Università degli Studi di Pavia

DIPARTIMENTO DI MATEMATICA "FELICE CASORATI"
Corso di Laurea Magistrale in Matematica

Nonstandard functional interpretations and categorical models

(Interpretazioni funzionali nonstandard e modelli categoriali)

Relatore:
Prof. Ludovico Pernazza

Correlatore:
Dr. Benno van den Berg

Tesi di laurea di:
Amar Hadzihasanovic

Anno Accademico 2012-2013

Sommario in italiano

L'argomento di questa tesi si colloca alla confluenza di due percorsi diversi della logica matematica. Da un lato, vi sono l'aritmetica e l'analisi nonstandard: un prodotto della teoria classica dei modelli, rimasto per lo più confinato in ambito classico, anche nel momento in cui, in seguito alla *teoria interna degli insiemi* di Nelson, si è riconosciuta la possibilità di un approccio *sintattico.* Dall'altro lato, vi è la teoria delle interpretazioni funzionali, una branca della teoria delle dimostrazioni, originata dall'interpretazione *Dialectica* di Gödel; in particolare, la sua recente riscoperta da parte del programma *proof mining* - l'estrazione di contenuto computazionale da dimostrazioni formali.

Dove questi si incontrano, è in una generale propensione nei confronti della *costruttivizzazione* della matematica. I primi modelli di aritmetica nonstandard utilizzavano, in modo essenziale, risultati non costruttivi, come l'esistenza di ultrafiltri non principali di insiemi; e anche nell'approccio sintattico, si è realizzato in breve che diversi utili principi nonstandard conducessero a istanze della legge del terzo escluso. È sembrato, per diversi anni, che l'analisi nonstandard fosse priva di alcun interesse per gli analisti costruttivi.

Ma non tutti si sono dati per vinti. Dopo un successo parziale di Palmgren, nel 1995, Moerdijk ha descritto il primo modello costruttivo dell'aritmetica nonstandard con un principio di *transfer* completo: un topos di fasci su una categoria di filtri e germi di funzioni "continue". Più di recente, nel 2012, van den Berg, Briseid e Safarik hanno definito un'interpretazione funzionale, *Dialectica nonstandard*, in grado di eliminare istanze di principi nonstandard da dimostrazioni dell'aritmetica intuizionistica in tutti i tipi finiti; dimostrando sintatticamente, nel contempo, la conservatività di tali principi rispetto al sistema di base.

I principi caratteristici dell'interpretazione Dialectica nonstandard hanno una peculiarità: sono *herbrandizzati.* In breve, laddove le tradizionali interpretazioni funzionali avrebbero prodotto un *singolo* realizzatore di un quantificatore esistenziale, questi principi producono una sequenza finita di realizzatori potenziali, fra i quali almeno uno è un realizzatore effettivo. Questa proprietà, che ricorda le disgiunzioni alla Herbrand nella logica classica, rovina il significato computazionale della disgiunzione intuizionistica, ma sembra inevitabile nell'interpretazione dell'aritmetica nonstandard.

Il Capitolo 1 della tesi è dedicato alla disamina di alcuni principi e regole dell'aritmetica nonstandard, con particolare riferimento alla loro accettabilità da un punto di vista costruttivo. Inizialmente, definiamo il sistema dell'aritmetica di Heyting in tutti i tipi finiti, e lo estendiamo con un predicato $\mathrm{st}_\sigma(x)$, "x è standard", per ogni tipo σ.

In seguito, discutiamo i fondamentali principi di *overspill* e *underspill*, e introduciamo una loro generalizzazione ai tipi finiti: overspill e underspill per sequenze. Di questi nuovi principi, mappiamo le relazioni con altri principi noti dell'analisi nonstandard, o rifiutati dall'analisi costruttiva alla Bishop. Passiamo poi in rassegna i principi e le regole di transfer, per dedicarci infine ai cosiddetti principi di saturazione, e ad alcuni principi (non classici) di *uniformità*.

Nel Capitolo 2, dopo una breve introduzione alle interpretazioni funzionali Dialectica e Diller-Nahm, apriamo una parentesi su un'idea ricorrente nella teoria delle dimostrazioni, risalente, in prima istanza, a un articolo di Lifschitz del 1985: avere due specie di quantificatori, una dotata di significato computazionale, e l'altra priva. Lifschitz propose l'introduzione di un predicato $\mathrm{K}(x)$, "x è calcolabile", per determinare la distinzione, con caratteristiche analoghe al nostro $\mathrm{st}(x)$.

Seguendo questa idea, presentiamo un'estensione di Diller-Nahm con due tipi di quantificatori, che chiamiamo *Diller-Nahm uniforme*; ne dimostriamo la validità, ne diamo una caratterizzazione, e stabiliamo alcune prima proprietà. Il sistema caratteristico di Diller-Nahm uniforme sembra adeguato all'aritmetica proposta da Lifschitz; è, nel contempo, una versione "de-herbrandizzata" del sistema caratteristico di Dialectica nonstandard. A quest'ultima è dedicata la parte finale del capitolo.

Nel Capitolo 3, dopo una breve introduzione alla logica categoriale, descriviamo la costruzione che porta al topos di Moerdijk, $\mathcal{N}$, e introduciamo una sua variante, che chiamiamo $\mathcal{U}$. Dimostriamo poi che, date alcune ragionevoli assunzioni riguardo alla metateoria, la logica del primo ordine in $\mathcal{U}$, rispettivamente in $\mathcal{N}$, riflette fedelmente il sistema caratteristico di Diller-Nahm uniforme, rispettivamente di Dialectica nonstandard. La caratterizzazione del rapporto fra $\mathcal{U}$ and $\mathcal{N}$ fornisce, dunque, un corrispettivo categoriale della herbrandizzazione.

Infine, passiamo in rassegna una classe completamente diversa di topoi elementari, *non* di Grothendieck, che sono stati presi in considerazione come modelli di interpretazioni funzionali, incluse Diller-Nahm e Dialectica nonstandard; la cui relazione con i topoi $\mathcal{U}$ e $\mathcal{N}$ ancora non è chiara.

Le conclusioni sono dedicate ai risultati concettuali che riteniamo più rilevanti, e alle domande che lasciamo aperte.

Introduction

This thesis is an account of the research I have done during one semester that I spent in Utrecht, as an exchange student, under the supervision of Benno van den Berg. Its focus stands at a confluence of two quite different paths in mathematical logic.

On one end, there is nonstandard arithmetic, and analysis: a subject that has been an upshot of classical model theory, and even after it was recognised that it was amenable to a syntactic treatment, as in Nelson's *internal set theory*, it mostly remained within the boundaries of classical set theory. On the other end, there is the markedly proof-theoretic topic of functional interpretations, stemming from Gödel's *Dialectica interpretation*; and, in particular, its recent revival through the programme of proof mining - the extraction of computational content from formalised proofs.

Where these ends meet, is in a general inclination towards the *constructivisation* of mathematics. The original models of nonstandard arithmetic were built from nonconstructive objects, such as nonprincipal ultrafilters of sets, and even in the syntactic approach, it was soon realised that many useful principles led to instances of the excluded middle - the nemesis of intuitionistic mathematics. But did nonstandard analysis really have *nothing* to offer to constructive analysts?

Not everyone was convinced, including, notably, Per Martin-Löf, who pushed the question in the early 1990s: first, Erik Palmgren succeeded in building a model with a restricted, yet useful transfer principle; then, in 1995, Ieke Moerdijk described the first constructive model of nonstandard arithmetic with a *full* transfer principle - a topos of sheaves over a category of filters. Later, by working in this topos, Palmgren provided simplified, nonstandard proofs of several theorems of constructive analysis, and so demonstrated the usefulness of this model.

But if nonstandard proofs do provide some constructive information, we might as well try to extract it in an automated fashion. In 2012, van den Berg, Briseid and Safarik succeeded in defining a functional interpretation, *nonstandard Dialectica*, which could eliminate nonstandard principles from proofs of intuitionistic arithmetic in all finite types, enriched with a predicate $\mathrm{st}_\sigma(x)$, "x is standard", for all types σ; also yielding a proof of conservativity of these principles over the base system.

Now, some of these were known to hold in Moerdijk's topos - including a form of Nelson's idealisation axiom, an underspill principle, and the undecidability of the standardness predicate. When I arrived in Utrecht, I was given the task of investigating how deep this connection would go.

And a deep connection it is: with the exception of one principle, which requires an assumption about the metatheory - in my opinion, not an unreasonable one -, *all* the characteristic principles of nonstandard Dialectica are true in the topos model, for free. Chapter 3 is devoted to showing this.

During this investigation, I also chanced upon two new principles, *sequence overspill* and *sequence underspill*. These appear to be more natural equivalents of principles that have been taken into consideration, earlier, in the context of proof-theoretic nonstandard arithmetic. In Chapter 1, I mapped their relation to other familiar principles from nonstandard and constructive analysis.

That is not all. Several characteristic principles of nonstandard Dialectica have a peculiarity: they are *herbrandised*. This is explained in more detail in Chapter 1; in short, where "traditional" functional interpretations would produce a *single* witness of an existential statement, these principles produce a *finite sequence* of *potential* witnesses, of which at least one is an actual witness. This property destroys the computational meaning of intuitionistic disjunction, yet seems unavoidable in the interpretation of nonstandard arithmetic.

The categorical analysis of nonstandard Dialectica supplied a very convenient way of "de-herbrandising", through a simple change in the Grothendieck topology, down from *finite* covers to *singleton* covers. Full transfer is lost - in the new topos, disjunction is stronger than in the metatheory - as well as the link to nonstandard arithmetic; but the de-herbrandised principles induce a new functional interpretation, which we call *uniform Diller-Nahm*, and is the main focus of Chapter 2.

Uniform Diller-Nahm has some striking similarities to light Dialectica, a variant of Dialectica with two different kinds of quantifiers - computational, and non computational - introduced in 2005 by Mircea-Dan Hernest, for the purpose of more efficient program extraction. Yet, irrespective of its technical value, the characteristic proof system of uniform Diller-Nahm might have a dignity of its own.

In 1985, Vladimir Lifschitz proposed a simple extension of Heyting arithmetic, where a distinction could be made between calculable, and non calculable natural numbers; a synthesis of classical and intuitionistic arithmetic. Under the interpretation of the predicate $\mathrm{st}(x)$ as "x is calculable", the proof system of uniform Diller-Nahm seems to be well-suited for Lifschitz's intended calculus. How useful that could turn out to be, I do not know; yet there is something intriguing about the "standardness" of nonstandard arithmetic being, conceivably, "herbrandised calculability".

As you may have noticed, in no way does the order in which these topics are presented in the thesis reflect the actual order in which I got to them; I had to opt for a reasonable progression of arguments, and the result is, inevitably, a fabrication of history. This introduction is a way of restoring some historical truth. It is also the only place where I refer to myself, before disappearing behind an impersonal "we"; it is hard to establish, for most statements and results, how much authorship one can personally claim, so this seemed a necessity to me.

So Chapter 1 and Chapter 2 are all about proof theory: the first is a discussion of principles of nonstandard arithmetic, and the second is about functional interpretations. All the category theory is stored in Chapter 3. There is a short summary of the contents of each section, at the beginning of each chapter; no need to be too detailed here.

Perhaps my main stylistic choice is that I introduce new notions, or technical results, only as soon as they are actually used: hence, no introductory sections, where all the lemmata and definitions are introduced, long before they are actually employed. This is particularly evident in Chapter 3, where this decision was also made in compliance with a pluralist view of metamathematics; so, for instance, no extra structure is imposed on the category of sets and functions, until it is really needed.

I tried to be formal enough in proofs, and informal in their discussion, and I did not shy away from analogy. Overall, I hope that I avoided, at least, an excess of unpleasantness in style.

Original contributions of this thesis

To the best of my knowledge, the following parts of this thesis are my original contributions.

> **Chapter 1.** Sections 1.3, 1.4: the definition of the sequence overspill and sequence underspill principles, and the equivalences $\mathsf{OS}^* \leftrightarrow \mathsf{I}$ and $\mathsf{US}^* \leftrightarrow \mathsf{HGMP}^{\mathrm{st}} + \mathsf{R}$; direct proofs of all their consequences, including EUS and EOS (but $\mathsf{I} \to \mathsf{LLPO}^{\mathrm{st}}$ was already known). Section 1.6: the definition of the nonstandard uniformity principle, and the discussion of its consequences.
>
> **Chapter 2.** Sections 2.2, 2.3 - all of them: the definition of the uniform Diller-Nahm interpretation; the soundness, characterisation and program extraction theorems; the discussion of their consequences, and the connection to Lifschitz's calculability.
>
> **Chapter 3.** Section 3.2: the definition of the topos $\mathcal{U}$, and the characterisation of the inclusion $\mathcal{N} \rightarrowtail \mathcal{U}$. Sections 3.3, 3.4: the adaptation of theorems about $\mathcal{N}$ to $\mathcal{U}$; the connection of $\mathcal{N}$ and $\mathcal{U}$ with nonstandard Dialectica, and uniform Diller-Nahm, respectively; the discussion of the validity of $\mathsf{FAN}^{\mathrm{st}}$ in $\mathcal{N}$. Section 3.5: the proof that $\mathcal{D}st$, $\mathcal{DN}_m$, $\mathcal{H}er$, and $\mathcal{M}od$ form a pullback square of geometric morphisms of triposes.

Regarding Section 3.4, I discovered only at a late stage that a characterisation of first order logic in $\mathcal{N}$ had already been given by Butz; nevertheless, the proofs that US^*, NCR, and $\mathsf{HAC}^{\mathrm{st}}$ hold in $\mathcal{N}$, which only utilise sheaf semantics, are original.

Contents

Chapter 1

Nonstandard principles of arithmetic

In this chapter:

- ▷ We describe the proof system of Heyting arithmetic in all finite types, and a definitional extension with types for finite sequences, with new forms of application and λ-abstraction. - *Section 1.1*
- ▷ We define an extension of Heyting arithmetic with a *standardness* predicate, inspired by Nelson's idea of creating a nonstandard universe through *syntax*. - *Section 1.2*
- ▷ Overspill and underspill are well-known nonstandard techniques. We introduce the new principles of *sequence* overspill and underspill, a useful generalisation to all finite types, and map their relations to familiar principles from nonstandard and constructive analysis alike. - *Sections 1.3, 1.4*
- ▷ We give an overview of transfer principles and rules, with their different degrees of acceptability for constructive mathematics. - *Section 1.5*
- ▷ After a brief comment on saturation principles, we explain the idea of *herbrandisation*, and provide an original analysis of the so-called nonclassical realisation principle as a herbrandised *uniformity* principle. - *Section 1.6*

1.1 Intuitionistic arithmetic

We want to build a framework in which to study the proof theory of (more or less) constructive, first order nonstandard arithmetic. Unless we wanted to get really quirky, this should be an enrichment of traditional proof systems of arithmetic; these have been the subject of excellent, and extremely detailed textbook presentations. In particular, [59] is the standard reference here; and [26] is also a very good source. The following is that mandatory section where we give a brief, poorer introduction in order to try and make this thesis self-contained.

Notation. In typed systems, we use σ, τ, ρ as symbols for arbitrary types. We write $a : \sigma$ to indicate that a term a is of type σ. For each type σ, we assume that we have a denumerable stock of variables, $x, y, z, \ldots : \sigma$. We use the primitive logical symbols $\bot$, $\wedge$, $\vee$, $\rightarrow$, $\forall x : \sigma$, $\exists x : \sigma$. We define $\neg A := A \rightarrow \bot$, $A \leftrightarrow B := (A \rightarrow B) \wedge (B \rightarrow A)$.

When introducing a proof system, there is usually a choice between a "Hilbert-style" presentation (many axioms, few rules) and a "natural deduction" one (basically, everything is a rule). In most cases, there is a way to switch from one to the other system, preserving provability. The rule of thumb, here, is that it is easier to prove things *about* Hilbert-style systems, and that it is easier to do formal proofs *in* natural deduction.

We are not going to do fully formal proofs. We will use informal intuitionistic reasoning, of the kind suggested by the *Brouwer-Heyting-Kolmogorov* (BHK) *interpretation*:

$\bot$ has no proof;

a proof of $A \wedge B$ is a pair $\langle a, b \rangle$ such that a proves A and b proves B;

a proof of $A \vee B$ is a pair $\langle z, p \rangle$ such that either $z = 0$ and p is a proof of A, or $z = 1$ and p is a proof of B;

a proof of $A \rightarrow B$ is a construction that transforms any proof of A into a proof of B;

a proof of $\forall x : \sigma\ A$ is a construction producing, for every a of type σ, a proof of $A[a/x]$;

a proof of $\exists x : \sigma\ A$ is a pair $\langle a, p \rangle$ such that $a : \sigma$ and p is a proof of $A[a/x]$.

Here, $A[a/x]$ denotes the formula where occurrences of x in A have been replaced by a.

Of course, this does not actually define anything rigorous; and formalisations of the notion of *proof* here, such as Kleene's recursive realisability (reviewed in [60]), can lead to incompatibility with classical mathematics, which is not what we are looking for. Nevertheless, resorting to BHK gives an intuition of what is valid intuitionistically, and what is not.

The most important principle that fails intuitionistically, and is in fact characteristic of classicality, is the law of the excluded middle:

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

Under BHK, this amounts to knowing for *every* proposition A whether it is A, or $\neg A$, that holds.

Anyway - since we are going to reason informally, but Chapter 2 will be all about soundness proofs, Hilbert-style is the natural choice.

Definition 1.1. The language of *Heyting arithmetic* , $\mathcal{L}(\mathrm{HA})$ is an untyped language including a constant 0, a unary function symbol S, and function symbols for all primitive recursive functions (see [54] for a primer of recursion theory).

The logical axioms and rules of HA are

1. the **structural axioms** of intuitionistic logic:

$$\begin{array}{lll} \mathsf{contraction}: & A \vee A \rightarrow A, & A \rightarrow A \wedge A; \\ \mathsf{weakening}: & A \rightarrow A \vee B, & A \wedge B \rightarrow A; \\ \mathsf{exchange}: & A \vee B \rightarrow B \vee A, & A \wedge B \rightarrow B \wedge A; \end{array}$$

2. the **logical axioms** of intuitionistic first order predicate logic:

$$\begin{aligned} \textsf{ex falso quodlibet :}\quad & \bot \to A; \\ & \forall x\, A \to A[a/x], \quad A[a/x] \to \exists x\, A, \end{aligned}$$

   with the usual side conditions on free variables;

3. the **logical rules** of intuitionistic first order predicate logic:

$$\begin{array}{llll} \textsf{modus ponens :} & \dfrac{A \quad A \to B}{B}; & \textsf{syllogism :} & \dfrac{A \to B \quad B \to C}{A \to C}; \\ \textsf{exportation :} & \dfrac{A \wedge B \to C}{A \to (B \to C)}; & \textsf{importation :} & \dfrac{A \to (B \to C)}{A \wedge B \to C}; \\ \textsf{expansion :} & \dfrac{A \to B}{C \vee A \to C \vee B}; & & \\ \dfrac{B \to A}{B \to \forall x\, A}, & \dfrac{A \to B}{\exists x\, A \to B}, & x \text{ not free in } B; & \end{array}$$

4. the **equality axioms**:

$$\begin{aligned} & x = x, \quad x = y \to y = x, \quad x = y \wedge y = z \to x = z, \\ & x = y \to fx = fy \end{aligned}$$

   for all function symbols f of $\mathcal{L}(\mathrm{HA})$.

The nonlogical axioms of HA are

1. the **successor axioms**:

$$\begin{aligned} & \neg \mathrm{S}x = 0, \\ & \mathrm{S}x = \mathrm{S}y \to x = y; \end{aligned}$$

2. defining equations for the primitive recursive functions;

3. the **induction schema**:

$$\textsf{IA :} \quad \big(\varphi(0) \wedge \forall x\, (\varphi(x) \to \varphi(\mathrm{S}x))\big) \to \forall x\, \varphi(x)\ .$$

We will often write $x + 1$ instead of $\mathrm{S}x$.

One then obtains *Peano arithmetic* as the "classicalisation" of HA - the system $\mathrm{HA} + \textsf{LEM}$.

1.1.1 Arithmetic in all finite types

The interpretations we are going to consider in Chapter 2 actually apply to stronger systems of many-sorted arithmetic *in all finite types.*

Definition 1.2. The type structure $\mathbf{T}$ of *finite types* is generated by the inductive clauses

▷ 0 is in **T**;

▷ if σ, τ are in **T**, then $\sigma \to \tau$ is in **T**.

When unbracketed, it is assumed that the type-forming operation is right-associative, e.g. $\rho \to \sigma \to \tau \equiv \rho \to (\sigma \to \tau)$.

The idea, here, is that 0 should be the type of natural numbers; $\sigma \to \tau$ should be a class of mappings from elements of type σ to elements of type τ. One speaks of a *type degree*, defined as

$$\deg(0) := 0 \ ,$$
$$\deg(\sigma \to \tau) := \max(\deg(\sigma) + 1, \deg(\tau)) \ .$$

All the finite types have a finite degree - that is why they are called so. Elements of degree higher than 1 are usually called *functionals*.

One immediate advantage of introducing types of higher degree is that one can formalise analysis, to some extent: once a coding of the rationals onto the natural numbers has been chosen, Cauchy sequences of rationals - hence, real numbers - can be represented by elements of type $0 \to 0$.

Definition 1.3. The language of N-HA$^\omega$ is many-sorted with **T** as its collection of types, and includes constants $0 : 0$ (zero), $\mathrm{S} : 0 \to 0$ (successor), and, for all types ρ, σ, τ,

$\Pi_{\sigma,\tau} : \sigma \to \tau \to \sigma$ (projector),

$\Sigma_{\rho,\sigma,\tau} : (\rho \to \tau \to \sigma) \to ((\rho \to \tau) \to (\rho \to \sigma))$ (combinator),

$\mathrm{R}_\sigma : \sigma \to (\sigma \to 0 \to \sigma) \to (0 \to \sigma)$ (recursor).

