, FernUniversität Hagen, Hagen, September 1992.

[47] WEIHRAUCH, K. *Computable analysis: an introduction*. Springer, Berlin, 2000.

A

Definition of the projection term

In Section 3.1.1, we mentioned the possibility to define in E-HA$^{\omega*}$ a projection term $(\cdot)_{\cdot,\rho}$ of type $\rho^* \to 0 \to \rho$ such that:

$$(\{\})_i = \mathcal{O}^\rho$$
$$(\mathsf{prep}\, xs)_0 = x$$
$$(\mathsf{prep}\, xs)_{(Si)} = (s)_i$$

We will now make its definition explicit. Recall the axioms for $\boldsymbol{R}$ and $\boldsymbol{L}$:

$(\boldsymbol{R})_{\boldsymbol{\rho}}$: Let $\boldsymbol{\rho} = \rho_1, \dots, \rho_k$ be any tuple of types. Let x^0, $\boldsymbol{y} = y_1, \dots, y_k$ with each y_i of type ρ_i, and $\boldsymbol{z} = z_1, \dots, z_k$ with each z_i of type $0 \to \rho_1 \to \dots \to \rho_k \to \rho_i$ be variables. The axioms are:

$$\begin{aligned} &(R_i)_{\boldsymbol{\rho}} 0\boldsymbol{yz} =_{\rho_i} y_i \\ &(R_i)_{\boldsymbol{\rho}} (Sx)\boldsymbol{yz} =_{\rho_i} z_i x(\boldsymbol{R}_{\boldsymbol{\rho}} x\boldsymbol{yz}) \end{aligned} \qquad \text{for } i \in \{1, \dots, k\}$$

$(\boldsymbol{L})_{\sigma,\boldsymbol{\rho}}$: Let $\boldsymbol{\rho} = \rho_1, \dots, \rho_k$ be any tuple of types. Let x^σ, s^{σ^*}, $\boldsymbol{y} = y_1, \dots, y_k$ with each y_i of type ρ_i, and $\boldsymbol{z} := z_1, \dots, z_k$ with each z_i of type $\sigma \to \rho_1 \to \dots \to \rho_k \to \rho_i$ be variables. The axioms are:

$$\begin{aligned} &(L_i)_{\sigma,\boldsymbol{\rho}} \{\}_\sigma \boldsymbol{yz} =_{\rho_i} y_i \\ &(L_i)_{\sigma,\boldsymbol{\rho}} (\mathsf{prep}_\sigma\, xs)\boldsymbol{yz} =_{\rho_i} z_i x(\boldsymbol{L}_{\sigma,\boldsymbol{\rho}} s\boldsymbol{yz}) \end{aligned} \qquad \text{for } i \in \{1, \dots, k\}$$

We define three auxiliary terms:

- First element of a sequence: first_ρ of type $\rho^* \to \rho$, defined as:

$$\mathsf{first}_\rho := \lambda\, s^{\rho^*} .\, L_{\rho,\rho} s \mathcal{O}^\rho (\lambda\, x^\rho, y^\rho \,.\, x)$$

  Using the axioms of L, we obtain:

$$\mathsf{first}\{\} = \mathcal{O}$$
$$\mathsf{first}(\mathsf{prep}\, xs) = x$$

- Identity and all except the first element of a sequence: $\mathsf{id}_\rho, \mathsf{rest}_\rho$ both of type $\rho^* \to \rho^*$, defined making use of simultaneous list recursion as:

$$\mathsf{id}_\rho := \lambda\, s^{\rho^*} .\, (L_1)_{\rho,\rho^*,\rho^*} s \{\}_\rho \mathcal{O}^{\rho^*} (\lambda\, x^\rho, u^{\rho^*}, v^{\rho^*} .\, \mathsf{prep}_\rho\, xu)(\lambda\, x^\rho, u^{\rho^*}, v^{\rho^*} .\, u)$$
$$\mathsf{rest}_\rho := \lambda\, s^{\rho^*} .\, (L_2)_{\rho,\rho^*,\rho^*} s \{\}_\rho \mathcal{O}^{\rho^*} (\lambda\, x^\rho, u^{\rho^*}, v^{\rho^*} .\, \mathsf{prep}_\rho\, xu)(\lambda\, x^\rho, u^{\rho^*}, v^{\rho^*} .\, u)$$

  Using the axioms of L_1, L_2, we obtain:

$$\mathsf{id}\{\} = \{\} \qquad \mathsf{rest}\{\} = \mathcal{O}$$
$$\mathsf{id}(\mathsf{prep}\, xs) = \mathsf{prep}\, xs \qquad \mathsf{rest}(\mathsf{prep}\, xs) = s$$

- End subsequence whose first element is the i-th element of the original sequence: sub_ρ of type $\rho^* \to 0 \to \rho^*$, defined as:

$$\mathsf{sub}_\rho := \lambda\, s^{\rho^*}, i^0 .\, R_{\rho^*} is (\lambda\, j^0, w^{\rho^*} .\, \mathsf{rest}_\rho\, w)$$

Using the axioms of R and the above, we obtain:

$$\mathsf{sub}\, s0 = s$$
$$\mathsf{sub}\, s(Si) = \mathsf{rest}(\mathsf{sub}\, si)$$

For example, informally writing a list as $\{s_0, s_1, s_2, s_3, s_4, s_5\}$:

$$\mathsf{sub}\{s_0, s_1, s_2, s_3, s_4, s_5\}(SSS0) = \mathsf{rest}(\mathsf{rest}(\mathsf{rest}\{s_0, s_1, s_2, s_3, s_4, s_5\})) = \{s_3, s_4, s_5\}$$

We are finally able to define the projection: $(\cdot)_{\cdot,\rho}$ of type $\rho^* \to 0 \to \rho$ as:

$$(\cdot)_{\cdot,\rho} := \lambda\, s^{\rho^*}, i^0 \,.\, L_{\rho,\rho} s \mathcal{O}^\rho (\lambda\, x^\rho, z^\rho \,.\, \mathsf{first}_\rho(\mathsf{sub}_\rho\, si))$$

We write $(s)_i$ instead of $(\cdot).si$. Using the axioms for L and the above, we obtain:

$$(\{\})_i = \mathcal{O}$$
$$(\mathsf{prep}\, xs)_0 = \mathsf{first}(\mathsf{sub}(\mathsf{prep}\, xs)0) = \mathsf{first}(\mathsf{prep}\, xs) = x$$
$$(\mathsf{prep}\, xs)_{(Si)} = \mathsf{first}(\mathsf{sub}(\mathsf{prep}\, xs)(Si)) = (s)_i$$

as desired.