The class of terms of N-HA$^\omega$ is defined inductively as follows:

▷ a constant $c : \sigma$ or a variable $x : \sigma$ is a term of type σ;

▷ if $a : \sigma$, $b : \sigma \to \tau$, then $ba : \tau$.

For all types σ, we have an equality symbol $=_\sigma$. The *formulae* of N-HA$^\omega$ are generated by the inductive clauses

▷ for all types σ and terms $s, t : \sigma$, $s =_\sigma t$ is a formula;

▷ if φ, ψ are formulae, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\forall x : \sigma\, \varphi$, and $\exists x : \sigma\, \varphi$ are formulae.

Notice that in arithmetic, we can define $\perp := 0 =_0 \mathrm{S}0$.

The logical axioms and rules of N-HA$^\omega$ are those of many-sorted intuitionistic first order predicate logic, as in the description of HA, with the appropriately typed quantifiers. The nonlogical axioms are

1. the successor axioms and the induction schema, restricted to type 0;
2. the defining axioms of $\Pi_{\sigma,\tau}$, $\Sigma_{\rho,\sigma,\tau}$, and R_σ:

$$\begin{array}{ll}
\Pi_{\sigma,\tau} xy =_\sigma x \ , & x : \sigma, \ y : \tau; \\
\Sigma_{\rho,\sigma,\tau} xyz =_\sigma xz(yz) \ , & x : \rho \to \sigma \to \tau, \ y : \rho \to \sigma, \ z : \rho; \\
\left\{ \begin{array}{l} \mathrm{R}_\sigma xy0 =_\sigma x \ , \\ \mathrm{R}_\sigma xy(\mathrm{S}z) =_\sigma y(\mathrm{R}_\sigma xyz)z \ , \end{array} \right. & x : \sigma, \ y : \sigma \to 0 \to \sigma, \ z : 0 \ .
\end{array}$$

From the projector and combinator, one can define the λ-*abstraction* operator, such that

$$\text{N-HA}^\omega \vdash (\lambda x.b)a =_\tau b[a/x] \;, \qquad x, a : \sigma, \; b : \tau \;.$$

Again, details can be found in [59].

The letter N in N-HA$^\omega$ stands for *neutral*, with respect to the treatment of *equality* in higher types. Depending on one's necessities, the system can be upgraded to one of the following.

- The system I-HA$^\omega$ with *intensional equality*. This is arguably the choice that is more in the constructive spirit, since it makes equality decidable for all types. It amounts to adding constants $E_\sigma : \sigma \to \sigma \to 0$ for all types σ, such that

$$\begin{aligned} &E_\sigma xy = 0 \leftrightarrow x =_\sigma y \;, \\ &E_\sigma xy = 0 \vee E_\sigma xy = 1 \;. \end{aligned}$$

- The system E-HA$^\omega_0$, where equality for types of higher degree is *extensional*, i.e. the axiom
$$\forall f, g : \sigma \to \tau \, (f =_{\sigma\to\tau} g \leftrightarrow \forall x : \sigma \, fx =_\tau gx)$$
holds for all σ, τ. This is the system that is more convenient for use with the functional interpretations of Chapter 2.

A variant of E-HA$^\omega_0$, called just E-HA$^\omega$, has higher-type equality as a *defined* notion: $=_0$ is the only equality symbol that is added to the language, and $f =_{\sigma\to\tau} g$ is read as an abbreviation of $\forall x : \sigma \, fx =_\tau gx$. The functionals are forced to behave extensionally by the axiom

$$\mathsf{EXT} : \quad x =_\sigma y \to fx =_\tau fy \;, \qquad x, y : \sigma \;, f : \sigma \to \tau.$$

Another variant we will use is WE-HA$^\omega$, where the extensionality axiom is replaced with the strictly weaker extensionality *rule*

$$\mathsf{EXT\text{-}R} : \quad \frac{fx_1 \ldots x_n =_0 gx_1 \ldots x_n \qquad \varphi_{\mathrm{qf}}(f)}{\varphi_{\mathrm{qf}}(g)} \;,$$

where φ_{qf} is a *quantifier-free* formula, and the variables are of appropriate types.

1.1.2 Finite sequences

By the combinatory completeness of the pair $(\Pi_{\sigma,\sigma}, \Sigma_{\sigma,\sigma,\sigma})$ for all types σ, it is possible to introduce in N-HA$^\omega$ a coding of *finite sequences* of elements of any type. As we will see, finite sequences (as first order stand-ins for sets) are quite ubiquitous in arguments of nonstandard arithmetic. Thus, in order to lighten the syntax, it seems preferable to introduce an extension of this system, with primitive types for finite sequences (hence, also variables running over finite sequences). This is a definitional extension, therefore conservative over the base system.

Following [64], we choose E-HA$^\omega_0$ as our base.

Definition 1.4. The type structure $\mathbf{T}^*$ is the closure of $\mathbf{T}$ under the additional clause

▷ if σ is in $\mathbf{T}^*$, then σ^* is in $\mathbf{T}^*$.

Notation. We use s, t, u, v (and $s', t', \ldots$) as variables of sequence type.

Definition 1.5. The system E-HA$^{\omega *}$ is the extension of E-HA$_0^\omega$ with types for finite sequences, whose language includes, for all types σ, τ, additional constants $\langle\rangle_\sigma : \sigma$ (empty sequence), $C : \sigma \to \sigma^* \to \sigma^*$ (prepending operator), and $\mathrm{L}_{\sigma,\tau} : \sigma \to (\sigma \to \tau \to \sigma) \to (\tau^* \to \sigma)$ (list recursor), with defining axioms

$$\mathsf{SA}: \quad \forall s : \sigma^* \left(s = \langle\rangle_\sigma \vee \exists x : \sigma\, \exists s' : \sigma^* \left(s = Cxs'\right)\right) ,$$

$$\begin{cases} \mathrm{L}_{\sigma,\tau} xy \langle\rangle_\tau =_\sigma x , \\ \mathrm{L}_{\sigma,\tau} xy(Czs) =_\sigma y(\mathrm{L}_{\sigma,\tau} xys)\langle z\rangle , \end{cases} \qquad x : \sigma,\ y : \sigma \to \tau \to \sigma,\ z : \tau,\ s : \tau^* ,$$

where $\langle z\rangle$ is the "singleton" $Cz\langle\rangle_\tau$. All the axioms and rules of E-HA$_0^\omega$, extended to the types in $\mathbf{T}^*$, are carried over to E-HA$^{\omega *}$.

Let us add to the language, for every type σ, a constant $\emptyset_\sigma$. Using the list recursor, one can define all the basic operations on finite sequences one needs in practice.

(i) A **length function** $|\cdot| : \sigma^* \to 0$, satisfying

$$|\langle\rangle_\sigma| = 0 , \qquad |Cas| = \mathrm{S}|s| ,$$

for $s : \sigma^*$, $a : \sigma$.

(ii) A **projection function** $(s, i) \mapsto s_i$ of type $\sigma^* \to 0 \to \sigma$, satisfying

$$\begin{aligned} (\langle\rangle_\sigma)_i &= \emptyset_\sigma \qquad \text{for all } i, \\ (Cas)_0 &= a , \\ (Cas)_{\mathrm{S}i} &= s_i . \end{aligned}$$

(iii) A **concatenation** operation $\cdot : \sigma^* \to \sigma^* \to \sigma^*$, such that

$$\langle\rangle_\sigma \cdot t = t , \qquad Cas \cdot t = Ca(s \cdot t) .$$

As expected, concatenation is provably associative, so we will iterate it without bothering with brackets.

As we mentioned, finite sequences are our replacement for sets.

Definition 1.6. Let $a : \sigma$, $s, s' : \sigma^*$. We define the abbreviations

(i) $a \in_\sigma s := \exists i < |s|\, (a =_\sigma s_i)$ (a is an *element* of s);

(ii) $s' \subseteq_\sigma s := \forall x : \sigma\, (x \in_\sigma s' \to x \in_\sigma s)$ (s' is *contained* in s).

We will drop subscripts in most occasions. We also extend the relation $\subseteq_\sigma$ to sequence-valued functionals, pointwise: for $s', s : \tau \to \sigma^*$,

(iii) $s' \subseteq s := \forall x : \tau\, (s'x \subseteq_\sigma sx)$.

The relation $\subseteq$ determines a preorder, provably in E-HA$^{\omega *}$.

Definition 1.7. A formula $\Phi(s)$ is *upwards closed* in $s : \sigma^*$ if

$$\Phi(s') \wedge s' \subseteq s \rightarrow \Phi(s) \ .$$

The following, easy properties are all established in [64].

Lemma 1.8. *(a)* $\text{E-HA}^{\omega *} \vdash \forall s : \sigma^* \, (|s| = 0 \leftrightarrow s = \langle\rangle_\sigma)$,

(b) $\text{E-HA}^{\omega *} \vdash \forall n : 0 \, \forall s : \sigma^* \, (|s| = \mathrm{S}n \leftrightarrow \exists x : \sigma \, \exists t : \sigma^* \, (s = Cxt \wedge |t| = n))$.

Proof. Let $s : \sigma^*$. By the sequence axiom SA, either $s = \langle\rangle_\sigma$ or $s = Cxt$ for some $x : \sigma$, $t : \sigma^*$. If $|s| = 0$, the latter case leads to a contradiction, for $|s| = \mathrm{S}|t| > 0$.

If $|s| = \mathrm{S}n$, then the former case leads to a contradiction, and we have proven the directions left to right. The converses are immediate. □

Proposition 1.9. $\text{E-HA}^{\omega *}$ *proves the induction schema for sequences*

$$\mathsf{IA}^* : \quad \big(\varphi(\langle\rangle_\sigma) \wedge \forall x : \sigma \, \forall s : \sigma^* \, (\varphi(s) \rightarrow \varphi(Cxs))\big) \rightarrow \forall s : \sigma^* \, \varphi(s) \ .$$

Proof. Suppose $\varphi(\langle\rangle_\sigma)$ and $\forall x : \sigma \, \forall s : \sigma^* \, (\varphi(s) \rightarrow \varphi(Cxs))$. By the previous lemma,

$$\forall s : \sigma^* \, (|s| = 0 \rightarrow \varphi(s)) \ .$$

Fix $n : 0$, and assume $\forall s : \sigma^* \, (|s| = n \rightarrow \varphi(s))$. Let s be of length $\mathrm{S}n$. Again by the previous lemma, $s = Cxt$ for some $x : \sigma$, and $t : \sigma^*$ of length n, and $\varphi(t)$ holds by hypothesis. Therefore, $\varphi(Cxt) \equiv \varphi(s)$ holds as well; and we have proved

$$\forall s : \sigma^* \, (|s| = n \rightarrow \varphi(s)) \rightarrow \forall s : \sigma^* \, (|s| = \mathrm{S}n \rightarrow \varphi(s)) \ .$$

By ordinary induction, it follows that $\forall n : 0, s : \sigma^* \, (|s| = n \rightarrow \varphi(s))$. □

Corollary 1.10. $\text{E-HA}^{\omega *}$ *proves that*

$$\forall s, t : \sigma^* \, |s \cdot t| = |s| + |t| \ ,$$

and that

$$\begin{cases} (s \cdot t)_i = s_i \ , & i < |s| \ , \\ (s \cdot t)_i = t_{i - |s|} & \text{otherwise} \ . \end{cases}$$

Proof. Follows from the recursive definition of length and concatenation, and the induction schema for sequences. □

An easy consequence is that, for all $s, t : \sigma^*$, $\text{E-HA}^{\omega *}$ proves that $s, t \subseteq s \cdot t$.

Definition 1.11. Let $s, t : \sigma^*$. We say that s and t are *extensionally equal*, and write $s =_e t$, if

$$|s| = |t| \wedge \forall i < |s| \, (s_i = t_i) \ .$$

Corollary 1.12. $\text{E-HA}^{\omega *} \vdash \forall s, t : \sigma^* \, (s =_e t \rightarrow s = t)$.

Proof. By induction for sequences. Suppose $s =_e t$. If $s = \langle\rangle_\sigma$, then $|s| = |t| = 0$, so, by Lemma 1.8, $t = \langle\rangle_\sigma$.

Otherwise, $s = Cxs'$ for some x, s'. Then $|s| = |t| = \mathrm{S}n$ for $n = |s'|$; again, by Lemma 1.8, $t = Cyt'$ for some y, t'. But $x = s_0 = t_0 = y$, and $s' =_e t'$; by the inductive hypothesis, $s' = t'$. Therefore, $s = Cxs' = Cyt' = t$. □

It is possible to define a form of *application* and of λ-*abstraction* for sequences.

Definition 1.13. Let $s : (\sigma \to \tau^*)^*$, $a : \sigma$, $t : \sigma \to \tau^*$. Then

$$s[a] := (s_0 a) \cdot \ldots \cdot (s_{|s|-1} a) : \tau^* \ ,$$
$$\Lambda x.t := C(\lambda x.t)\langle\rangle : (\sigma \to \tau^*)^* \ .$$

The following proposition justifies the definition we just gave.

Proposition 1.14. E-HA$^{\omega *}$ *proves that for all* $s : \sigma \to \tau^*$, $a : \sigma$,

$$(\Lambda x.s)[a] = (\lambda x.s)a = sa \ .$$

An important feature of the sequence application is that it is provably *monotone* in the first component.

Lemma 1.15. E-HA$^{\omega *}$ *proves that for all* $s, s' : (\sigma \to \tau^*)^*$, $a : \sigma$,

$$s \subseteq s' \to s[a] \subseteq s'[a] \ .$$

Proof. Let $i < |s[a]|$. By definition of sequence application, there exist a $j < |s|$ and a $k < |s_j a|$ such that

$$s[a]_i = (s_j a)_k \ .$$

Since $s \subseteq s'$, there is a $\ell < |s'|$ such that $s_j = s'_\ell$, hence $s[a]_i = (s'_\ell a)_k$ by extensionality.

But $s'_\ell a \subseteq s'[a]$, so there exists some $m < |s'[a]|$ such that

$$s[a]_i = (s_j a)_k = s'[a]_m \ .$$

Therefore, $s[a] \subseteq s'[a]$. □

1.2 The syntactic approach to the nonstandard

It has been known at least since 1934, Skolem's [57], that the Peano axioms of arithmetic (*a fortiori*, the Heyting axioms as well) admit models $^*\mathbb{N}$ containing *infinite* natural numbers; that is, numbers that are larger than those obtained by a finite iteration of the successor. These models all contain an initial segment isomorphic to the *standard* model, the good old $\mathbb{N}$.

It took, however, a little longer to figure out that these were good for something. Once we do arithmetic inside a nonstandard model, we would like to derive properties that are also true of the standard model; in other words, a "link" is needed between truth in the nonstandard model and truth in the standard model. This link is commonly called a *transfer* theorem.

From a nonstandard model of the natural numbers, one can build nonstandard models of the rationals and of the reals, including *infinitesimal* as well as infinite numbers; and, again, one would want to obtain a nice transfer theorem.

In 1958, Schmieden and Laugwitz [56] proposed the first explicit model of nonstandard analysis. It had the advantage of being fully constructive, but a quite weak transfer property.

The breakthrough and limited popularisation of nonstandard arithmetic and analysis came in the '60s, the first edition of Robinson's *Non-standard analysis* [53] describing a nonstandard model of the natural numbers that is an *elementary extension* of the standard model; that is, an extension preserving truth of first order statements.

Moreover, Robinson's models allowed for a formal "calculus of infinitesimals", that was reminiscent of early, Euler-style analysis; leading to proofs of known theorems that were arguably shorter and leaner - and even to some new results, e.g. [5]. In subsequent years, nonstandard analysis won some illustrious supporters - notably, Gödel [16, pp. 311-312]:

> *"There are good reasons to believe that nonstandard analysis, in some version or other, will be the analysis of the future."*

Nevertheless, there were many issues, besides habit, which contributed to the new methods never gaining mainstream adoption.

(i) The construction of the models relied on results, such as the compactness theorem, or the existence of nonprincipal ultrafilters of subsets, which make an essential use of the axiom of choice.

   It is a well-known result, due to Diaconescu [11], that the full set-theoretic axiom of choice implies LEM, that is classicality. This led to full rejection from the constructivist community, championed by Bishop [7]; but a certain suspiciousness was more widespread, towards these objects that "exist, but have been seen by nobody".

(ii) More generally, the proliferation of non-equivalent models, which, however, supported similar modes of reasoning, led to a feeling of disconnectedness of the new methods from the usual set-theoretic foundations.

Already in the ending of [53], Robinson suggested that one could view nonstandard analysis as introducing "*new deductive procedures* rather than new mathematical entities", i.e. proof-theoretically rather than model-theoretically. This route was followed as early as in 1969 by Kreisel [29]; and, perhaps most notably, led to the development of Nelson's *internal set theory* [38] in 1977.

Definition 1.16. The theory IST (internal set theory) is an extension of ZFC (Zermelo-Fraenkel set theory with the axiom of choice - for which see any textbook on set theory, e.g. [22]) including

- a unary predicate $\mathrm{st}(x)$ (x is standard), and associated quantifiers

$$\begin{aligned} \forall^{\mathrm{st}} x \ldots &:= \forall x\, (\mathrm{st}(x) \to \ldots)\ , \\ \exists^{\mathrm{st}} x \ldots &:= \exists x\, (\mathrm{st}(x) \wedge \ldots)\ ; \end{aligned}$$

- the three axiom schemata

$$\begin{aligned} \textsf{idealisation}: &\quad \forall^{\mathrm{st\,fin}} z\, \exists x\, \forall y \in z\, \varphi(x,y) \leftrightarrow \exists x\, \forall^{\mathrm{st}} y\, \varphi(x,y)\ , \\ \textsf{standardisation}: &\quad \forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \forall^{\mathrm{st}} z\, (z \in y \leftrightarrow z \in x \wedge \Phi(z))\ , \\ \textsf{transfer}: &\quad \forall^{\mathrm{st}} t_1, \ldots, t_n\, (\forall^{\mathrm{st}} x\, \varphi(x) \to \forall x\, \varphi(x))\ , \end{aligned}$$

  where φ ranges over *internal formulae*, i.e. formulae not containing the standardness predicate; and Φ ranges over all formulae.

The names of the axiom schemata also form the acronym IST, which is quite convenient.

We will later meet these principles in some form, and examine their connection to other nonstandard principles; yet this is no place for a thorough exploration of IST, to which [52] is a very readable introduction, for those interested. What matters, for now, is that IST supports most of the usual forms of nonstandard reasoning; that it does so in a way that feels somewhat like ordinary mathematics (apart from some annoyances with set formation) - as if the standard elements were "already there", but could not be accessed without the standardness predicate; and that it is a *conservative extension* of ZFC.

Nelson provided two proofs of the latter fact: a model-theoretic one, in his original article; and a proof-theoretic one [40], in the guise of a *reduction algorithm*, translating nonstandard into standard proofs. This "interpretation" of IST into ZFC was the cue for the authors of [64] to look into functional interpretations of nonstandard arithmetic.

As we are interested in constructive nonstandard arithmetic, it is noteworthy that standardisation is, in some way, both less acceptable *and* less relevant than the other two axiom schemata of IST. It is the origin, in IST, of the so-called *standard part* map defined on limited real numbers, which is pinpointed in [68] as the source of most nonconstructiveness in nonstandard proofs. However, in [39], Nelson developed a large chunk of nonstandard probability theory using idealisation and transfer, and dispensing with full standardisation, which therefore seems to be less essential.

Of course, ZFC is the epitome of a tremendously strong system, and its underlying logic is classical. Something more manageable, and constructive, such as E-HA$^{\omega*}$, seems a better starting point, if one is interested in computational content hidden within nonstandard proofs.

Definition 1.17. The system E-HA$^{\omega*}_{\mathrm{st}}$ is an extension of E-HA$^{\omega*}$, whose language includes a (unary) predicate $\mathrm{st}_\sigma(x)$, $x : \sigma$, for all types σ of $\mathbf{T}^*$; and the *external quantifiers* $\forall^{\mathrm{st}} x : \sigma$, $\exists^{\mathrm{st}} x : \sigma$.

> **Notation.** Following Nelson, so-called *internal* formulae - those in the language of E-HA$^{\omega*}$ - are always denoted with small Greek letters, and generic, *external* formulae with capital Greek letters.

The following axioms are added to those of E-HA$^{\omega*}$:

1. the defining axioms of the external quantifiers:
$$\forall^{\mathrm{st}} x : \sigma\, \Phi(x) \leftrightarrow \forall x : \sigma\, (\mathrm{st}_\sigma(x) \to \Phi(x)) \ ,$$
$$\exists^{\mathrm{st}} x : \sigma\, \Phi(x) \leftrightarrow \exists x : \sigma\, (\mathrm{st}_\sigma(x) \wedge \Phi(x)) \ ;$$
2. axioms for the standardness predicate:
$$\begin{aligned} &\mathrm{st}_\sigma(x) \wedge x =_\sigma y \to \mathrm{st}_\sigma(y) \ , && \\ &\mathrm{st}_\sigma(a) && \text{for all } \textit{closed } a : \sigma \ , \\ &\mathrm{st}_{\sigma\to\tau}(f) \wedge \mathrm{st}_\sigma(x) \to \mathrm{st}_\tau(fx) \ ; && \end{aligned}$$
3. the **external induction** schema:
$$\mathsf{IA}^{\mathrm{st}} : \quad \big(\Phi(0) \wedge \forall^{\mathrm{st}} x : 0\, (\Phi(x) \to \Phi(\mathrm{S}x))\big) \to \forall^{\mathrm{st}} x : 0\, \Phi(x) \ .$$

The "internal" induction schema IA is assumed to hold for internal formulae only.

Finally, we define disjunction differently than in the base system:

$$\Phi \vee \Psi := \exists^{\mathrm{st}} z : 0\,(z = 0 \to \Phi \wedge \neg\, z = 0 \to \Psi)\ .$$

The basic linguistic blocks are in place, but there is nothing inherently nonstandard about the system we have so far defined. In fact, one could interpret $\mathrm{st}_\sigma(x)$ as $x =_\sigma x$, and all the new axioms would be provable in E-HA$^{\omega*}$. This simple fact also implicates that E-HA$^{\omega*}_{\mathrm{st}}$ is a conservative extension of E-HA$^{\omega*}$.

However, there are some simple facts, of the kind we would expect from a "standardness property", that can already be proved.

Proposition 1.18. *For every formula $\Phi(x)$,* E-HA$^{\omega*}_{\mathrm{st}}$ *proves*

$$\Phi(x) \wedge x = y \to \Phi(y)\ .$$

Proof. Easy induction on the logical structure of Φ, utilising the fact that the standardness predicate is extensional. □

Proposition 1.19. E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall n, m : 0\,(\mathrm{st}_0(n) \wedge m \leq n \to \mathrm{st}_0(m))$.

Proof. Apply external induction to the formula $\Phi(n) := \forall m : 0\,(m \leq n \to \mathrm{st}_0(m))$. □

Definition 1.20. Let $s : \sigma^*$. We say that s is (properly) *finite* if $|s|$ is standard. We say that s is *hyperfinite* if $|s|$ is not standard.

We show that basically anything one can get from standard sequences is standard.

Lemma 1.21. *(a)* E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s : \sigma^*\,(\mathrm{st}(s) \to \mathrm{st}(|s|))$,

(b) E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s : \sigma^*\,(\mathrm{st}(s) \to \forall i < |s|\ \mathrm{st}(s_i))$,

(c) E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s : \sigma^*\,\forall x : \sigma\,(\mathrm{st}(s) \wedge x \in_\sigma s \to \mathrm{st}(x))$,

(d) E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s, t : \sigma^*\,(\mathrm{st}(s) \wedge \mathrm{st}(t) \to \mathrm{st}(s \cdot t))$,

(e) E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall f : 0 \to \sigma^*\,\forall n : 0\,\big(\mathrm{st}(f) \wedge \mathrm{st}(n) \to \mathrm{st}(f0 \cdot \ldots \cdot fn)\big)$.

Proof. Everything follows from the standardness axioms, coupled with the fact that the list recursor is standard. □

A simple consequence of the lemma is that the operations of sequence application and abstraction, as defined in the previous section, preserve standardness.

Corollary 1.22. *(a)* E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s : (\sigma \to \tau^*)^*\,\forall x : \sigma\,\big(\mathrm{st}(s) \wedge \mathrm{st}(x) \to \mathrm{st}(s[x])\big)$,

(b) E-HA$^{\omega*}_{\mathrm{st}} \vdash \forall s : \sigma \to \tau^*\,(\mathrm{st}(s) \to \mathrm{st}(\Lambda x.s))$.

Finally, we prove that finite sequences of standard elements are standard; the converse is already a consequence of Lemma 1.21.(a)-(b).

Lemma 1.23. E-HA$^{\omega*}_{\mathrm{st}}$ *proves that*

$$\forall s : \sigma^*\,\big(\mathrm{st}(|s|) \wedge \forall i < |s|\ \mathrm{st}(s_i) \to \mathrm{st}(s)\big)\ .$$

Proof. Suppose $s : \sigma^*$ is finite, and that, for all $i < |s|$, s_i is standard. By an iteration of Lemma 1.21.(d), $s' := s_0 \cdot \ldots \cdot s_{|s|-1}$ is also standard. Clearly, s and s' are extensionally equal; by Corollary 1.12, $s = s'$. Thus, s is standard. $\square$

This, in turn, is used to prove an external induction schema for sequences.

Proposition 1.24. E-HA$^{\omega*}_{\text{st}}$ *proves the external induction schema for sequences*

$$\mathsf{IA}^{*\text{st}} : \quad \big(\Phi(\langle\rangle_\sigma) \wedge \forall^{\text{st}} x : \sigma\, \forall^{\text{st}} s : \sigma^* \,(\Phi(s) \to \Phi(Cxs))\big) \to \forall^{\text{st}} s : \sigma^*\, \Phi(s)\ .$$

Proof. From the previous lemma, one obtains that if $s = Cxt$ and s is standard, then x and t are also standard. Then one argues precisely as in Proposition 1.9, applying external instead of ordinary induction. $\square$

Enough with the preliminaries - we are now ready to discuss actual nonstandard arithmetic; starting with the principle that supplies us with “enough” nonstandard elements.

1.3 The overspill principle

Of course, we want nonstandard numbers to exist; no need to explain why. But how should the transition from standard to nonstandard happen?

We are already able to prove, in E-HA$^{\omega*}_{\text{st}}$, that there is no smallest nonstandard natural number. By ordinary induction, one proves that every number $n > 0$ has a predecessor, $n - 1$. Suppose $\neg\,\text{st}_0(n)$. If $n - 1$ were standard, then also $n - 1 + 1 = n$ would be standard, for the successor function is. Therefore, $\neg\,\text{st}_0(n - 1)$.

More generally, we do not want to be able to discriminate between standard and nonstandard by *internal* means; for otherwise, what would be the point of introducing the external language at all? So, for instance, an internal property that holds up to *any* standard number should also hold up to some nonstandard, “hyperfinite” number.

The **overspill principle** takes care of this for type 0:

$$\mathsf{OS}_0 : \quad \forall^{\text{st}} n : 0\, \varphi(n) \to \exists n : 0\, (\neg\,\text{st}(n) \wedge \varphi(n))\ .$$

Remark. Unless we specify otherwise, all the principles we consider may have additional parameters besides those explicitly shown.

The first benchmark of any respectable nonstandard proof system is met.

Proposition 1.25. *In* E-HA$^{\omega*}_{\text{st}}$*, the principle* OS_0 *implies that nonstandard natural numbers exist:*

$$\exists n : 0\, \neg\,\text{st}(n)\ .$$

Proof. Apply OS_0 with $\varphi(n) := n = n$, or any other tautology. $\square$

One would want to extend overspill to all finite types. The obvious strategy would be to just formulate it with no type restriction:

$$\mathsf{OS} : \quad \forall^{\text{st}} x : \sigma\, \varphi(x) \to \exists x : \sigma\, (\neg\,\text{st}(x) \wedge \varphi(x))\ .$$

As it turns out, this does not seem to be very useful. The reason is that overspill is almost always applied to formulae like

$$\forall^{\text{st}} n : 0\, \forall k < n\, \varphi(k)\ ,$$

stating that some internal property holds *up to* any finite number.

Example 1.26. This is taken from [46]. Let G be a countably infinite graph; in our setting, a function $G : 0 \to 0 \to 0$, such that

$$\forall i, j : 0 \, (Gij = 0 \vee Gij = 1) \; .$$

For all $n : 0$, we say that a *n-colouring* of G is a function $c : 0 \to 0$ such that

$$\forall i : 0 \, (ci < n) \wedge \forall i, j : 0 \, (Gij = 1 \to \neg \, (ci = cj)) \; .$$

Suppose that any *finite* subgraph of G can be n-coloured. This can be stated as follows:

$$\forall^{\mathrm{st}} k : 0 \, \exists c : 0 \to 0 \, \big(\forall i < k \, (ci < n) \wedge \forall i, j < k \, (Gij = 1 \to \neg \, (ci = cj))\big) \; .$$

Then, by overspill, there exist a *nonstandard* k and an n-colouring c such that

$$\forall i < k \, (ci < n) \wedge \forall i, j < k \, (Gij = 1 \to \neg \, (ci = cj)) \; ,$$

implying

$$\forall^{\mathrm{st}} i : 0 \, (ci < n) \wedge \forall^{\mathrm{st}} i, j : 0 \, (Gij = 1 \to \neg \, (ci = cj)) \; .$$

With a suitable transfer principle (namely, $\mathsf{TP}_\forall$ of Section 1.5), we would be able to turn external into ordinary quantifiers, and derive that G has a n-colouring. This is known as the *de Bruijn-Erdős* theorem.

So, the specificity of type 0 is that, when something holds up to a nonstandard number, it holds for all standard numbers. Given a nonstandard $n : 0$, we can define a hyperfinite sequence

$$s := \langle 0 \rangle \cdot \ldots \cdot \langle n \rangle$$

of type 0^*, with the property that s contains all standard natural numbers. This can be fruitfully generalised.

Definition 1.27. Let $s : \sigma^*$. We say that s is a *hyperfinite enumeration* of the type σ if

$$\forall^{\mathrm{st}} x : \sigma \, (x \in s) \; .$$

We introduce, for all types σ, a defined predicate $\mathrm{hyper}(s)$, such that

$$\forall s : \sigma^* \, (\mathrm{hyper}(s) \leftrightarrow \forall^{\mathrm{st}} x : \sigma \, (x \in s)) \; ,$$

as well as quantifiers ranging over hyperfinite enumerations, with defining axioms

$$\forall^{\mathrm{hyp}} s : \sigma^* \, \Phi(s) \leftrightarrow \forall s : \sigma^* \, (\mathrm{hyper}_\sigma(s) \to \Phi(s)) \; ,$$
$$\exists^{\mathrm{hyp}} s : \sigma^* \, \Phi(s) \leftrightarrow \exists s : \sigma^* \, (\mathrm{hyper}_\sigma(s) \wedge \Phi(s)) \; .$$

This yields a definitional extension of $\text{E-HA}^{\omega *}_{\mathrm{st}}$, for which we keep the name.

Lemma 1.28. $\text{E-HA}^{\omega *}_{\mathrm{st}}$ *proves that*

$$\forall^{\mathrm{st}} s : \sigma^* \, \exists^{\mathrm{st}} x : \sigma \, \neg \, x \in s \; .$$

Proof. Since standard sequences are finite, this amounts to showing that the standard elements of finite types are not *Kuratowski-finite*, i.e. do not admit finite enumerations. This is trivial for type 0; then one proceeds by induction on the type structure. □

Corollary 1.29. $\text{E-HA}^{\omega*}_{\text{st}} \vdash \forall s : \sigma^* \neg(\text{st}(s) \wedge \text{hyper}(s))$.

We introduce the principle of **sequence overspill**:

$$\mathsf{OS}^* : \quad \forall^{\text{st}} s : \sigma^* \varphi(s) \to \exists^{\text{hyp}} s : \sigma^* \varphi(s) \ .$$

Proposition 1.30. *In* $\text{E-HA}^{\omega*}_{\text{st}}$, *the principle* OS^* *implies that hyperfinite enumerations exist for all types:*

$$\exists s : \sigma^* \,\text{hyper}(s) \ .$$

Proof. As in Proposition 1.25: apply OS^* with $\varphi(s) := s = s$. □

Proposition 1.31. *In* $\text{E-HA}^{\omega*}_{\text{st}}$, *the principle* OS^* *implies for all types* σ *the following equivalence, dubbed* enumeration overspill*:*

$$\mathsf{EOS} : \quad \forall^{\text{st}} x : \sigma\, \varphi(x) \leftrightarrow \exists^{\text{hyp}} s : \sigma^* \,\forall x \in s\, \varphi(x).$$

Proof. Suppose $\forall^{\text{st}} x : \sigma\, \varphi(x)$. By Lemma 1.21.(c), this implies

$$\forall^{\text{st}} s : \sigma^* \,\forall x \in s\, \varphi(x) \ .$$

Then $\exists^{\text{hyp}} s : \sigma^* \,\forall x \in s\, \varphi(x)$ follows by sequence overspill. The right to left direction is trivial. □

The relevance of sequence overspill will now become apparent. Consider the following, typed version of Nelson's $\mathsf{idealisation}$:

$$\mathsf{I} : \quad \forall^{\text{st}} s : \sigma^* \,\exists y : \tau \,\forall x \in s\, \varphi(x,y) \to \exists y : \tau \,\forall^{\text{st}} x : \sigma\, \varphi(x,y) \ .$$

It so happens that these two principles are *equivalent*!

Proposition 1.32. $\text{E-HA}^{\omega*}_{\text{st}} \vdash \mathsf{I} \leftrightarrow \mathsf{OS}^*$.

Proof. Assume I, and suppose $\forall^{\text{st}} s : \sigma^* \varphi(s)$. Let $t : (\sigma^*)^*$ be a standard sequence of sequences; then $s := t_0 \cdot \ldots \cdot t_{|t|-1}$ is again standard, so $\varphi(s)$ holds. Furthermore, by construction, for all $i < |t|$, $t_i \subseteq s$; in other words,

$$\forall^{\text{st}} t : (\sigma^*)^* \,\exists s : \sigma^* \,\forall t' \in t\, (t' \subseteq s \wedge \varphi(s)) \ .$$

By idealisation, we obtain

$$\exists s : \sigma^* \,\forall^{\text{st}} t : \sigma^* \,(t \subseteq s \wedge \varphi(s)) \ .$$

It remains to prove that $\forall^{\text{st}} t : \sigma^* \,(t \subseteq s) \leftrightarrow \text{hyper}(s)$, an easy consequence of Lemma 1.21.

Conversely, assume OS^*, and suppose $\forall^{\text{st}} s : \sigma^* \,\exists y : \tau \,\forall x \in s\, \varphi(x,y)$. By sequence overspill, it follows that

$$\exists y : \tau \,\exists^{\text{hyp}} s : \sigma^* \,\forall x \in s\, \varphi(x,y) \ ,$$

which implies

$$\exists y : \tau \,\forall^{\text{st}} x : \sigma^* \,\varphi(x,y) \ .$$

This concludes the proof. □

Several consequences of I are listed in [46] and in [64], which, by the previous proposition, are also consequences of OS^*.

Proposition 1.33. *In* E-HA$^{\omega*}_{\text{st}}$, OS^* *implies that*

$$\forall^{\text{hyp}} s : \sigma^* \, \exists x \in s \, (\neg \, \text{st}(x)) \ ;$$

every hyperfinite enumeration also enumerates some nonstandard elements.

Proof. Assume, equivalently, idealisation, and let s be a hyperfinite enumeration of the type σ. By Corollary 1.29, s cannot be extensionally equal to any standard sequence t; hence,

$$\forall^{\text{st}} t : \sigma^* \, \exists x : \sigma \, \forall y \in t \, (x \in s \wedge \neg \, x = y) \ .$$

By idealisation, it follows that

$$\exists x : \sigma \, \forall^{\text{st}} y : \sigma \, (x \in s \wedge \neg \, x = y) \ ,$$

which implies $\exists x \in s \, (\neg \, \text{st}(x))$. □

Corollary 1.34. *In* E-HA$^{\omega*}_{\text{st}}$*, the principle* OS^* *implies* OS.

Proof. Suppose $\forall^{\text{st}} x : \sigma \, \varphi(x)$. By enumeration overspill, $\exists^{\text{hyp}} s : \sigma^* \, \forall x \in s \, \varphi(x)$; by the previous proposition, any such sequence s contains a nonstandard element. □

Corollary 1.35. *In* E-HA$^{\omega*}_{\text{st}}$*, the principle* OS^* *implies that nonstandard elements of any type exist:*

$$\exists x : \sigma \, \neg \, \text{st}(x) \ .$$

The following is the first example of an external version of a principle that is rejected by strict constructivism, yet implied by a nonstandard principle. The functional interpretations of next chapter will, nevertheless, give it a constructive justification.

The *lesser limited principle of omniscience* for natural numbers is usually stated as follows:

$$\mathsf{LLPO}_0 : \quad \forall a, b : 0 \to 0 \, \big(\forall n, m : 0 \, (an = 0 \vee bm = 0) \to (\forall n : 0 \, an = 0 \vee \forall n : 0 \, bn = 0)\big) \ .$$

Informally, this amounts to the following. Take any pair of sequences of natural numbers. Pick a position in the first, and a position in the second. If, for all possible choices, you find at least one zero, then one of the sequences is constantly zero.

This is called a "principle of omniscience", because it enables one to pass from a partial, finite view of the sequence, to a global, "infinitary" knowledge of it. Within Bishop-style, informal constructive reasoning, it is known to be equivalent to a surprising number of classical theorems, including

(i) the real numbers being an integral domain;

(ii) the Weak König Lemma, that every infinite binary tree has an infinite path;

(iii) the intermediate value theorem of classical analysis;

(iv) the Hahn-Banach theorem of functional analysis.

Even more equivalences are listed in [21]. We will see now that sequence overspill implies the following, external generalisation of LLPO_0 to all finite types:

$$\mathsf{LLPO}^{\mathrm{st}}: \quad \forall^{\mathrm{st}} x, y : \sigma\, (\varphi(x) \vee \psi(y)) \to (\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \;\vee\; \forall^{\mathrm{st}} x : \sigma\, \psi(x)) \,.$$

Proposition 1.36. $\text{E-HA}^{\omega *}_{\mathrm{st}} + \mathsf{OS}^* \vdash \mathsf{LLPO}^{\mathrm{st}}$.

Proof. Suppose $\forall^{\mathrm{st}} x, y : \sigma\, (\varphi(x) \vee \psi(y))$. We prove by external sequence induction that

$$\forall^{\mathrm{st}} s : \sigma^*\, (\forall x \in s\, \varphi(x) \vee \forall x \in s\, \psi(x)) \,. \tag{1.1}$$

For $s = \langle\rangle_\sigma$, $\forall x \in s\, \varphi(x) \vee \forall x \in s\, \psi(x)$ is vacuously true. Suppose it is true for some arbitrary, standard s, and pick any standard $a : \sigma$.

Suppose $\forall x \in s\, \varphi(x)$, and pick $b \in Cas$. Since

$$\forall^{\mathrm{st}} x : \sigma\, (\varphi(a) \vee \psi(x)) \,,$$

either $\varphi(a)$ holds, in which case we are done, or $\psi(b)$ holds. In the latter case, since b was arbitrary in Cas,

$$\forall x \in Cas\, \psi(x) \,,$$

and again we achieve the desired disjunction. Now, applying sequence overspill to (1.1) gives

$$\exists^{\mathrm{hyp}} s : \sigma^*\, (\forall x \in s\, \varphi(x) \vee \forall x \in s\, \psi(x)) \,,$$

which implies $\mathsf{LLPO}^{\mathrm{st}}$. □

Notice that OS_0 alone would have sufficed to prove the restriction of $\mathsf{LLPO}^{\mathrm{st}}$ to type 0; unsurprisingly, as sequence overspill collapses to ordinary overspill in type 0.

It seems likely to us that $\mathsf{LLPO}^{\mathrm{st}}$ implies results that are analogous to those proved by its internal version, but confirming this goes beyond the scope of this thesis.

1.4 The underspill principle

Classically, the overspill principle is equivalent to its dual, the **underspill principle**:

$$\mathsf{US}_0: \quad \forall n : 0\, (\neg\, \mathrm{st}(n) \to \varphi(n)) \to \exists^{\mathrm{st}} n : 0\, \varphi(n) \,.$$

Intuitionistically, this does not seem to be the case.

Underspill also has a direct generalisation to higher types, viz.

$$\mathsf{US}: \quad \forall x : \sigma\, (\neg\, \mathrm{st}(x) \to \varphi(x)) \to \exists^{\mathrm{st}} x : \sigma\, \varphi(x) \,,$$

for which, much like OS, we did not find any use.

One of the most common uses of overspill and underspill is giving suggestive nonstandard characterisations of notions from standard mathematics, as in the following example.

Example 1.37. For once, we will make a completely informal use of constructive reasoning; no particular metatheory is implied.

Let (X, d_X) and (Y, d_Y) be metric spaces. Then the usual definition of *uniform continuity* of a function $f : X \to Y$, translated to the nonstandard setting, is

$$\forall^{\mathrm{st}} n \, \exists^{\mathrm{st}} m \, \forall x, x' \in X \Big(d_X(x, x') < \frac{1}{m} \to d_Y(f(x), f(x')) < \frac{1}{n} \Big) .$$

However, a tale is often told that what this actually *means* is that f carries "infinitely close" points of X to "infinitely close" points of Y. With nonstandard analysis, one can make this rigorous, and introduce a relation

$$x \simeq_X x' \quad \text{if and only if} \quad \forall^{\mathrm{st}} n \; d_X(x, x') < \frac{1}{n} ,$$

to be read "x and x' are infinitely close in X". Suppose that f is such that

$$\forall x, x' \in X \, (x \simeq_X x' \to f(x) \simeq_Y f(x')) .$$

Expanding the definition, applying the overspill principle to the premise of the implication, and using intuitionistic logic to move quantifiers to the front, we obtain

$$\forall^{\mathrm{st}} n \, \forall m \left(\neg \, \mathrm{st}(m) \to \forall x, x' \in X \left(d_X(x, x') < \frac{1}{m} \to d_Y(f(x), f(x')) < \frac{1}{n} \right) \right) .$$

By underspill, this gives the traditional definition of uniform continuity. The converse statement is easy, so the two characterisations really are equivalent.

On to the useful generalisation of underspill to higher types - **sequence underspill**:

$$\mathsf{US}^* : \quad \forall^{\mathrm{hyp}} s : \sigma^* \, \varphi(s) \to \exists^{\mathrm{st}} s : \sigma^* \, \varphi(s) .$$

Proposition 1.38. *In* E-HA$^{\omega *}_{\mathrm{st}}$*, the principle* US^* *implies for all types* σ *the following equivalence, dubbed* enumeration underspill*:*

$$\mathsf{EUS} : \quad \forall^{\mathrm{hyp}} s : \sigma^* \, \exists x \in s \, \varphi(x) \leftrightarrow \exists^{\mathrm{st}} x : \sigma \, \varphi(x).$$

Proof. Suppose $\forall^{\mathrm{hyp}} s : \sigma^* \, \exists x \in s \, \varphi(x)$. By sequence underspill, $\exists^{\mathrm{st}} s : \sigma^* \, \exists x \in s \, \varphi(x)$. Lemma 1.21.(c) then leads to the conclusion. The converse is immediate. □

Unlike sequence overspill, sequence underspill does not seem to imply the "useless" version in all types; it still needs overspill for that purpose.

Proposition 1.39. E-HA$^{\omega *}_{\mathrm{st}}$ $+ \mathsf{OS}^* \vdash \mathsf{EUS} \to \mathsf{US}$.

Proof. Suppose $\forall x : \sigma \, (\neg \, \mathrm{st}(x) \to \varphi(x))$. By Proposition 1.33, sequence overspill implies $\forall^{\mathrm{hyp}} s : \sigma^* \, \exists x \in s \, (\neg \, \mathrm{st}(x))$; therefore, by the hypothesis,

$$\forall^{\mathrm{hyp}} s : \sigma^* \, \exists x \in s \, \varphi(x) .$$

By enumeration underspill, this is equivalent to $\exists^{\mathrm{st}} x : \sigma \, \varphi(x)$. □

Enumeration underspill has the intuitive meaning that "the intersection of all hyperfinite enumerations are the standard elements"; there are no elements that stand out, leaking into as soon as we step into nonstandard territory.

There is an intriguing consequence of the pair $\mathsf{EOS} + \mathsf{EUS}$. The functional interpretations from Chapter 2 will translate external formulae Φ into formulae that look like

$$\exists^{\mathrm{st}} x : \sigma\, \forall^{\mathrm{st}} y : \tau\, \varphi(x, y) \ ,$$

with φ internal. With EOS and EUS, these turn out to be equivalent to

$$\exists^{\mathrm{hyp}} t : \tau^*\, \forall^{\mathrm{hyp}} s : \sigma^*\, \exists x \in s\, \forall y \in t\, \varphi(x, y) \ .$$

We do not know, however, whether there is something deeper in this observation.

Idealisation also has a dual, named *realisation* in [64]:

$$\mathsf{R} : \quad \forall y : \tau\, \exists^{\mathrm{st}} x : \sigma\, \varphi(x, y) \to \exists^{\mathrm{st}} s : \sigma^*\, \forall y : \tau\, \exists x \in s\, \varphi(x, y) \ .$$

Since I is equivalent to OS^*, it would make sense if R were equivalent to US^*; yet things are not so simple. In fact, only one implication seems to hold.

Proposition 1.40. $\text{E-HA}^{\omega *}_{\mathrm{st}} + \mathsf{US}^* \vdash \mathsf{R}$.

Proof. Suppose $\forall y : \tau\, \exists^{\mathrm{st}} x : \sigma\, \varphi(x, y)$. By enumeration underspill, this is equivalent to

$$\forall^{\mathrm{hyp}} s : \sigma^*\, \forall y : \tau\, \exists x \in s\, \varphi(x, y) \ ,$$

which, by sequence underspill, implies $\exists^{\mathrm{st}} s : \sigma^*\, \forall y : \tau\, \exists x \in s\, \varphi(x, y)$. □

What is missing, in order to obtain an equivalence, is the following principle, boldly called the *herbrandised generalised Markov's principle* in [64]:

$$\mathsf{HGMP}^{\mathrm{st}} : \quad (\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \psi) \to \exists^{\mathrm{st}} s : \sigma^*\, (\forall x \in s\, \varphi(x) \to \psi) \ .$$

There will be time, later, to explain "herbrandised"; *Markov's principle*, on the other hand, is another axiom schema strict constructivists reject: in type 0, it says

$$\mathsf{MP}_0 : \quad \big(\forall n : 0\, (\varphi(n) \vee \neg\, \varphi(n)) \wedge \neg\neg\, \exists n : 0\, \varphi(n)\big) \to \exists n : 0\, \varphi(n) \ .$$

One important consequence of Markov's principle is that if we derive a contradiction from a real number *not* being greater than 0, then it *is* greater than 0:

$$\forall x \in \mathbb{R}\, (\neg\neg\, x > 0 \to x > 0) \ .$$

Replacing ψ with a contradiction, e.g. $0 =_0 1$, and choosing a *negated* $\varphi(x)$, we see that $\mathsf{HGMP}^{\mathrm{st}}$ implies an external version of Markov's principle in all finite types:

$$\mathsf{MP}^{\mathrm{st}} : \quad \big(\forall^{\mathrm{st}} x : \sigma\, (\varphi(x) \vee \neg\, \varphi(x)) \wedge \neg\neg\, \exists^{\mathrm{st}} x : \sigma\, \varphi(x)\big) \to \exists^{\mathrm{st}} x : \sigma\, \varphi(x) \ ,$$

which justifies its name.

Proposition 1.41. $\text{E-HA}^{\omega *}_{\mathrm{st}} + \mathsf{US}^* \vdash \mathsf{HGMP}^{\mathrm{st}}$.

Proof. Suppose $\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \psi$. Then

$$\exists^{\mathrm{hyp}} s : \sigma^* \, \forall x \in s\, \varphi(x) \to \psi \ ,$$

which is intuitionistically equivalent to

$$\forall^{\mathrm{hyp}} s : \sigma^* \, (\forall x \in s\, \varphi(x) \to \psi) \ .$$

An application of sequence underspill leads to the conclusion. □

As for overspill and LLPO, notice that US_0 is sufficient to prove the external Markov's principle for the type of natural numbers,

$$\mathsf{MP}_0^{\mathrm{st}} : \quad \big(\forall^{\mathrm{st}} n : 0\, (\varphi(n) \vee \neg\, \varphi(n)) \wedge \neg\neg\, \exists^{\mathrm{st}} n : 0\, \varphi(n)\big) \to \exists^{\mathrm{st}} n : 0\, \varphi(n) \ .$$

The previous result can actually be slightly strenghtened, to also provide a form of *independence of premise*:

Proposition 1.42. *In* $\text{E-HA}^{\omega*}_{\mathrm{st}}$, US^* *implies the principle*

$$(\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \exists^{\mathrm{st}} y : \tau\, \psi(y)) \to \exists^{\mathrm{st}} s : \sigma^* \, \exists^{\mathrm{st}} t : \tau^* \, (\forall x \in s\, \varphi(x) \to \exists y \in t\, \psi(y)) \ .$$

Proof. Suppose $\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \exists^{\mathrm{st}} y : \tau\, \psi(y)$. Then

$$\exists^{\mathrm{hyp}} s : \sigma^* \, \forall x \in s\, \varphi(x) \to \forall^{\mathrm{hyp}} t : \tau^* \, \exists y \in t\, \psi(y) \ ,$$

which, by intuitionistic logic, is equivalent to

$$\forall^{\mathrm{hyp}} s : \sigma^* \, \forall^{\mathrm{hyp}} t : \tau^* \, (\forall x \in s\, \varphi(x) \to \exists y \in t\, \psi(y)) \ .$$

Applying sequence underspill yields the desired result. □

We now complete the characterisation of US^*.

Proposition 1.43. $\text{E-HA}^{\omega*}_{\mathrm{st}} + \mathsf{HGMP}^{\mathrm{st}} + \mathsf{R} \vdash \mathsf{US}^*$.

Proof. Suppose $\forall^{\mathrm{hyp}} s : \sigma^* \, \varphi(s)$; that is,

$$\forall s : \sigma^* \, (\forall^{\mathrm{st}} x : \sigma\, (x \in s) \to \varphi(s)) \ .$$

By the herbrandised generalised Markov's principle, this is equivalent to

$$\forall s : \sigma^* \, \exists^{\mathrm{st}} t : \sigma^* \, (t \subseteq s \to \varphi(s)) \ ;$$

which, by realisation and intuitionistic logic, implies

$$\exists^{\mathrm{st}} t : (\sigma^*)^* \, \forall s : \sigma^* \, (\forall t' \in t\, (t' \subseteq s) \to \varphi(s)) \ .$$

Take a standard $t : (\sigma^*)^*$ as in (1.4), and pick $s := t_0 \cdot \ldots \cdot t_{|t|-1}$. By Lemma 1.21, s is standard, and, for all $t' \in t$, $t' \subseteq s$; therefore, it holds that $\varphi(s)$. We thus prove

$$\exists^{\mathrm{st}} s : \sigma^* \, \varphi(s) \ ,$$

and the sequence overspill principle. □

In [64], I, $\mathsf{HGMP}^{\mathrm{st}}$ and a stronger, nonclassical version of R (see Section 1.6.2) were used as characteristic principles for the D_{st} interpretation (Chapter 2). These can now be eliminated in favour of US^* and OS^*, which seem to us more intuitive, and more in the spirit of nonstandard arithmetic in type 0.

Since both overspill and underspill force some nonconstructive modes of reasoning, we wonder whether one could weaken them, in such a way that the resulting logic would mirror more closely the usual constructivist practice; perhaps by introducing "layers" of nonstandardness, such that overspill and underspill only "go through" some of them.

Some form of layering is suggested in [55], but we have not investigated the issue any further.

1.5 Transfer principles and rules

In the model-theoretic approach to nonstandard arithmetic, transfer is usually formulated as a *theorem*, expressing the fact that a certain embedding $\iota : \mathbb{N} \to {}^*\mathbb{N}$ of the standard model into a nonstandard model is *elementary*: for all first order formulae of arithmetic $\varphi(x_0, \dots, x_k)$, and all natural numbers $n_0, \dots, n_k$,

$$\mathbb{N} \models \varphi(n_0, \dots, n_k) \quad \text{if and only if} \quad {}^*\mathbb{N} \models \varphi(\iota n_0, \dots, \iota n_k) .$$

Of course, this is not viable in the syntactic approach; so internal set theory has a transfer *principle*. In our setting, $\mathsf{transfer}$ gets decomposed into

$$\mathsf{TP}_\forall : \quad \forall^{\mathrm{st}} y_1 : \tau_1 \dots \forall^{\mathrm{st}} y_n : \tau_n \, (\forall^{\mathrm{st}} x : \sigma \, \varphi(x, y_1, \dots, y_n) \to \forall x : \sigma \, \varphi(x, y_1, \dots, y_n)) ,$$

and the classically, but not intuitionistically equivalent

$$\mathsf{TP}_\exists : \quad \forall^{\mathrm{st}} y_1 : \tau_1 \dots \forall^{\mathrm{st}} y_n : \tau_n \, (\exists x : \sigma \, \varphi(x, y_1, \dots, y_n) \to \exists^{\mathrm{st}} x : \sigma \, \varphi(x, y_1, \dots, y_n)) ,$$

where, for once, no unwritten parameters are allowed.

The constructive character of transfer principles has got somewhat more attention in the literature than other nonstandard principles, being the main focus of [2] and of parts of [37]. And for good reasons: it appears that even mild forms of transfer, in conjunction with overspill and underspill, force classicality, or violate conservativity over Heyting arithmetic.

The following is an adaptation, due to [64], of a result from [37].

Proposition 1.44. *In* $\text{E-HA}^{\omega *}_{\mathrm{st}} + \mathsf{TP}_\forall$*, the existence of hyperfinite enumerations implies the law of the excluded middle for all internal formulae.*

Proof. We proceed by induction on the number of internal quantifiers of a formula φ. If φ is atomic, then the axiom of extensionality can be used to eliminate equality at higher types in favour of equality at type 0, which is decidable. This implies that all atomic formulae are already decidable.

The propositional connectives preserve decidability, so it remains to discuss the case of the quantifiers. Suppose $\varphi(x, y)$ is an internal, decidable formula, where $x : \sigma$, and $y : \tau$ is a placeholder for any additional parameters. By ordinary sequence induction, we can prove

$$\forall^{\mathrm{st}} y : \tau \, \forall s : \sigma^* \, (\forall x \in s \, \varphi(x, y) \vee \exists x \in s \, \neg \varphi(x, y)) . \tag{1.2}$$

Suppose a hyperfinite enumeration s of the type σ exists. Then (1.2) implies

$$\forall^{\mathrm{st}} y : \tau\, (\forall^{\mathrm{st}} x : \sigma\, \varphi(x,y) \vee \exists x : \sigma\, \neg\, \varphi(x,y)) \ .$$

Applying $\mathsf{TP}_\forall$ once, we obtain

$$\forall^{\mathrm{st}} y : \tau\, (\forall x : \sigma\, \varphi(x,y) \vee \exists x : \sigma\, \neg\, \varphi(x,y)) \ ,$$

and, applying it again,

$$\forall y : \tau\, (\forall x : \sigma\, \varphi(x,y) \vee \exists x : \sigma\, \neg\, \varphi(x,y)) \ ,$$

which completes the induction step. □

As usual, by the interchangeability of hyperfinite enumerations of the type 0 with nonstandard natural numbers, the existence of the latter suffices to prove the law of the excluded middle for all internal *arithmetical* formulae.

With underspill principles, things go even more wrong, as shown by [2].

Lemma 1.45. *There exists a primitive recursive formula $\varphi(n)$ of arithmetic such that*

$$\neg\, \forall n\, \neg\, \varphi(n) \to \exists n\, \varphi(n)$$

is not provable in HA.

Proof. Let $\forall n\, \neg\, \varphi(n)$ be a *Rosser sentence* for HA, i.e. a sentence such that HA proves neither it, nor its negation. Suppose

$$\mathrm{HA} \vdash \neg\, \forall n\, \neg\, \varphi(n) \to \exists n\, \varphi(n) \ .$$

As shown in [59, 3.1.7], HA is closed under the following independence of premise *rule*:

$$\mathsf{IPR}^c : \quad \frac{\neg\, \psi \to \exists n\, \varphi(n)}{\exists n\, (\neg\, \psi \to \varphi(n))} \ ,$$

where ψ is required to be closed. Thus, by the *existence property* of HA, we can find $\bar{n}$ such that

$$\neg\, \forall n\, \neg\, \varphi(n) \to \varphi(\bar{n}) \ . \tag{1.3}$$

Since $\varphi(n)$ is primitive recursive, either $\mathrm{HA} \vdash \varphi(\bar{n})$ or $\mathrm{HA} \vdash \neg\, \varphi(\bar{n})$. In the first case,

$$\mathrm{HA} \vdash \neg\, \forall n\, \neg\, \varphi(n) \ ;$$

in the second, by (1.3),

$$\mathrm{HA} \vdash \neg\, \neg\, \forall n\, \neg\, \varphi(n) \ ,$$

or, equivalently, $\mathrm{HA} \vdash \forall n\, \neg\, \varphi(n)$. But this contradicts $\forall n\, \neg\, \varphi(n)$ being a Rosser sentence. □

Proposition 1.46. *In* $\text{E-HA}^{\omega *}_{\mathrm{st}} + \mathsf{US}_0$*, there exist primitive recursive arithmetical formulae $\varphi(n)$, $\psi(n)$ such that either one of the axioms*

$$\forall^{\mathrm{st}} n : 0\, \varphi(n) \to \forall n : 0\, \varphi(n) \ , \tag{1.4}$$

$$\exists n : 0\, \psi(n) \to \exists^{\mathrm{st}} n : 0\, \psi(n) \tag{1.5}$$

destroys conservativity over $\text{E-HA}^{\omega *}$.

Proof. By our comment beneath Proposition 1.41,

$$\text{E-HA}^{\omega*}_{\text{st}} + \mathsf{US}_0 \vdash \mathsf{MP}^{\text{st}}_0 \ .$$

Therefore, taking $\varphi(n)$ as in the previous lemma, we can prove in $\text{E-HA}^{\omega*}_{\text{st}} + \mathsf{US}_0$

$$\neg \forall^{\text{st}} n \neg \varphi(n) \to \exists^{\text{st}} n \, \varphi(n) \ .$$

Assuming (1.4), we would be able to derive

$$\neg \forall n \neg \varphi(n) \to \exists n \, \varphi(n) \ ,$$

which, in $\text{E-HA}^{\omega*}$, we can not. The claim about (1.5) also follows with $\psi(n) := \neg \varphi(n)$, by $\neg\neg$-stability of primitive recursive formulae. □

It would seem that transfer *principles* are a no-go - one can hardly get any weaker. So, why not make them *rules*?

$$\mathsf{TR}_\forall : \quad \frac{\forall^{\text{st}} x : \sigma \, \varphi(x)}{\forall x : \sigma \, \varphi(x)} \ ,$$

$$\mathsf{TR}_\exists : \quad \frac{\exists x : \sigma \, \varphi(x)}{\exists^{\text{st}} x : \sigma \, \varphi(x)} \ .$$

We will see in Chapter 2, thanks to the D_{st} interpretation, that this really is feasible; unless otherwise specified, we will take for granted that the systems of nonstandard arithmetic we consider are closed under these rules.

However, [2] again rules out several possible strenghtenings, at least in the presence of an *overspill* principle.

Lemma 1.47. *There exist primitive recursive arithmetical formulae $\varphi(n)$, $\psi(n)$ such that*

$$\text{HA} \vdash \forall n, m \, (\varphi(n) \vee \psi(m))$$

yet

$$\text{HA} \nvdash \forall n \, \varphi(n) \vee \forall n \, \psi(n) \ .$$

Proof. See [2, Lemma 5.5]. □

Proposition 1.48. *There exist primitive recursive arithmetical formulae $\varphi(n)$, $\psi(n)$ such that*

$$\text{E-HA}^{\omega*}_{\text{st}} + \mathsf{OS}_0 \vdash \forall^{\text{st}} n : 0 \, \varphi(n) \vee \forall^{\text{st}} n : 0 \, \psi(n) \ ,$$

yet

$$\text{E-HA}^{\omega*} \nvdash \forall n : 0 \, \varphi(n) \vee \forall n : 0 \, \psi(n) \ .$$

Proof. Take $\varphi(n)$, $\psi(n)$ as in the previous lemma. Clearly,

$$\text{E-HA}^{\omega*}_{\text{st}} \vdash \forall^{\text{st}} n, m : 0 \, (\varphi(n) \vee \psi(m)) \ .$$

By our comment beneath Proposition 1.36,

$$\text{E-HA}^{\omega*}_{\text{st}} + \mathsf{OS}_0 \vdash \mathsf{LLPO}^{\text{st}}_0 \ ,$$

whence

$$\text{E-HA}^{\omega*}_{\text{st}} + \mathsf{OS}_0 \vdash \forall^{\text{st}} n : 0 \, \varphi(n) \vee \forall^{\text{st}} n : 0 \, \psi(n) \ .$$

The thesis follows by the previous lemma. □

Corollary 1.49. *There exists a primitive recursive arithmetical formula* $\chi(z,n)$ *such that*

$$\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}_0 \vdash \exists^{\text{st}} z : 0\, \forall^{\text{st}} n : 0\, \chi(z,n) \ ,$$

yet

$$\text{E-HA}^{\omega *} \nvdash \exists z : 0\, \forall n : 0\, \chi(z,n) \ .$$

Proof. Take $\chi(z,n) := (z = 0 \wedge \varphi(n)) \vee (z = 1 \wedge \psi(n))$, with $\varphi(n)$ and $\psi(n)$ as in the previous proposition. □

Corollary 1.50. *The system* $\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}_0$ *does not have the disjunction property.*

Proof. By Proposition 1.48, the system proves $\forall^{\text{st}} n : 0\, \varphi(n) \vee \forall^{\text{st}} n : 0\, \psi(n)$. By closure under $\mathsf{TR}_\forall$, though, it cannot prove either disjunct. □

Corollary 1.51. *The system* $\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}_0$ *does not have the existence property - not even restricted to formulae*

$$\exists^{\text{st}} n : 0\, \psi(n) \ .$$

Proof. Follows from Corollary 1.49. □

Finally, one could consider the following rule:

$$\mathsf{TR}_{\forall\exists} : \quad \frac{\forall x : \sigma\, \exists y : \tau\, \varphi(x,y)}{\forall^{\text{st}} x : \sigma\, \exists^{\text{st}} y : \tau\, \varphi(x,y)} \ .$$

In fact, this is satisfied by the topos model of nonstandard arithmetic from Chapter 3, which may indicate that it is constructively acceptable. However, for now we do not know if it is derivable from our proof systems.

1.6 Saturation and uniformity

1.6.1 Saturation principles

One of the most celebrated results of nonstandard analysis has been the construction of *Loeb measures* [33], which have some interesting applications to probability theory and abstract measure theory. This construction utilises a property of certain nonstandard extensions, called *countable saturation.*

In our context, countable saturation is the principle

$$\mathsf{CSAT} : \quad \forall^{\text{st}} n : 0\, \exists y : \tau\, \Phi(n,y) \to \exists f : 0 \to \tau\, \forall^{\text{st}} n : 0\, \Phi(n, fn) \ .$$

This has a straightforward generalisation to arbitrary finite types:

$$\mathsf{SAT} : \quad \forall^{\text{st}} x : \sigma\, \exists y : \tau\, \Phi(x,y) \to \exists f : \sigma \to \tau\, \forall^{\text{st}} x : \sigma\, \Phi(x, fx) \ .$$

As mentioned in [64], countable saturation has a D_{st} interpretation. No such result is known for the general saturation principle, and we have not looked into it.

We have adapted, though, the result of [48] that a combination of saturation and overspill implies an extension of $\mathsf{LLPO}^{\text{st}}$ to certain external formulae.

Proposition 1.52. *In* E-HA$^{\omega *}_{\mathrm{st}} + \mathsf{US}^*$*, the principle* SAT *implies the following principle of omniscience: suppose* $\Phi(x)$*,* $\Psi(x)$*,* $x : \sigma$*, are external formulae such that*

$$\forall^{\mathrm{st}} x : \sigma \, (\Phi(x) \vee \neg \, \Phi(x)) \; , \qquad \forall^{\mathrm{st}} x : \sigma \, (\Psi(x) \vee \neg \, \Psi(x)) \; ;$$

then

$$\forall^{\mathrm{st}} x, y : \sigma \, (\Phi(x) \vee \Psi(y)) \to (\forall^{\mathrm{st}} x : \sigma \, \Phi(x) \; \vee \; \forall^{\mathrm{st}} x : \sigma \, \Psi(x)) \; .$$

Proof. Proceeding by external sequence induction, much like in Proposition 1.36, we can prove

$$\forall^{\mathrm{st}} s : \sigma^* \, (\forall x \in s \, \Phi(x) \vee \forall x \in s \, \Psi(x)) \; . \tag{1.6}$$

By the decidability of Φ, Ψ for standard values of x, we have

$$\forall^{\mathrm{st}} x : \sigma \, \exists z : 0 \, (z = 0 \leftrightarrow \Phi(x)) \; , \qquad \forall^{\mathrm{st}} x : \sigma \, \exists z : 0 \, (z = 0 \leftrightarrow \Psi(x)) \; .$$

By saturation, then, we can find $f, g : \sigma \to 0$ such that

$$\forall^{\mathrm{st}} x : \sigma \, (fx = 0 \leftrightarrow \Phi(x)) \; , \qquad \forall^{\mathrm{st}} x : \sigma \, (gx = 0 \leftrightarrow \Psi(x)) \; ,$$

and, substituting into (1.6),

$$\forall^{\mathrm{st}} s : \sigma^* \, (\forall x \in s \, fx = 0 \vee \forall x \in s \, gx = 0) \; .$$

The formula between brackets is now internal: we can apply sequence overspill, reaching the desired conclusion. □

As usual, US_0 and CSAT are sufficient for the corresponding result in type 0.

1.6.2 Uniformity principles

This section makes a slight detour from the theme of this chapter: the principles we are going to discuss have nothing particularly "nonstandard" about them. The functional interpretations from next chapter will, however, validate them; so this may be regarded as a bridge between what has been, and what follows.

In Section 1.4, we have used the adjective *herbrandised*, perhaps improperly, with reference to $\mathsf{HGMP}^{\mathrm{st}}$. Although there is no need to make it rigorous, the term deserves an explanation.

An idea, that will be further developed in Chapter 2, is that the distinction between external and ordinary quantifiers should be a distinction between *computational* and *non-computational* quantifiers, respectively; that is, quantifiers that carry some constructive information, and others that are void of it. So it should be the *former* that correspond to the usual intuitionistic quantifiers.

In the context of functional interpretations, many useful principles enable one to derive, from the premise

$$\ldots \; \exists^{\mathrm{st}} y : \tau \, \Phi(x, y) \; \ldots \; ,$$

where $x : \sigma$ is a placeholder for any set of additional variables, the consequence

$$\exists^{\mathrm{st}} f : \sigma \to \tau \, (\ldots \; \Phi(x, fx) \; \ldots) \; . \tag{1.7}$$

A soundness theorem then guarantees that one can extract from proofs a closed term $\bar{f}$ such that

$$\ldots\ \Phi(x, \bar{f}x)\ \ldots\ ;$$

such an f is called a *realiser* for the premise, and confirms the computational quality of the quantifier.

The *herbrandised* version of such a principle is the one where (1.7) is replaced by

$$\exists^{\mathrm{st}} f : (\sigma \to \tau^*)^* \,(\ldots\ \exists y \in f[x]\, \Phi(x, y)\ \ldots)\ ;$$

so that, in order the realise the premise, one needs to provide a *finite sequence* of *potential* realisers, of which at least one is an actual realiser. This is reminiscent of *Herbrand disjunctions* in classical logic - whence the name.

For instance, this is the external *axiom of choice*:

$$\mathsf{AC}^{\mathrm{st}} : \quad \forall^{\mathrm{st}} x : \sigma\, \exists^{\mathrm{st}} y : \tau\, \Phi(x, y) \to \exists^{\mathrm{st}} f : \sigma \to \tau\, \forall^{\mathrm{st}} x : \sigma\, \Phi(x, fx)\ ;$$

and this is the *herbrandised* axiom of choice:

$$\mathsf{HAC}^{\mathrm{st}} : \quad \forall^{\mathrm{st}} x : \sigma\, \exists^{\mathrm{st}} y : \tau\, \Phi(x, y) \to \exists^{\mathrm{st}} (f : \sigma \to \tau^*)^*\, \forall^{\mathrm{st}} x : \sigma\, \exists y \in f[x]\, \Phi(x, y)\ .$$

Such versions are the ones that turn out to be compatible with conservativity of systems of nonstandard arithmetic.

We are interested in the following principle - **nonclassical realisation**:

$$\mathsf{NCR} : \quad \forall y : \tau\, \exists^{\mathrm{st}} x : \sigma\, \Phi(x, y) \to \exists^{\mathrm{st}} s : \sigma^*\, \forall y : \tau\, \exists x \in s\, \Phi(x, y)\ .$$

A remarkable consequence of NCR is that it forces nonclassicality.

Proposition 1.53. *In* E-HA$^{\omega *}_{\mathrm{st}}$*, the principle* NCR *implies, for all types, that the standardness predicate is undecidable:*

$$\neg\, \forall x : \sigma\, (\mathrm{st}(x) \vee \neg\, \mathrm{st}(x))\ .$$

Proof. Suppose $\forall x : \sigma\, (\mathrm{st}(x) \vee \neg\, \mathrm{st}(x))$. It would then follow that

$$\forall y : \sigma\, \exists^{\mathrm{st}} x : \sigma\, (\mathrm{st}(y) \to x = y)\ .$$

By nonclassical realisation,

$$\exists^{\mathrm{st}} s : \sigma^*\, \forall y : \sigma\, \exists x \in s\, (\mathrm{st}(y) \to x = y)\ ,$$

which is the statement that there are only finitely many elements of type σ; a contradiction. $\square$

In [64], NCR has been classified as a strenghtening of R, which is reasonable. However, the analysis of the next chapters suggests that it is, more properly, the herbrandisation of the following **nonstandard uniformity** principle:

$$\mathsf{NU} : \quad \forall y : \tau\, \exists^{\mathrm{st}} x : \sigma\, \Phi(x, y) \to \exists^{\mathrm{st}} x : \sigma\, \forall y : \tau\, \Phi(x, y)\ .$$

Nonstandard uniformity is strongly nonclassical, and is, in fact, inconsistent with E-HA$^{\omega *}_{\mathrm{st}}$.

Proposition 1.54. *In* E-HA$^{\omega *}_{\mathrm{st}}$*, the principle* NU *implies*

$$\neg \forall n : 0 \, (n = 0 \vee \neg\, n = 0) \ ,$$

hence a contradiction.

Proof. Suppose $\forall n : 0 \, (n = 0 \vee \neg\, n = 0)$. This is intuitionistically equivalent to

$$\forall n : 0 \, \exists^{\mathrm{st}} z : 0 \, (z = 0 \to n = 0 \wedge \neg\, z = 0 \to \neg\, n = 0) \ ,$$

which, by nonstandard uniformity, implies

$$\exists^{\mathrm{st}} z : 0 \, \forall n : 0 \, (z = 0 \to n = 0 \wedge \neg\, z = 0 \to \neg\, n = 0) \ ,$$

the statement that all natural numbers are zero, or all are non-zero; a contradiction. □

The way out of contradiction is to consider the system E-HA$^{\omega *}_{\mathrm{st}\vee}$, where the ordinary induction schema is limited to internal and $\vee$-*free* formulae.

> **Notation.** If P is an axiom schema where certain schematic variables range over *internal* formulae of E-HA$^{\omega *}_{\mathrm{st}}$, we write $\mathsf{P}_{\vee}$ for the same axiom schema, where "internal" is replaced by "internal and $\vee$-free".

So, in E-HA$^{\omega *}_{\mathrm{st}\vee}$, IA is replaced by $\mathsf{IA}_{\vee}$. Notice that, by external induction, one can still derive

$$\forall^{\mathrm{st}} n : 0 \, (n = 0 \vee \neg\, n = 0)$$

and all other statements provable by ordinary induction, with quantifiers restricted to standard elements. Conservativity over Heyting arithmetic is now limited to formulae without disjunctions.

The reason why we called this a *uniformity* principle is the similarity of

$$\forall s : 0^* \, \exists^{\mathrm{st}} n : 0 \, \Phi(s, n) \to \exists^{\mathrm{st}} n : 0 \, \forall s : 0^* \, \Phi(s, n)$$

to Troelstra's *uniformity principle* [60, Proposition 8.21]

$$\mathsf{UP} : \quad \forall S \subseteq \mathbb{N} \, \exists n \in \mathbb{N} \, \Phi(S, n) \to \exists n \in \mathbb{N} \, \forall S \subseteq \mathbb{N} \, \Phi(S, n) \ ,$$

a second-order principle that is validated by higher-order versions of recursive realisability.

At this point, of course, we are way past nonstandard arithmetic, and fully within the domain of next chapter.

Chapter 2

Functional interpretations

In this chapter:

- ▷ We recall a few basic facts about Gödel's Dialectica, the forefather of all functional interpretations, and its variant by Diller and Nahm, dispensing with decidability of atomic predicates. - *Section 2.1*

- ▷ We trace a concise history of the idea of having two kinds of quantifiers: computational, and noncomputational, or *uniform*; from Lifschitz's philosophical motivations, to its revival through optimised *proof mining*. - *Section 2.2*

- ▷ We define a new functional interpretation, an extension of the Diller-Nahm interpretation with uniform quantifiers. We prove its basic properties, and describe its virtues and shortcomings. - *Sections 2.2, 2.3*

- ▷ We review the *nonstandard Dialectica* interpretation of [64], as a herbrandised version of uniform Diller-Nahm. - *Section 2.4*

2.1 The Dialectica and Diller-Nahm interpretations

The term *functional interpretation* is often used synonimously with Gödel's *Dialectica interpretation*. This is the subject of a chapter in many books on proof theory, e.g. [1], so one should refer to those, more complete treatments; but it seems fit to say something at least, as an introduction to our nonstandard variants, about the functional interpretation *par excellence*.

It was introduced in 1958, [15], a rare example of a mathematical object getting its popular name from the journal where it was first published. It provided a translation of formulae of what we now know as Heyting arithmetic in all finite types, into a quantifier-free fragment of the same system; such that, from an intuitionistic proof of a formula φ, one could extract a *closed term* - a functional of finite type - that "realises" its translation φ_D.

As discussed in [14], the motivation behind Gödel's article was two-fold. There was, of course, an immediate technical advantage - first of all, a simple *relative consistency* proof: the standard contradiction $0 =_0 1$ is interpreted as itself, so if one system proves it, than the other does too; and then a plethora of closure and consistency results for Heyting arithmetic, exhaustively catalogued in [59].

But a *foundational* intent appeared predominant, of substituting the underdefined intuitionistic notion of "proof" with that of "computable functional of finite type", which, albeit abstract as well, seems more precise, and more in the spirit of Hilbert's programme of finitary foundations of mathematics.

When it seemed that Dialectica had run out of possibilities, only a partial success with respect to its goals, it came to an unexpected revival, through the so-called programme of *proof mining*.

Proof mining is the recent rebranding [27] of Kreisel's programme of *unwinding proofs*, which he advanced as early as in 1951 [28]; the idea that some constructive content may be "hidden" within non-constructive proofs, and that through syntactic elaboration one can extract it. With its capability of *locally* composing realisers at each step of a proof, as opposed to techniques such as cut-elimination, the Dialectica interpretation (possibly composed with the *negative translation* of Peano into Heyting arithmetic) turned out to be a valid, and computationally feasible tool for that purpose. Much of [26] is devoted to these applications.

We will now define the Dialectica translation; but first, we need to address some formalities about notation for *tuples* of types and terms, which we postponed in the first chapter.

Notation. We write $\underline{\sigma} := \sigma_1, \dots, \sigma_n$, $\underline{x} : \underline{\sigma} := x_0 : \sigma_0, \dots, x_n : \sigma_n$ for tuples of types and terms. $[\,]$ stands for the empty tuple. We write

$$f\underline{x} := (\dots(fx_0)x_1)\dots)x_n\ ,$$

with the appropriate types; while, if $\underline{f} := f_0, \dots, f_m$, $\underline{f}\underline{x}$ stands for $f_0\underline{x}, \dots, f_m\underline{x}$. We will have, correspondingly,

$$\lambda\underline{x}.\underline{f} := \lambda\underline{x}.f_0, \dots \lambda\underline{x}.f_m\ ,$$

and the same for sequence application.

Relations distribute as expected: for instance, if $\underline{y} := y_0, \dots, y_n$, with the same length *and* types as $\underline{x}$,

$$\underline{x} =_{\underline{\sigma}} \underline{y} := \bigwedge_{i=0}^{n} x_i =_{\sigma_i} y_i\ ;$$

and if $\underline{s} := s_0 : \sigma_0^*, \dots, s_n : \sigma_n^*$ is a tuple of sequences,

$$\underline{x} \in_{\underline{\sigma}} \underline{s} := \bigwedge_{i=0}^{n} x_i \in_{\sigma_i} s_i\ .$$

Definition 2.1. To every formula $\varphi(\underline{a})$ of $\mathcal{L}(\text{WE-HA}^\omega)$, with free variables $\underline{a}$, we associate inductively its *Dialectica translation* $\varphi(\underline{a})^D = \exists\underline{x}\,\forall\underline{y}\,\varphi_D(\underline{x}, \underline{y}, \underline{a})$.

▷ $\varphi(\underline{a})^D := \varphi_D(\underline{a}) := \varphi(\underline{a})$, for φ atomic.

Let $\varphi(\underline{a})^D = \exists\underline{x}\,\forall\underline{y}\,\varphi_D(\underline{x}, \underline{y}, \underline{a})$, $\psi(\underline{b})^D = \exists\underline{u}\,\forall\underline{v}\,\psi_D(\underline{u}, \underline{v}, \underline{b})$:

▷ $(\varphi(\underline{a}) \wedge \psi(\underline{b}))^D := \exists\underline{x}, \underline{u}\,\forall\underline{y}, \underline{v}\,\big(\varphi_D(\underline{x}, \underline{y}, \underline{a}) \wedge \psi_D(\underline{u}, \underline{v}, \underline{b})\big)\ ;$

▷ $(\varphi(\underline{a}) \vee \psi(\underline{b}))^D := \exists z : 0, \underline{x}, \underline{u}\,\forall\underline{y}, \underline{v}\,\big(z = 0 \to \varphi_D(\underline{x}, \underline{y}, \underline{a}) \wedge \neg\, z = 0 \to \psi_D(\underline{u}, \underline{v}, \underline{b})\big)\ ;$

- ▷ $(\varphi(\underline{a}) \to \psi(\underline{b}))^D := \exists \underline{U}, \underline{Y}\, \forall \underline{x}, \underline{v} \left(\varphi_D(\underline{x}, \underline{Yxv}, \underline{a}) \to \psi_D(\underline{Ux}, \underline{v}, \underline{b})\right)$;
- ▷ $(\exists z\, \varphi(z, \underline{a}))^D := \exists z, \underline{x}\, \forall \underline{y}\, \varphi_D(\underline{x}, \underline{y}, z, \underline{a})$;
- ▷ $(\forall z\, \varphi(z, \underline{a}))^D := \exists \underline{X}\, \forall \underline{y}, z\, \varphi_D(\underline{X}z, \underline{y}, z, \underline{a})$.

All formulae are *classically* equivalent to their Dialectica-translated counterparts, provided one includes the *axiom of choice* in "classicality"; and, in fact, even intuitionistically, the clauses for $\wedge$, $\vee$ and $\exists z$ preserve equivalence.

The clause for implication is manifestly the most interesting one, and deserves a step-by-step analysis. One starts with

$$\exists \underline{x}\, \forall \underline{y}\, \varphi_D(\underline{x}, \underline{y}) \to \exists \underline{u}\, \forall \underline{v}\, \psi_D(\underline{u}, \underline{v}) \;;$$

this is intuitionistically equivalent to

$$\forall \underline{x} \left(\forall \underline{y}\, \varphi_D(\underline{x}, \underline{y}) \to \exists \underline{u}\, \forall \underline{v}\, \psi_D(\underline{u}, \underline{v})\right) . \tag{2.1}$$

Suppose we have at hand the following, restricted independence of premise principle:

$$\mathsf{IP}_\forall : \quad (\forall x : \sigma\, \varphi_{\mathrm{qf}}(x) \to \exists y : \tau\, \psi(y)) \to \exists y : \tau\, (\forall x : \sigma\, \varphi_{\mathrm{qf}}(x) \to \psi(y)) \;,$$

where $\varphi_{\mathrm{qf}}(x)$ is quantifier-free (that is, $\forall x : \sigma\, \varphi_{\mathrm{qf}}(x)$ is *purely universal*); then (2.1) becomes equivalent to

$$\forall \underline{x}\, \exists \underline{u}\, \forall \underline{v} \left(\forall \underline{y}\, \varphi_D(\underline{x}, \underline{y}) \to \psi_D(\underline{u}, \underline{v})\right) .$$

By classical logic, we would derive $\forall \underline{x}\, \exists \underline{u}\, \forall \underline{v}\, \exists \underline{y} \left(\varphi_D(\underline{x}, \underline{y}) \to \psi_D(\underline{u}, \underline{v})\right)$; thus, intuitionistically, we obtain

$$\forall \underline{x}\, \exists \underline{u}\, \forall \underline{v}\, \neg\neg\, \exists \underline{y} \left(\varphi_D(\underline{x}, \underline{y}) \to \psi_D(\underline{u}, \underline{v})\right) .$$

In order to proceed, we need yet another nonconstructive principle; a version of Markov's principle:

$$\mathsf{MP}' : \quad \neg\neg\, \exists x : \sigma\, \varphi_{\mathrm{qf}}(x) \to \exists x : \sigma\, \varphi_{\mathrm{qf}}(x) \;.$$

Notice that, in this setting, quantifier free formulae are decidable. With MP', we derive

$$\forall \underline{x}\, \exists \underline{u}\, \forall \underline{v}\, \exists \underline{y} \left(\varphi_D(\underline{x}, \underline{y}) \to \psi_D(\underline{u}, \underline{v})\right) ;$$

so with two applications of the axiom of choice

$$\mathsf{AC} : \quad \forall x : \sigma\, \exists y : \tau\, \varphi(x, y) \to \exists f : \sigma \to \tau\, \forall x : \sigma\, \varphi(x, fx)$$

we get the Dialectica interpretation of implication.

Of course, there are many classically equivalent choices for $\exists\forall$-prenexation, so one may wonder why we should choose these particular ones. However, as explained in [59, 3.5.3], these choices are those with the weakest logical complexity, and in a quite strong sense: any other choice leads to the interpretation of $A \to A$, for suitable A, requiring non-recursive realisers.

The following is the fundamental theorem about the Dialectica translation.

Theorem 2.2 (Soundness of the Dialectica interpretation)**.** *Suppose*

$$\text{WE-HA}^\omega \vdash \varphi(\underline{a}) \ ,$$

with $\varphi(\underline{a})^D = \exists \underline{x} \forall \underline{y}\, \varphi_D(\underline{x}, \underline{y}, \underline{a})$. *Then from the proof we can extract a tuple of closed terms* $\underline{t}$ *such that*

$$\text{WE-HA}^\omega \vdash \forall \underline{y}\ \varphi_D(\underline{t}, \underline{y}, \underline{a}) \ .$$

Proof. See [59, 3.5.4]. □

We have tacitly assumed the *weakly* extensional variant of Heyting arithmetic in all finite types as our base system. This is because the extensionality axioms for types of degree higher than 1 are *not* Dialectica-interpretable, as shown by [18]. The intensional variant would also be fine, but it is not our concern.

The principles we introduced earlier to justify the interpretation of implication are in fact characteristic of the D-translation.

Theorem 2.3 (Characterisation of the Dialectica interpretation)**.** *For all formulae* φ *in the language of* WE-HA$^\omega$*,*

$$\text{WE-HA}^\omega + \mathsf{IP}_\forall + \mathsf{MP}' + \mathsf{AC} \vdash \varphi \leftrightarrow \varphi^D \ .$$

Moreover, if $\text{WE-HA}^\omega + \mathsf{IP}_\forall + \mathsf{MP}' + \mathsf{AC} \vdash \varphi(\underline{a})$*, then one can extract from the proof closed terms* $\underline{t}$ *such that*

$$\text{WE-HA}^\omega \vdash \forall \underline{y}\ \varphi_D(\underline{t}, \underline{y}, \underline{a}) \ .$$

Proof. See [59, 3.5.10]. But we already handled the least obvious case of implication. □

It follows that the Dialectica translation can also be used as a means to *eliminate* these nonconstructive principles from proofs, providing a sort of an intuitionistic justification to their use.

The aches with extensionality, however, are not the only issue with Dialectica. One may want, for instance, to apply a Dialectica-like translation to theories whose atomic formulae are *not* all decidable. With the plain D-translation of implication, one cannot; but this is not due to some fundamental character of the interpretation, compelling one to call for decidability in all sorts of situations - it is a *single* axiom that turns out to be problematic, namely contraction.

The D-translation of contraction is

$$\exists \underline{X}_1, \underline{X}_2, \underline{Y}\, \forall \underline{x}, \underline{y}_1, \underline{y}_2 \left(\varphi(\underline{x}, \underline{Y}\underline{x}\underline{y}_1\underline{y}_2) \to \varphi(\underline{X}_1\underline{x}, \underline{y}_1) \wedge \varphi(\underline{X}_2\underline{x}, \underline{y}_2)\right) ,$$

for φ quantifier-free. For the realisers $\underline{X}_1$ and $\underline{X}_2$ there is only one obvious choice, namely $\underline{X}_1 \equiv \underline{X}_2 := \lambda \underline{x}.\underline{x}$.

But $\underline{Y}$ needs to be a *conditional* operator, such that

$$\underline{Y}\underline{x}\underline{y}_1\underline{y}_2 = \underline{y}_1$$

if it holds that $\neg\, \varphi(\underline{x}, \underline{y}_1)$, and

$$\underline{Y}\underline{x}\underline{y}_1\underline{y}_2 = \underline{y}_2$$

otherwise. This, of course, is possible if and only if φ is decidable.

Incidentally - since the recognition of $\mathsf{contraction}$ as a "problematic" axiom is at the root of *linear logic*, one may guess that Dialectica can be refined to an interpretation of the latter; and so it is [43], the result being quite more symmetric for linear implication than for intuitionistic implication.

So, the D-translation of proofs forces a *decision* every time the contraction axiom is used. Depending on one's needs, this may be considered a feature, since it keeps redundancy to a minimum. However, in other cases, either decidability is not an option, or we want more control over which formulae are interpreted by themselves; the simplest solution is then to interpret $\mathsf{contraction}$ as

$$\exists \underline{X}_1, \underline{X}_2, \underline{Y}\, \forall \underline{x}, \underline{y}_1, \underline{y}_2 \left(\forall \underline{y} \in \underline{Y}\underline{x}\underline{y}_1\underline{y}_2\, \varphi(\underline{x}, \underline{y}) \to \varphi(\underline{X}_1\underline{x}, \underline{y}_1) \wedge \varphi(\underline{X}_2\underline{x}, \underline{y}_2)\right) ,$$

and pick $\underline{Y}\underline{x}\underline{y}_1\underline{y}_2 := \lambda\underline{x}, \underline{y}_1, \underline{y}_2. \langle \underline{y}_1 \rangle \cdot \langle \underline{y}_2 \rangle$, effectively circumventing any decision.

This gives the *Diller-Nahm* variant of the Dialectica interpretation [12].

Definition 2.4. The *Diller-Nahm interpretation* associates to every formula $\varphi(\underline{a})$ of $\mathcal{L}(\text{WE-HA}^{\omega*})$, with free variables $\underline{a}$, a formula $\varphi(\underline{a})^\wedge = \exists \underline{x}\, \forall \underline{y}\, \varphi_\wedge(\underline{x}, \underline{y}, \underline{a})$. The inductive clauses are identical to those of the Dialectica interpretation, except for implication:

$\triangleright$ $(\varphi(\underline{a}) \to \psi(\underline{b}))^\wedge := \exists \underline{U}, \underline{Y}\, \forall \underline{x}, \underline{v} \left(\forall \underline{y} \in \underline{Y}\underline{x}\underline{v}\, \varphi_\wedge(\underline{x}, \underline{y}, \underline{a}) \to \psi_\wedge(\underline{U}\underline{x}, \underline{v}, \underline{b})\right)$.

As sequences do not play such a big role in the Diller-Nahm interpretation, it is not really customary to use a system enriched with types for finite sequences, like WE-HA$^{\omega*}$, as a base; but, having gone to great lengths to set it up, we might as well use it.

Theorem 2.5 (Soundness of the Diller-Nahm interpretation)**.** *Suppose*

$$\text{WE-HA}^{\omega*} \vdash \varphi(\underline{a}) ,$$

with $\varphi(\underline{a})^\wedge = \exists \underline{x}\, \forall \underline{y}\, \varphi_\wedge(\underline{x}, \underline{y}, \underline{a})$. *Then from the proof we can extract a tuple of closed terms* $\underline{t}$ *such that*

$$\text{WE-HA}^{\omega*} \vdash \forall \underline{y}\; \varphi_\wedge(\underline{t}, \underline{y}, \underline{a}) .$$

Proof. We have already handled the case of $\mathsf{contraction}$, which is the only difference with respect to the D-translation. ☐

You may have noticed that, in order to obtain Diller-Nahm implication, one needs, instead of MP', a principle like

$$\mathsf{HGMP} : \quad (\forall x : \sigma\, \varphi(x) \to \psi) \to \exists s : \sigma^*\, (\forall x \in s\, \varphi(x) \to \psi) ;$$

an "internalisation" of $\mathsf{HGMP}^{\mathsf{st}}$, where φ, ψ have the appropriate range. Since, in the nonstandard context, the latter is a consequence of the "natural" principle US^*, our path to nonstandard functional interpretations leads us to Diller-Nahm, rather than Dialectica-style implication.

2.2 Uniform Diller-Nahm

In [32], Lifschitz made the following observation: even in a classical framework, recursion theory has made precise the notion of a calculable *function*; however, there is no such way of speaking about calculable *numbers*. So philosophically puzzling situations arise, such as a function provably having a Gödel number, hence being recursive, and us not knowing which number that is.

So, either one reasons in Heyting arithmetic, where *all* numbers are calculable numbers, or one adopts classical logic, where the distinction is completely lost. Lifschitz's proposal is to enrich the language of Heyting arithmetic with a predicate $\mathrm{K}(n)$, "n is calculable"; and then extend the definition of Kleene's recursive realisability with the clause

▷ $x \mathrel{\mathsf{r}_{\mathrm{K}}} \mathrm{K}(n)$ if and only if $x = n$,

all the while interpreting quantifiers *uniformly*:

▷ $x \mathrel{\mathsf{r}_{\mathrm{K}}} \forall n\, \varphi(n)$ if and only if $\forall n\, (x \mathrel{\mathsf{r}_{\mathrm{K}}} \varphi(n))$,

▷ $x \mathrel{\mathsf{r}_{\mathrm{K}}} \exists n\, \varphi(n)$ if and only if $\exists n\, (x \mathrel{\mathsf{r}_{\mathrm{K}}} \varphi(n))$.

So, by themselves, in this definition quantifiers are completely void of any computational meaning; it is by invoking quantifiers *restricted to calculable numbers*, $\forall n\, (\mathrm{K}(n) \to \ldots)$, and $\exists n\, (\mathrm{K}(n) \wedge \ldots)$, that one restores it.

There is a remarkable similarity between Lifschitz's minimal axioms for the calculability predicate - stability under application, and a restricted, "external" induction schema - and those we imposed on the standardness predicate. But the analogy does not stop here.

In [52], Robert suggests to regard "standardness" as "accessibility" in order to obtain an intuitive understanding of internal set theory; so, for instance, the existence of hyperfinite enumerations becomes the statement that whatever is accessible can be enumerated, corresponding to

"...a practical finiteness feeling for the notion of accessibility." [52, p. 13]

One could substitute "calculable" for "accessible", and it would still all make sense.

A couple of decades later, Lifschitz's demand was rediscovered, from a completely different perspective, in the area of proof mining. Rather than the foundational issue of injecting a certain "modular constructiveness" into classical reasoning, it was the practical problem of more efficient *program extraction* from proofs that was addressed.

Even in fully intuitionistic proofs, a fine-grained analysis reveals instances of formulae with quantifiers that are *computationally redundant*; i.e. the constructive content that is encoded in the quantifiers is never used in the program extracted with the aid of a functional interpretation. This always happens, in particular, when - in a natural deduction setting - an implication introduction discharges more then one instance of the same formula, so that the contraction rule needs to be used.

One would want a way to flag such quantifiers, telling the extraction program to just "pass through" them. This is the function performed by Berger's *uniform* quantifiers [4] and, with an eye to Dialectica extraction, by Hernest's quantifiers *without computational meaning* [17]. But, realisability being a rudimentary functional interpretation - in particular Kreisel's *modified* brand, see [42] - this is also what Lifschitz's calculability predicate achieved!

So, once we accept the analogy between calculability and standardness, we have two ways of looking at the dichotomy between the ordinary quantifiers $\forall x$, $\exists x$ and the external quantifiers $\forall^{\mathrm{st}} x$, $\exists^{\mathrm{st}} x$:

- one that is mediated by nonstandard arithmetic and analysis: $\forall x$, $\exists x$ are the ordinary quantifiers, and it is the universe of types that has been (syntactically) divided into layers;
- one that is inspired by Berger and Hernest: the external quantifiers have computational meaning, and correspond to the usual intuitionistic quantifiers, while the ordinary quantifiers should be interpreted uniformly.

These two viewpoints reach a perfect synthesis in the nonstandard Dialectica, or D_{st} interpretation; but we will get there gradually. Our categorical analysis from Chapter 3 suggested that we could obtain D_{st} in two steps: extending Diller-Nahm with uniform quantifiers; and then "herbrandising", in order to fix a problem with disjunctions.

The connection of this *uniform Diller-Nahm* interpretation to nonstandard arithmetic is quite feeble: it lives in the system E-$\mathrm{HA}^{\omega*}_{\mathrm{st}\vee}$, which, due to the induction schema $\mathsf{IA}_\vee$ being restricted to $\vee$-free formulae, is not even an extension of Heyting arithmetic; so conservativity ("transfer") is restricted. However, forgetting for a moment the usual semantics of the standardness predicate, and thinking of it as more alike to Lifschitz's calculability, one could still find a use for this interpretation in optimised proof mining.

Notice that if one equates "standard" to "calculable", the failure of

$$\forall n : 0\,(n = 0 \vee \neg\, n = 0)$$

starts making sense; for how could we know whether a non-calculable number is zero or non zero?

Definition 2.6. To every formula $\Phi(\underline{a})$ of $\mathcal{L}(\text{E-HA}^{\omega*}_{\mathrm{st}\vee})$, with free variables $\underline{a}$, we associate inductively its *uniform Diller-Nahm* translation $\Phi(\underline{a})^U = \exists^{\mathrm{st}}\underline{x}\,\forall^{\mathrm{st}}\underline{y}\,\varphi_U(\underline{x},\underline{y},\underline{a})$, where φ_U is internal and $\vee$-free.

- ▷ $\varphi(\underline{a})^U := \varphi_U(\underline{a}) := \varphi(\underline{a})$, for φ internal atomic;
- ▷ $\mathrm{st}_\sigma(x)^U := \exists^{\mathrm{st}} y : \sigma\,(y = x)$.

Let $\Phi(\underline{a})^U = \exists^{\mathrm{st}}\underline{x}\,\forall^{\mathrm{st}}\underline{y}\,\varphi_U(\underline{x},\underline{y},\underline{a})$, $\Psi(\underline{b})^U = \exists^{\mathrm{st}}\underline{u}\,\forall^{\mathrm{st}}\underline{v}\,\psi_U(\underline{u},\underline{v},\underline{b})$:

- ▷ $(\Phi(\underline{a}) \wedge \Psi(\underline{b}))^U := \exists^{\mathrm{st}}\underline{x},\underline{u}\,\forall^{\mathrm{st}}\underline{y},\underline{v}\,\big(\varphi_U(\underline{x},\underline{y},\underline{a}) \wedge \psi_U(\underline{u},\underline{v},\underline{b})\big)$;
- ▷ $(\Phi(\underline{a})\vee\Psi(\underline{b}))^U := \exists^{\mathrm{st}} z : 0,\underline{x},\underline{u}\,\forall^{\mathrm{st}}\underline{y},\underline{v}\,\big(z = 0 \to \varphi_U(\underline{x},\underline{y},\underline{a})\wedge\neg\, z = 0 \to \psi_U(\underline{u},\underline{v},\underline{b})\big)$;
- ▷ $(\Phi(\underline{a}) \to \Psi(\underline{b}))^U := \exists^{\mathrm{st}}\underline{U},\underline{Y}\,\forall^{\mathrm{st}}\underline{x},\underline{v}\,\big(\forall \underline{y} \in \underline{Yxv}\,\varphi_U(\underline{x},\underline{y},\underline{a}) \to \psi_U(\underline{Ux},\underline{v},\underline{b})\big)$;
- ▷ $(\exists z\,\Phi(z,\underline{a}))^U := \exists^{\mathrm{st}}\underline{x}\,\forall^{\mathrm{st}}\underline{y}\,\exists z\,\forall \underline{y}' \in \underline{y}\,\varphi_U(\underline{x},\underline{y}',z,\underline{a})$;
- ▷ $(\forall z\,\Phi(z,\underline{a}))^U := \exists^{\mathrm{st}}\underline{x}\,\forall^{\mathrm{st}}\underline{y}\,\forall z\,\varphi_U(\underline{x},\underline{y},z,\underline{a})$;
- ▷ $(\exists^{\mathrm{st}} z\,\Phi(z,\underline{a}))^U := \exists^{\mathrm{st}} z,\underline{x}\,\forall^{\mathrm{st}}\underline{y}\,\varphi_U(\underline{x},\underline{y},z,\underline{a})$;
- ▷ $(\forall^{\mathrm{st}} z\,\Phi(z,\underline{a}))^U := \exists^{\mathrm{st}}\underline{X}\,\forall^{\mathrm{st}}\underline{y}, z\,\varphi_U(\underline{X}z,\underline{y},z,\underline{a})$.

The first thing to notice is that, if this interpretation is restricted to formulae that contain only external quantifiers - or, if you prefer, everything is declared standard/calculable - it is the same as the usual Diller-Nahm translation.

Secondly, the interpretation is *idempotent*: formulae of the form

$$\exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, \underline{a})$$

with φ internal and $\vee$-free are interpreted as themselves, as shown by an easy induction on their structure. This is a feature that we will lose with herbrandisation.

Except for a minor change in the interpretation of the uniform existential quantifier, this interpretation is to the Diller-Nahm variant what Hernest's *light Dialectica interpretation* [17] is to Dialectica. The main difference between the two approaches is that our *external* quantifiers correspond to the ordinary quantifiers in Hernest's setting; the distinction is obtained there with an extension of intuitionistic natural deduction with new rules for the noncomputational quantifiers.

Instead, it is by using the standardness predicate in the spirit of Lifschitz's calculability that we will see the analogy to nonstandard arithmetic.

2.2.1 The soundness theorem

Time for a soundness theorem - it is a long thing, but it needs to be done once. We will not need, however, to handle everything explicitly: except for those concerning the quantifiers, all the logical axioms and rules admit the same realisers as those for the Diller-Nahm interpretation. We will show, though, that uniform Diller-Nahm interprets what we will recognise later as its characteristic principles; these include a form of the following independence of premise principle:

$$\mathsf{IP}^{\mathrm{st}}_{\forall} : \quad (\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \exists^{\mathrm{st}} y : \tau\, \Psi(y)) \to \exists^{\mathrm{st}} y : \tau\, (\forall^{\mathrm{st}} x : \sigma\, \varphi(x) \to \Psi(y)) \ .$$

Recall that, if P is an axiom schema with variables ranging over internal formulae, the same variables range over internal and $\vee$-free formulae in $\mathsf{P}_{\vee}$. We write E-HA$^{\omega*}_{\vee}$ for the system E-HA$^{\omega*}$ with $\mathsf{IA}_{\vee}$ in place of IA.

Theorem 2.7 (Soundness of uniform Diller-Nahm)**.** *Suppose*

$$\text{E-HA}^{\omega*}_{\mathrm{st}\vee} + \mathsf{OS}^*_{\vee} + \mathsf{US}^*_{\vee} + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee} + \Delta_{\vee} \vdash \Phi(\underline{a}) \ ,$$

where $\Delta_{\vee}$ is a set of internal, $\vee$-free sentences. Let $\Phi(\underline{a})^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi_U(\underline{x}, \underline{y}, \underline{a})$. Then from the proof we can extract a tuple of closed terms $\underline{t}$ such that

$$\text{E-HA}^{\omega*}_{\vee} + \Delta_{\vee} \vdash \forall \underline{y}\, \varphi_U(\underline{t}, \underline{y}, \underline{a}) \ .$$

Proof. We proceed by induction on the length of the derivation.

1. *The logical axioms and rules of intuitionistic first order predicate logic.* We consider the quantifier axioms and rules, and give another couple of examples, referring again to [59, 3.5.4] for the rest.

   (i) Example - $\mathsf{weakening}$: $\quad A \to A \vee B$.
   Suppose $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, \underline{a})$, $B^U = \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \psi(\underline{u}, \underline{v}, \underline{b})$. Then

   $$\begin{aligned}(A \to A \vee B)^U = \exists^{\mathrm{st}} \underline{Z}, \underline{X'}, \underline{U}, \underline{S}\, \forall^{\mathrm{st}} \underline{x}, \underline{y'}, \underline{v}\, \big(&\forall \underline{y} \in \underline{S}\underline{x}\underline{y'}\underline{v}\, \varphi(\underline{x}, \underline{y}, \underline{a}) \to \\ &(\underline{Z}\underline{x} = 0 \to \varphi(\underline{X'}\underline{x}, \underline{y'}, \underline{a}) \wedge \neg\, \underline{Z}\underline{x} = 0 \to \psi(\underline{U}\underline{x}, \underline{v}, \underline{b}))\big) \ ,\end{aligned}$$

   and we can take

   $$\begin{aligned}&\underline{Z} := \lambda \underline{x}.0 \ , \quad \underline{X'} := \lambda \underline{x}.\underline{x} \ , \\ &\underline{U} \text{ arbitrary}, \quad \underline{S} := \lambda \underline{x}, \underline{y'}, \underline{v}.\langle \underline{y'} \rangle \ .\end{aligned}$$

(ii) $\forall z\, A \to A[b/z]$.
Suppose $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, z, \underline{a})$. Then

$$(\forall z\, A \to A[b/z])^U = \exists^{\mathrm{st}} \underline{X}', \underline{S}\, \forall^{\mathrm{st}} \underline{x}, \underline{y}' \\ \big(\forall \underline{y} \in \underline{S}\underline{x}\underline{y}'\, \forall z\, \varphi(\underline{x}, \underline{y}, z, \underline{a}) \to \varphi(\underline{X}'\underline{x}, \underline{y}', b, \underline{a})\big)\ ,$$

so we can take

$$\underline{X}' := \lambda \underline{x}.\underline{x}\ , \quad \underline{S} := \lambda \underline{x}, \underline{y}'.\langle \underline{y}' \rangle\ .$$

(iii) $A[b/z] \to \exists z\, A$.
Suppose $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, z, \underline{a})$. Then

$$(A[b/z] \to \exists z\, A)^U = \exists^{\mathrm{st}} \underline{X}', \underline{S}\, \forall^{\mathrm{st}} \underline{x}, \underline{t} \\ \big(\forall \underline{y} \in \underline{S}\underline{x}\underline{t}\, \varphi(\underline{x}, \underline{y}, b, \underline{a}) \to \exists z\, \forall \underline{y}' \in \underline{t}\, \varphi(\underline{X}'\underline{x}, \underline{y}', z, \underline{a})\big)\ ,$$

and we can take

$$\underline{X}' := \lambda \underline{x}.\underline{x}\ , \quad \underline{S} := \lambda \underline{x}, \underline{t}.\underline{t}\ .$$

(iv) Example - modus ponens.
Suppose that $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, \underline{a})$, $B^U = \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \psi(\underline{u}, \underline{v}, \underline{b})$, and that we have terms $\underline{t}_1$ realising the interpretation of A^U and $\underline{T}_2, \underline{T}_3$ realising the interpretation of $(A \to B)^U$.
This means we have

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{y}\, \varphi(\underline{t}_1, \underline{y}, \underline{a})\ ,$$

and

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{x}, \underline{v}\, \big(\forall \underline{y} \in \underline{T}_3\underline{x}\underline{v}\, \varphi(\underline{x}, \underline{y}, \underline{a}) \to \psi(\underline{T}_2\underline{x}, \underline{v}, \underline{b})\big)\ .$$

Taking $\underline{t}_4 := \underline{T}_2\underline{t}_1$, we obtain

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{v}\, \psi(\underline{t}_4, \underline{v}, \underline{b})\ ,$$

as desired.

(v) $\dfrac{B \to A}{B \to \forall z\, A}$.
Suppose that $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, z, \underline{a})$, $B^U = \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \psi(\underline{u}, \underline{v}, \underline{b})$, where z is not free in ψ, and that we have terms $\underline{T}_1, \underline{T}_2$ realising $(B \to A)^U$. Then,

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{u}, \underline{y}\, \big(\forall \underline{v} \in \underline{T}_2\underline{u}\underline{y}\, \psi(\underline{u}, \underline{v}, \underline{b}) \to \varphi(\underline{T}_1\underline{u}, \underline{y}, z, \underline{a})\big)\ .$$

Then $\underline{T}_3 := \underline{T}_1$ and $\underline{T}_4 := \underline{T}_2$ realise the interpretation of $B \to \forall z\, A$.

(vi) $\dfrac{A \to B}{\exists z\, A \to B}$.
Suppose that $A^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x}, \underline{y}, z, \underline{a})$, $B^U = \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \psi(\underline{u}, \underline{v}, \underline{b})$, where z is not free in ψ, and that we have terms $\underline{T}_1, \underline{T}_2$ realising $(A \to B)^U$. Then,

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{x}, \underline{v}\, \big(\forall \underline{y} \in \underline{T}_2\underline{x}\underline{v}\, \varphi(\underline{x}, \underline{y}, z, \underline{a}) \to \psi(\underline{T}_1\underline{x}, \underline{v}, \underline{b})\big)\ .$$

We have

$$(\exists z\, A \to B)^U = \exists^{\mathrm{st}} \underline{U}, \underline{S}\, \forall^{\mathrm{st}} \underline{x}, \underline{v} \\ \big(\forall s \in \underline{Sxv}\, \exists z\, \forall y \in s\, \varphi(\underline{x}, \underline{y}, z, \underline{a}) \to \psi(\underline{Ux}, \underline{v}, \underline{b})\big)\ ;$$

so we can take $\underline{T}_3 := \underline{T}_1$, and $\underline{T}_4 := \lambda \underline{x}, \underline{v}.\langle \underline{T_2 xv} \rangle$, to obtain

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{x}, \underline{v}\, \big(\forall s \in \underline{T_4 xv}\, \exists z\, \forall y \in s\, \varphi(\underline{x}, \underline{y}, z, \underline{a}) \to \psi(\underline{T_3 x}, \underline{v}, \underline{b})\big)\ .$$

2. *The nonlogical axioms of extensional Heyting arithmetic in all finite types* (with the restricted induction schema $\mathsf{IA}_{\vee}$). These are all internal and $\vee$-free, hence are realised by the empty tuple.

3. *The defining axioms of the external quantifiers.* Let $\Phi(x)^U := \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \varphi(\underline{u}, \underline{v}, x)$ here.

   (i) $\forall^{\mathrm{st}} x\, \Phi(x) \leftrightarrow \forall x\, (\mathrm{st}(x) \to \Phi(x))$.
   The interpretation of $\forall^{\mathrm{st}} x\, \Phi(x) \to \forall x\, (\mathrm{st}(x) \to \Phi(x))$ is

$$\exists^{\mathrm{st}} \underline{U'}, S, \underline{T}\, \forall^{\mathrm{st}} \underline{U}, y, \underline{v'}\, \big(\forall x \in S\underline{U}y\underline{v'}\, \forall \underline{v} \in \underline{TU}y\underline{v'}\, \varphi(\underline{U}x, \underline{v}, x) \to \\ \forall x\, (x = y \to \varphi(\underline{U'}\underline{U}y, \underline{v'}, x))\big)\ ;$$

   so we can take

$$\underline{U'} := \lambda \underline{U}, y.\underline{U}y\ , \qquad S := \lambda \underline{U}, y, \underline{v'}.\langle y \rangle\ , \\ \underline{T} := \lambda \underline{U}, y, \underline{v'}.\langle \underline{v'} \rangle\ .$$

   On the other hand, the interpretation of $\forall x\, (\mathrm{st}(x) \to \Phi(x)) \to \forall^{\mathrm{st}} x\, \Phi(x)$ is

$$\exists^{\mathrm{st}} \underline{U'}, S, \underline{T}\, \forall^{\mathrm{st}} x', \underline{U}, \underline{v'}\, \big(\forall y \in Sx'\underline{U}\underline{v'}\ \forall \underline{v} \in \underline{T}x'\underline{U}\underline{v'}\ \forall x \\ (x = y \to \varphi(\underline{U}y, \underline{v}, x)) \to \varphi(\underline{U'}\underline{U}x', \underline{v'}, x')\big)\ ,$$

   and we can take

$$\underline{U'} := \lambda \underline{U}, x'.\underline{U}x'\ , \qquad S := \lambda x', \underline{U}, \underline{v'}.\langle x' \rangle\ , \\ \underline{T} := \lambda x', \underline{U}, \underline{v'}.\langle \underline{v'} \rangle\ .$$

   (ii) $\exists^{\mathrm{st}} x\, \Phi(x) \leftrightarrow \exists x\, (\mathrm{st}(x) \wedge \Phi(x))$.
   The interpretation of $\exists^{\mathrm{st}} x\, \Phi(x) \to \exists x\, (\mathrm{st}(x) \wedge \Phi(x))$ is

$$\exists^{\mathrm{st}} Y, \underline{U'}, \underline{T}\, \forall^{\mathrm{st}} x, \underline{u}, \underline{s}\, \big(\forall \underline{v} \in \underline{T}x\underline{us}\ \varphi(\underline{u}, \underline{v}, x) \to \\ \exists x'\, \forall \underline{v'} \in \underline{s}\, (Yx\underline{u} = x' \wedge \varphi(\underline{U'}x\underline{u}, \underline{v'}, x'))\big)\ ;$$

   so we can take

$$Y := \lambda x, \underline{u}.x\ , \qquad \underline{U'} := \lambda x, \underline{u}.\underline{u}\ , \\ \underline{T} := \lambda x, \underline{u}, \underline{s}.\underline{s}\ .$$

   The interpretation of its converse $\exists x\, (\mathrm{st}(x) \wedge \Phi(x)) \to \exists^{\mathrm{st}} x\, \Phi(x)$ is

$$\exists^{\mathrm{st}} X, \underline{U}, \underline{S}\, \forall^{\mathrm{st}} y, \underline{u'}, \underline{v}\, \big(\forall \underline{s} \in \underline{S}y\underline{u'}\underline{v}\ \exists x'\ \forall \underline{v'} \in \underline{s} \\ (y = x' \wedge \varphi(\underline{u'}, \underline{v'}, x')) \to \varphi(\underline{U}y\underline{u'}, \underline{v}, Xy\underline{u'})\big)\ ,$$

and we can take

$$X := \lambda y, \underline{u}'.y \;, \qquad \underline{U} := \lambda y, \underline{u}'.\underline{u}' \;,$$
$$\underline{S} := \lambda y, \underline{u}', \underline{v}.\langle\langle \underline{v} \rangle\rangle \;.$$

4. *The axioms for the standardness predicate.*

   (i) $\mathrm{st}(x) \wedge x = y \to \mathrm{st}(y)$.
   The interpretation of this axiom is

   $$\exists^{\mathrm{st}} Y' \, \forall^{\mathrm{st}} x' \, (x = x' \wedge x = y \to y = Y'x') \;,$$

   so we can take $Y' := \lambda x'.x'$.

   (ii) $\mathrm{st}(a)$ for all closed terms a.
   We have $(\mathrm{st}(a))^U = \exists^{\mathrm{st}} x \, (a = x)$, so we can take $x := a$.

   (iii) $\mathrm{st}(f) \wedge \mathrm{st}(x) \to \mathrm{st}(fx)$.
   The interpretation of this axiom is

   $$\exists^{\mathrm{st}} Y \, \forall^{\mathrm{st}} f', x' \, (f = f' \wedge x = x' \to fx = Y f' x') \;,$$

   so we can take $Y := \lambda f', x'.f'x'$.

5. *The external induction schema.*

   As in [64], we consider the equivalent external induction *rule*

   $$\mathsf{IR}^{\mathrm{st}} : \quad \frac{\Phi(0) \quad \forall^{\mathrm{st}} n : 0 \, (\Phi(n) \to \Phi(n+1))}{\forall^{\mathrm{st}} n : 0 \, \Phi(n)} \;,$$

   from which the external induction schema is obtained by taking $\Phi(m) := \Psi(0) \wedge \forall^{\mathrm{st}} n : 0 \, (\Psi(n) \to \Psi(n+1)) \to \Psi(m)$.

   So, suppose that $(\Phi(n))^U = \exists^{\mathrm{st}} \underline{x} \, \forall^{\mathrm{st}} \underline{y} \, \varphi(\underline{x}, \underline{y}, n, \underline{a})$, and that we have realisers $\underline{t}_1$, and $\underline{T}_2, \underline{T}_3$ for the premises; i.e.

   $$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall \underline{y} \, \varphi(\underline{t}_1, \underline{y}, 0, \underline{a}) \;,$$

   and

   $$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall n, \underline{x}, \underline{y}' \, (\forall \underline{y} \in \underline{T}_3 n \underline{x} \underline{y}' \, \varphi(\underline{x}, \underline{y}, n, \underline{a}) \to \varphi(\underline{T}_2 n \underline{x}, \underline{y}', n+1, \underline{a})) \;.$$

   By taking $\underline{T}_4 := \lambda n.\mathrm{R} \underline{t}_1 \underline{T}_2 n$, we obtain, by induction for $\vee$-free formulae in E-$\text{HA}^{\omega *}_{\vee}$, that

   $$\text{E-HA}^{\omega *}_{\vee} + \Delta_{\vee} \vdash \forall n, \underline{y} \, \varphi(\underline{T}_4 n, \underline{y}, n, \underline{a}) \;,$$

   which was to be proved.

6. *The principles* $\mathsf{OS}^*_{\vee}, \mathsf{US}^*_{\vee}, \mathsf{NU}, \mathsf{AC}^{\mathrm{st}}, \mathsf{IP}^{\mathrm{st}}_{\forall \vee}$.

   (i) $\mathsf{OS}^*_{\vee} : \quad \forall^{\mathrm{st}} s \, \varphi(s) \to \exists s \, (\forall^{\mathrm{st}} x \, (x \in s) \wedge \varphi(s))$, with φ internal and $\vee$-free.
   This is interpreted as

   $$\exists^{\mathrm{st}} S \, \forall^{\mathrm{st}} s' \, \big(\forall s \in S s' \, \varphi(s) \to \exists s \, (s' \subseteq s \wedge \varphi(s))\big) \;,$$

   and we can take $S := \lambda s'.\langle s' \rangle$.

(ii) $\mathsf{US}^*_\vee$: $\forall s\,(\forall^{\mathrm{st}} x\,(x \in s) \to \varphi(s)) \to \exists^{\mathrm{st}} s\,\varphi(s)$, with φ internal and $\vee$-free.
The interpretation of this axiom schema is

$$\exists^{\mathrm{st}} S\,\forall^{\mathrm{st}} s'\,\big(\forall s\,(s' \subseteq s \to \varphi(s)) \to \varphi(Ss')\big)\ ;$$

so we can take $S := \lambda s'.s'$.

For the principles $\mathsf{NU}, \mathsf{AC}^{\mathrm{st}}, \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$, we can just observe that the premise and the conclusion have identical interpretations, so it is trivial to find realisers for the implication. We do the first as an example.

(iii) NU : $\forall y\,\exists^{\mathrm{st}} x\,\Phi(x,y) \to \exists^{\mathrm{st}} x\,\forall y\,\Phi(x,y)$.
Let $\Phi(x,y)^U := \exists^{\mathrm{st}} \underline{u}\,\forall^{\mathrm{st}} \underline{v}\,\varphi(\underline{u},\underline{v},x,y)$. Both the premise and the conclusion are interpreted as

$$\exists^{\mathrm{st}} x, \underline{u}\,\forall^{\mathrm{st}} \underline{v}\,\forall y\,\varphi(\underline{u},\underline{v},x,y)\ ;$$

so the implication is interpreted as

$$\exists^{\mathrm{st}} X', \underline{U}', \underline{S}\,\forall^{\mathrm{st}} x, \underline{u}, \underline{v}'\,\big(\forall \underline{v} \in \underline{S}x\underline{u}\underline{v}'\,\forall y\,\varphi(\underline{u},\underline{v},x,y) \to \forall y\,\varphi(\underline{U}'x\underline{u}, \underline{v}', X'x\underline{u}, y)\big)\ ,$$

and we can take

$$X' := \lambda x, \underline{u}.x\ , \qquad \underline{U}' := \lambda x, \underline{u}.\underline{u}\ ,$$
$$\underline{S} := \lambda x, \underline{u}, \underline{v}'.\langle \underline{v}' \rangle\ .$$

This concludes the proof. □

As usual, soundness of an interpretation leads to a conservation result.

Corollary 2.8. *The system*

$$\text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$$

is conservative with respect to $\vee$-free formulae of $\text{E-HA}^{\omega *}_\vee$.

Proof. Follows immediately from the previous theorem. □

It may seem odd, at first, that in order to extend the Diller-Nahm interpretation with uniform quantifiers, we would end up adopting a system like $\text{E-HA}^{\omega *}_\vee$, which, due to the lack of the full induction schema, is not even a proper system of arithmetic. The point is that, unlike the situation in Chapter 1, where we started with ordinary Heyting arithmetic, and then introduced new deductive procedures by means of the standardness predicate, here it is the fragment where *all quantifiers are external* that corresponds to the usual intuitionistic arithmetic; and the ordinary quantifiers, with their uniform interpretation, are "new".

In fact, it is through the translation of $\text{N-HA}^{\omega *}$ into $\text{E-HA}^{\omega *}_{\mathrm{st}\vee}$ where $\forall x : \sigma \mapsto \forall^{\mathrm{st}} x : \sigma$ that one retrieves the usual Diller-Nahm interpretation. Notice that, while there are no problems with the U-interpretation of the extensionality axioms with uniform quantifiers, still we could not interpret an axiom like

$$f =_{\sigma\to\tau} g \leftrightarrow \forall^{\mathrm{st}} x : \sigma\; fx =_\tau gx\ .$$

An external extensionality *rule*, though, should be acceptable.

2.3 Characterisation and properties

We are now in the position to prove that the principles examined above are characteristic of uniform Diller-Nahm. So, let

$$\mathrm{H} := \text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee} \ .$$

Theorem 2.9 (Characterisation of uniform Diller-Nahm)**.** *Let Φ be a formula in the language of* $\text{E-HA}^{\omega *}_{\mathrm{st}\vee}$*.*

(a) $\mathrm{H} \vdash \Phi \leftrightarrow \Phi^U$.

(b) If for all formulae Ψ of $\mathcal{L}(\text{E-HA}^{\omega *}_{\mathrm{st}\vee})$, with $\Psi^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \psi(\underline{x}, \underline{y})$,

$$\mathrm{H} + \Phi \vdash \Psi$$

implies that there exist closed terms $\underline{t}$ such that

$$\text{E-HA}^{\omega *}_{\vee} \vdash \forall \underline{y}\, \psi(\underline{t}, \underline{y})$$

holds, then $\mathrm{H} \vdash \Phi$ *.*

Proof. We prove (a) by induction on the logical structure of Φ. For $\Phi \equiv \varphi$ internal atomic, obviously $\mathrm{H} \vdash \varphi \leftrightarrow \varphi^U$.

Let $\Phi \equiv \mathrm{st}(x)$. If x is standard, it follows that $\exists^{\mathrm{st}} y\, (x = y)$, by taking $y := x$. Conversely, if $\exists^{\mathrm{st}} y\, (x = y)$, by the first axiom for the standardness predicate it follows that x is standard. Hence,

$$\mathrm{H} \vdash \mathrm{st}(x) \leftrightarrow \exists^{\mathrm{st}} y\, (x = y) \ .$$

For the induction hypothesis, using an appropriate embedding of tuples of types into higher types, and a compatible coding of tuples of terms [59, 1.6.17], we can assume, given formulae Φ and Ψ, that there exist internal, $\vee$-free formulae φ, ψ such that

$$\mathrm{H} \vdash \Phi(x) \leftrightarrow \exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x, y) \ ,$$
$$\mathrm{H} \vdash \Psi(x) \leftrightarrow \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \psi(u, v) \ .$$

(i) For $\wedge$, by intuitionistic logic,

$$\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x, y) \wedge \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \psi(u, v)$$

is equivalent to

$$\exists^{\mathrm{st}} x, u\, \forall^{\mathrm{st}} y, v\, (\varphi(x, y) \wedge \psi(u, v)) \ .$$

(ii) For $\vee$,

$$\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x, y) \vee \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \psi(u, v)$$

is equivalent in H to

$$\exists^{\mathrm{st}} z : 0\, (z = 0 \to \exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x, y) \wedge \neg\, z = 0 \to \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \psi(u, v)) \ .$$

By $\mathsf{IP}^{\mathrm{st}}_{\forall\vee}$, this is equivalent to

$$\exists^{\mathrm{st}} z : 0\, \big(\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, (z = 0 \to \varphi(x, y)) \wedge \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, (\neg\, z = 0 \to \psi(u, v))\big) \ ,$$

and we are back to the case of conjunction.

(iii) For $\to$, we proceed as with Diller-Nahm implication. By intuitionistic logic and the principle $\mathsf{IP}^{\mathrm{st}}_{\forall\vee}$,

$$\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x,y) \to \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \psi(u,v)$$

is equivalent to

$$\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, (\forall^{\mathrm{st}} y\, \varphi(x,y) \to \psi(u,v))\ .$$

Now, adapting Proposition 1.41, we see that $\text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{US}^*_\vee \vdash \mathsf{HGMP}^{\mathrm{st}}_\vee$, so this is equivalent to

$$\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} u\, \forall^{\mathrm{st}} v\, \exists^{\mathrm{st}} s\, (\forall y \in s\, \varphi(x,y) \to \psi(u,v))\ .$$

Two applications of $\mathsf{AC}^{\mathrm{st}}$ then lead to

$$\exists^{\mathrm{st}} U, S\, \forall^{\mathrm{st}} x, v\, (\forall y \in Sxv\, \varphi(x,y) \to \psi(Ux,v))\ .$$

(iv) For $\exists z$, adapting Proposition 1.32, we see that $\text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee \vdash \mathsf{I}_\vee$; therefore

$$\exists z\, \exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x,y,z)$$

is equivalent to

$$\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} s\, \exists z\, \forall y \in s\, \varphi(x,y,z)\ .$$

(v) For $\forall z$, we use that by NU

$$\forall z\, \exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x,y,z)$$

is equivalent to

$$\exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \forall z\, \varphi(x,y,z)\ .$$

(vi) For $\exists^{\mathrm{st}} z$, nothing really needs to be done.

(vii) For $\forall^{\mathrm{st}} z$, we just use $\mathsf{AC}^{\mathrm{st}}$ once in order to obtain that

$$\forall^{\mathrm{st}} z\, \exists^{\mathrm{st}} x\, \forall^{\mathrm{st}} y\, \varphi(x,y,z)$$

is equivalent to

$$\exists^{\mathrm{st}} X\, \forall^{\mathrm{st}} y, z\, \varphi(Xz,y,z)\ .$$

This proves item (a). For (b), suppose Φ satisfies the condition, and that $\Phi^U = \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x},\underline{y})$. Then, from

$$\mathrm{H} + \Phi \vdash \Phi$$

it follows that there exist closed terms $\underline{t}$ such that

$$\text{E-HA}^{\omega *}_{\vee} \vdash \forall \underline{y}\, \varphi(\underline{t},\underline{y})\ .$$

From this, we obtain $\text{E-HA}^{\omega *}_{\mathrm{st}\vee} \vdash \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{t},\underline{y})$, whence

$$\text{E-HA}^{\omega *}_{\mathrm{st}\vee} \vdash \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x},\underline{y})\ ;$$

so $\mathrm{H} \vdash \exists^{\mathrm{st}} \underline{x}\, \forall^{\mathrm{st}} \underline{y}\, \varphi(\underline{x},\underline{y})$ as well. But then, by the equivalence of (a),

$$\mathrm{H} \vdash \Phi$$

which was to be proved. $\square$

We now show how the uniform Diller-Nahm interpretation may be used to extract programs from proofs, and eliminate instances of its characteristic principles.

Theorem 2.10 (Program extraction by the U-interpretation)**.** *Let $\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \varphi(x,y)$ be a sentence of* $\text{E-HA}^{\omega *}_{\mathrm{st}\vee}$*, with φ internal and $\vee$-free, and let $\Delta_\vee$ be a set of internal, $\vee$-free sentences. If*

$$\text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee} + \Delta_\vee \vdash \forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \varphi(x,y)\ ,$$

then from the proof we can extract a closed term T such that

$$\text{E-HA}^{\omega *}_{\vee} + \Delta_\vee \vdash \forall x\, \varphi(x, Tx)\ .$$

Proof. The U-translation of $\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \varphi(x,y)$ is

$$\exists^{\mathrm{st}} f\, \forall^{\mathrm{st}} x\, \varphi(x, fx)\ ,$$

so the thesis immediately follows from the soundness theorem. □

Finally, we derive a few properties of the system $\text{E-HA}^{\omega *}_{\mathrm{st}\vee}$, which follow from the properties of the uniform Diller-Nahm interpretation.

Proposition 2.11. *The system* $\mathrm{H} := \text{E-HA}^{\omega *}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$ *is closed under the restricted transfer rules*

$$\mathsf{TR}_{\forall\vee} : \quad \frac{\forall^{\mathrm{st}} x : \sigma\, \varphi(x)}{\forall x : \sigma\, \varphi(x)}\ ,$$

$$\mathsf{TR}_{\exists\vee} : \quad \frac{\exists x : \sigma\, \varphi(x)}{\exists^{\mathrm{st}} x : \sigma\, \varphi(x)}\ ,$$

where φ ranges over internal $\vee$-free formulae.

Proof. This is an adaptation of [64, Proposition 5.12]. Suppose

$$\mathrm{H} \vdash \forall^{\mathrm{st}} x\, \varphi(x)\ .$$

By the soundness theorem, it follows that

$$\text{E-HA}^{\omega *}_{\vee} \vdash \forall x\, \varphi(x)\ ,$$

which, since H is an extension of $\text{E-HA}^{\omega *}_{\vee}$, implies $\mathrm{H} \vdash \forall^{\mathrm{st}} x\, \varphi(x)$.

Now, suppose

$$\mathrm{H} \vdash \exists x\, \varphi(x)\ ;$$

by conservativity, this implies $\text{E-HA}^{\omega *}_{\vee} \vdash \exists x\, \varphi(x)$. Being a subsystem of $\text{E-HA}^{\omega *}$, $\text{E-HA}^{\omega *}_{\vee}$ inherits the existence property; so we can find a closed term t such that

$$\text{E-HA}^{\omega *}_{\vee} \vdash \varphi(t)\ .$$

Since t is provably standard in H, this implies $\mathrm{H} \vdash \exists^{\mathrm{st}} x\, \varphi(x)$. □

Proposition 2.12. *The system* $\mathrm{H} := \text{E-HA}^{\omega*}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$ *has the following form of the existence property: if*

$$\mathrm{H} \vdash \exists^{\mathrm{st}} x\, \Phi(x) \ ,$$

then there exists a closed term t such that $\mathrm{H} \vdash \Phi(t)$.

Proof. Let $\Phi(x)^U = \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \varphi(x, \underline{u}, \underline{v})$. By the characterisation theorem, H proves that Φ is equivalent to its U-translation; so, if $\mathrm{H} \vdash \exists^{\mathrm{st}} x\, \Phi(x)$,

$$\mathrm{H} \vdash \exists^{\mathrm{st}} x, \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \varphi(x, \underline{u}, \underline{v}) \ .$$

By soundness of uniform Diller-Nahm, we can extract closed terms $t, \underline{T}$ such that

$$\text{E-HA}^{\omega*}_{\vee} \vdash \forall \underline{v}\, \varphi(t, \underline{T}, \underline{v}) \ ;$$

which, by conservativity, and weakening the quantifier, implies

$$\mathrm{H} \vdash \forall^{\mathrm{st}} \underline{v}\, \varphi(t, \underline{T}, \underline{v}) \ .$$

Since the terms in $\underline{T}$ are provably standard in H, we obtain

$$\mathrm{H} \vdash \exists^{\mathrm{st}} \underline{u}\, \forall^{\mathrm{st}} \underline{v}\, \varphi(t, \underline{u}, \underline{v}) \ ,$$

which, again by the characterisation theorem, implies $\mathrm{H} \vdash \Phi(t)$. □

Corollary 2.13. *The system* $\mathrm{H} := \text{E-HA}^{\omega*}_{\mathrm{st}\vee} + \mathsf{OS}^*_\vee + \mathsf{US}^*_\vee + \mathsf{NU} + \mathsf{AC}^{\mathrm{st}} + \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$ *has the disjunction property.*

Proof. Follows from the validity of $\Phi \vee \Psi \leftrightarrow \exists^{\mathrm{st}} z : 0\, (z = 0 \to \Phi \wedge \neg\, z = 0 \to \Psi)$ in H, and the previous proposition. □

This is the crux of an approach to nonstandard arithmetic starting from the Diller-Nahm interpretation. We have seen that the OS_0 principle - the least that we would want from a proper system of nonstandard arithmetic - implies $\mathsf{LLPO}^{\mathrm{st}}_0$; yet, by Corollary 1.50, this means giving up the disjunction property, and, *a fortiori*, the existence property.

So, in a way, the aforementioned system H, with its restricted induction schema and overspill and underspill principles, is as close as one can get to actual nonstandard arithmetic, while retaining the full computational strength of the intuitionistic quantifiers. To proceed farther, one needs to relax this requirement; the next best thing is asking for a *sequence* of potential realisers, of which at least one has to work, for each existential statement, instead of a single, actual realiser - what we called *herbrandisation*.

2.4 The nonstandard Dialectica interpretation

The introduction of the nonstandard Dialectica interpretation in [64] had nonstandard analysis as its main motivation: the benchmark to meet was eliminating overspill and underspill from proofs, retrieving what computational content they may have.

Almost all the results from this section can be found there. But we tried, in this chapter, to describe a different route towards the same result: one where we start with the Dialectica interpretation, and progressively apply small "patches", fixing whichever shortcomings might arise. We did keep the nonstandard nomenclature, for there was no point in pretending we had different goals; but it seems quite noteworthy, to us, that one could come, in principle, to this interpretation without ever thinking of nonstandard arithmetic.

So Dialectica requires decidability: we may be unhappy with that, and turn to Diller-Nahm. Yet we could speed up our proof mining, if only we had two kinds of quantifiers; which leads us to uniform Diller-Nahm. Then, we notice that the system we obtain - E-HA$^{\omega*}_{\mathrm{st}\vee}$ and characteristic principles - is just one connective away from being a system of intuitionistic arithmetic. Also, fixing that may require that we weaken the existence property.

We could come up with the following requirements.

1. Realisers will have to be sequences, so the translated formulae will be of the form $\exists^{\mathrm{st}}\underline{s}\,\forall^{\mathrm{st}}\underline{y}\,\varphi(\underline{s},\underline{y})$, where all the variables in $\underline{s}$ are of sequence type.

2. Also, the more, the better: if one $x \in s$ is an actual realiser, then any $s' \supseteq s$ is also good. So φ will also have to be *upwards closed* in the first argument; and we will need the monotonic *sequence* application as our habitual application.

3. Accordingly, the characteristic principles will also receive the Herbrand treatment: goodbye, $\mathsf{NU}, \mathsf{AC}^{\mathrm{st}}, \mathsf{IP}^{\mathrm{st}}_{\forall\vee}$; welcome, $\mathsf{NCR}, \mathsf{HAC}^{\mathrm{st}}$, and

$$\mathsf{HIP}^{\mathrm{st}}_{\forall}: \quad (\forall^{\mathrm{st}}x:\sigma\,\varphi(x)\to\exists^{\mathrm{st}}y:\tau\,\Psi(y))\to\exists^{\mathrm{st}}t:\tau^*\,(\forall^{\mathrm{st}}x:\sigma\,\varphi(x)\to\exists y\in t\,\Psi(y))\,.$$

4. Finally, we want to replace "internal and $\vee$-free" with just "internal", where needed, and be able to interpret IA, OS^* and US^*.

In our alternate history, this leads us straight to the following definition.

Definition 2.14. To every formula $\Phi(\underline{a})$ of $\mathcal{L}(\text{E-HA}^{\omega*}_{\mathrm{st}})$, with free variables $\underline{a}$, we associate inductively its *nonstandard Dialectica* translation $\Phi(\underline{a})^{D_{\mathrm{st}}} = \exists^{\mathrm{st}}\underline{s}\,\forall^{\mathrm{st}}\underline{y}\,\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},\underline{a})$, where $\varphi_{D_{\mathrm{st}}}$ is internal.

▷ $\varphi(\underline{a})^{D_{\mathrm{st}}} := \varphi_{D_{\mathrm{st}}}(\underline{a}) := \varphi(\underline{a})$, for φ internal atomic;

▷ $\mathrm{st}_\sigma(x)^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}s:\sigma^*\,(x\in s)$.

Let $\Phi(\underline{a})^{D_{\mathrm{st}}} = \exists^{\mathrm{st}}\underline{s}\,\forall^{\mathrm{st}}\underline{y}\,\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},\underline{a})$, $\Psi(\underline{b})^{D_{\mathrm{st}}} = \exists^{\mathrm{st}}\underline{t}\,\forall^{\mathrm{st}}\underline{v}\,\psi_{D_{\mathrm{st}}}(\underline{t},\underline{v},\underline{b})$:

▷ $(\Phi(\underline{a})\wedge\Psi(\underline{b}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{s},\underline{t}\,\forall^{\mathrm{st}}\underline{y},\underline{v}\,\big(\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},\underline{a})\wedge\psi_{D_{\mathrm{st}}}(\underline{t},\underline{v},\underline{b})\big)$;

▷ $(\Phi(\underline{a})\vee\Psi(\underline{b}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{s},\underline{t}\,\forall^{\mathrm{st}}\underline{y},\underline{v}\,\big(\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},\underline{a})\vee\psi_{D_{\mathrm{st}}}(\underline{t},\underline{v},\underline{b})\big)$;

▷ $(\Phi(\underline{a})\to\Psi(\underline{b}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{T},\underline{Y}\,\forall^{\mathrm{st}}\underline{s},\underline{v}\,\big(\forall\underline{y}\in\underline{Y}[\underline{s},\underline{v}]\,\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},\underline{a})\to\psi_{D_{\mathrm{st}}}(\underline{T}[\underline{s}],\underline{v},\underline{b})\big)$;

▷ $(\exists z\,\Phi(z,\underline{a}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{s}\,\forall^{\mathrm{st}}\underline{t}\,\exists z\,\forall\underline{y}\in\underline{t}\,\varphi_{D_{\mathrm{st}}}(\underline{x},\underline{y},z,\underline{a})$;

▷ $(\forall z\,\Phi(z,\underline{a}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{s}\,\forall^{\mathrm{st}}\underline{y}\,\forall z\,\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},z,\underline{a})$;

▷ $(\exists^{\mathrm{st}}z\,\Phi(z,\underline{a}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}u,\underline{s}\,\forall^{\mathrm{st}}\underline{t}\,\exists^{\mathrm{st}}z\in u\,\forall\underline{y}\in\underline{t}\,\varphi_{D_{\mathrm{st}}}(\underline{s},\underline{y},z,\underline{a})$;

▷ $(\forall^{\mathrm{st}}z\,\Phi(z,\underline{a}))^{D_{\mathrm{st}}} := \exists^{\mathrm{st}}\underline{S}\,\forall^{\mathrm{st}}\underline{y},z\,\varphi_{D_{\mathrm{st}}}(\underline{S}[z],\underline{y},z,\underline{a})$.

The D_{st} translation is very similar to uniform Diller-Nahm in most of its clauses, the biggest differences - unsurprisingly, since we are discharging some of their computational content - being disjunction and the $\exists^{\text{st}} z$ quantifier. The latter, in particular, leads to the interpretation *not* being idempotent: $(\exists^{\text{st}} z\, \Phi(z))^{D_{\text{st}}}$ is not, in general, the same as $((\exists^{\text{st}} z\, \Phi(z))^{D_{\text{st}}})^{D_{\text{st}}}$.

However, it is true that formulae of the form $\forall^{\text{st}} \underline{x}\, \varphi(\underline{x})$, with φ internal, are left unchanged by the interpretation.

It can be easily shown that our second requirement is satisfied.

Proposition 2.15. *Let $\Phi(\underline{a})$ be a formula of $\mathcal{L}(\text{E-HA}^{\omega *}_{\text{st}})$, $\Phi(\underline{a})^{D_{\text{st}}} = \exists^{\text{st}} \underline{s}\, \forall^{\text{st}} \underline{y}\, \varphi(\underline{s}, \underline{y}, \underline{a})$. Then* E-HA$^{\omega *}$ *proves that φ is upwards closed in $\underline{s}$:*

$$\text{E-HA}^{\omega *} \vdash \varphi(\underline{s}, \underline{y}, \underline{a}) \wedge \underline{s} \subseteq \underline{s}' \to \varphi(\underline{s}', \underline{y}, \underline{a}) \;.$$

Proof. By induction on the logical structure of $\Phi(\underline{a})$, using Lemma 1.15 in the clauses for $\to$ and $\forall^{\text{st}} z$. □

We now state the soundness theorem for the D_{st} interpretation, without a complete proof, which would be unnecessary.

Theorem 2.16 (Soundness of the nonstandard Dialectica interpretation)**.** *Suppose*

$$\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}^* + \mathsf{US}^* + \mathsf{NCR} + \mathsf{HAC}^{\text{st}} + \mathsf{HIP}^{\text{st}}_{\forall} + \Delta_{\text{int}} \vdash \Phi(\underline{a}) \;,$$

where Δ_{int} is a set of internal sentences. Let $\Phi(\underline{a})^{D_{\text{st}}} = \exists^{\text{st}} \underline{s}\, \forall^{\text{st}} \underline{y}\, \varphi_{D_{\text{st}}}(\underline{s}, \underline{y}, \underline{a})$. Then from the proof we can extract a tuple of closed terms $\underline{t}$ such that

$$\text{E-HA}^{\omega *} + \Delta_{\text{int}} \vdash \forall \underline{y}\, \varphi_{D_{\text{st}}}(\underline{t}, \underline{y}, \underline{a}) \;.$$

Proof. As usual, the proof proceeds by induction on the length of the derivation. We only analyse one of the defining axioms of the external quantifiers, as an illustrative example of the specificities of D_{st}, and the principles OS^* and US^*, which were not taken into consideration in the original article; and refer to [64, Theorem 5.5] for the rest.

Consider $\exists^{\text{st}} x\, \Phi(x) \leftrightarrow \exists x\, (\text{st}(x) \wedge \Phi(x))$, and let $\Phi(x)^{D_{\text{st}}} = \exists^{\text{st}} \underline{s}\, \forall^{\text{st}} \underline{y}\, \varphi(\underline{s}, \underline{y}, x)$. We have

$$(\exists^{\text{st}} x\, \Phi(x))^{D_{\text{st}}} = \exists^{\text{st}} u, \underline{s}\, \forall^{\text{st}} \underline{t}\, \exists x \in u\, \forall \underline{y} \in \underline{t}\, \varphi(\underline{s}, \underline{y}, x) \;,$$

and

$$\big(\exists x\, (\text{st}(x) \wedge \Phi(x))\big)^{D_{\text{st}}} = \exists^{\text{st}} u, \underline{s}\, \forall^{\text{st}} \underline{t}\, \exists x\, \forall \underline{y} \in \underline{t}\, (x \in u \wedge \varphi(\underline{s}, \underline{y}, x)) \;.$$

Then the interpretation of $\exists^{\text{st}} x\, \Phi(x) \to \exists x\, (\text{st}(x) \wedge \Phi(x))$ is

$$\begin{aligned} \exists^{\text{st}} U', \underline{S}', \underline{T}\, \forall^{\text{st}} \underline{s}, u, \underline{t}'\, \big(\forall \underline{t} \in \underline{T}[\underline{s}, u, \underline{t}']\, \exists x \in u\, \forall \underline{y} \in \underline{t}\, \varphi(\underline{s}, \underline{y}, \underline{x}) \to \\ \exists x\, \forall \underline{y} \in \underline{t}'\, (x \in U'[\underline{s}, u] \wedge \varphi(\underline{S}'[\underline{s}, u], \underline{y}, x)\big) \;, \end{aligned}$$

and we can take

$$\begin{aligned} &U' := \Lambda \underline{s}, u.u \;, \qquad \underline{S}' := \Lambda \underline{s}, u.\underline{s} \;, \\ &\underline{T} := \Lambda \underline{s}, u, \underline{t}'.\langle \underline{t}' \rangle \;. \end{aligned}$$

The converse $\exists x\,(\mathrm{st}(x) \wedge \Phi(x)) \to \exists^{\mathrm{st}} x\, \Phi(x)$, on the other hand, is interpreted as

$$\exists^{\mathrm{st}} U', \underline{S}', \underline{T}\, \forall^{\mathrm{st}} \underline{s}, u, \underline{t}' \left(\forall \underline{t} \in \underline{T}[\underline{s}, u, \underline{t}']\, \exists x\, \forall \underline{y} \in \underline{t}' \right.$$
$$\left. (x \in U'[\underline{s}, u] \wedge \varphi(\underline{S}'[\underline{s}, u], \underline{y}, x) \to \exists x \in u\, \forall \underline{y} \in \underline{t}\, \varphi(\underline{s}, \underline{y}, \underline{x})\right) ,$$

and the *same* realisers will work as with the first implication.

Now, consider OS^* : $\forall^{\mathrm{st}} s\, \varphi(s) \to \exists s\, (\forall^{\mathrm{st}} x\, (x \in s) \wedge \varphi(s))$, with φ internal. Its interpretation is

$$\exists^{\mathrm{st}} S\, \forall^{\mathrm{st}} s' \left(\forall s \in S[s']\, \varphi(s) \to \exists s\, (s' \subseteq s \wedge \varphi(s))\right) ,$$

and we can take $S := \Lambda s'.\langle s' \rangle$.

Finally, US^* : $\forall s\, (\forall^{\mathrm{st}} x\, (x \in s) \to \varphi(s)) \to \exists^{\mathrm{st}} s\, \varphi(s)$ is interpreted as

$$\exists^{\mathrm{st}} T\, \forall^{\mathrm{st}} s'' \left(\forall s\, \exists s' \in s''\, (s' \subseteq s \to \varphi(s)) \to \exists t \in T[s'']\, \varphi(t)\right);$$

since $\forall s\, \exists s' \in s''\, (s' \subseteq s \to \varphi(s))$ implies $\varphi(s''_0 \cdot \ldots \cdot s''_{|s''|-1})$, we can take

$$T := \Lambda s''.(s''_0 \cdot \ldots \cdot s''_{|s''|-1}) . \qquad \square$$

Corollary 2.17. *The system*

$$\mathrm{H} := \text{E-HA}^{\omega*}_{\mathrm{st}} + \mathsf{OS}^* + \mathsf{US}^* + \mathsf{NCR} + \mathsf{HAC}^{\mathrm{st}} + \mathsf{HIP}^{\mathrm{st}}_{\forall}$$

is a conservative extension of E-HA$^{\omega*}$*, hence of* E-HA$^{\omega}$*.*

So the magic happened: we have attained a system which, including sequence overspill and underspill, we are not ashamed to call nonstandard arithmetic; and, by the last conservation result, we have a full "transfer theorem".

Then, we have the analogue of Theorem 2.9, with the herbrandised principles.

Theorem 2.18 (Characterisation of nonstandard Dialectica)**.** *Let Φ be a formula in the language of* E-HA$^{\omega*}_{\mathrm{st}}$*.*

(a) $\mathrm{H} \vdash \Phi \leftrightarrow \Phi^{D_{\mathrm{st}}}$.

(b) If for all formulae Ψ of $\mathcal{L}$(E-HA$^{\omega*}_{\mathrm{st}}$), with $\Psi^{D_{\mathrm{st}}} = \exists^{\mathrm{st}} \underline{s}\, \forall^{\mathrm{st}} \underline{y}\, \psi(\underline{s}, \underline{y})$,

$$\mathrm{H} + \Phi \vdash \Psi$$

implies that there exist closed terms $\underline{t}$ such that

$$\text{E-HA}^{\omega*} \vdash \forall \underline{y}\, \psi(\underline{t}, \underline{y})$$

holds, then $\mathrm{H} \vdash \Phi$ *.*

Proof. For item (a), again one should proceed by induction on the logical structure of Φ. We only examine the cases that are distinct enough from uniform Diller-Nahm, namely the standardness predicate, $\vee$ and $\exists^{\mathrm{st}} z$, and refer to [64, Theorem 5.8] for the rest.

For $\mathrm{st}(x)$, if x is standard, $\exists^{\mathrm{st}} s\, (x \in s)$ follows with $s := \langle x \rangle$. If $\exists^{\mathrm{st}} s\, (x \in s)$, then $\mathrm{st}(x)$ follows by Lemma 1.21.(c).

Now, given Φ and Ψ, suppose that there are internal φ, ψ such that

$$\mathrm{H} \vdash \Phi(x) \leftrightarrow \exists^{\mathrm{st}} s\, \forall^{\mathrm{st}} y\, \varphi(s, y) ,$$
$$\mathrm{H} \vdash \Psi(x) \leftrightarrow \exists^{\mathrm{st}} t\, \forall^{\mathrm{st}} v\, \psi(t, v) .$$

(i) For $\vee$, observe that

$$\exists^{\mathrm{st}} s\, \forall^{\mathrm{st}} y\, \varphi(s,y) \vee \exists^{\mathrm{st}} t\, \forall^{\mathrm{st}} v\, \psi(t,v)$$

is intuitionistically equivalent to

$$\exists^{\mathrm{st}} s,t\, (\forall^{\mathrm{st}} y\, \varphi(s,y) \vee \forall^{\mathrm{st}} v\, \psi(t,v)) \ .$$

By 1.36, $\mathsf{LLPO}^{\mathrm{st}}$ holds in H, so this in turn is equivalent to

$$\exists^{\mathrm{st}} s,t\, \forall^{\mathrm{st}} y,v\, (\varphi(s,y) \vee \psi(t,v)) \ .$$

(ii) For $\exists^{\mathrm{st}} z$, we use that

$$\exists^{\mathrm{st}} z,s\, \forall^{\mathrm{st}} y\, \varphi(s,y,z)$$

is equivalent, by Lemma 1.21, to

$$\exists^{\mathrm{st}} u,s\, \exists z \in u\, \forall^{\mathrm{st}} y\, \varphi(s,y,z) \ .$$

By idealisation - a principle equivalent to sequence overspill - this is the same as

$$\exists^{\mathrm{st}} u,s\, \forall^{\mathrm{st}} t\, \exists z \in u\, \forall y \in t\, \varphi(s,y,z) \ .$$

The proof of item (b) is identical to that of Theorem 2.9. $\square$

With the characterisation theorem, we can prove that H is closed under both transfer rules.

Proposition 2.19. *The system* $\text{E-HA}^{\omega*}_{\mathrm{st}} + \mathsf{OS}^* + \mathsf{US}^* + \mathsf{NCR} + \mathsf{HAC}^{\mathrm{st}} + \mathsf{HIP}^{\mathrm{st}}_{\forall}$ *is closed under the rules* $\mathsf{TR}_{\forall}$ *and* $\mathsf{TR}_{\exists}$.

Proof. See [64, Proposition 5.12]. $\square$

By Corollary 1.50, we already know that H does not have the disjunction property, nor the existence property. Still, we are able to extract programs from proofs, albeit in a herbrandised fashion.

Theorem 2.20 (Program extraction by the D_{st}-interpretation)**.** *Let* $\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \varphi(x,y)$ *be a sentence of* $\text{E-HA}^{\omega*}_{\mathrm{st}}$*, with* φ *internal, and let* Δ_{int} *be a set of internal sentences. If*

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

then from the proof we can extract a closed term T *such that*

$$\text{E-HA}^{\omega*} + \Delta_{\mathrm{int}} \vdash \forall x\, \exists y \in Tx\, \varphi(x,y) \ .$$

Proof. The D_{st}-translation of $\forall^{\mathrm{st}} x\, \exists^{\mathrm{st}} y\, \varphi(x,y)$ is

$$\exists^{\mathrm{st}} f\, \forall^{\mathrm{st}} x\, \exists y \in f[x]\, \varphi(x,y) \ .$$

By the soundness theorem, we can extract from the proof a closed term S such that

$$\text{E-HA}^{\omega*} + \Delta_{\mathrm{int}} \vdash \forall x\, \exists y \in S[x]\, \varphi(x,y) \ ;$$

so it suffices to pick $T := \lambda x.S[x]$. $\square$

Classical models of nonstandard analysis were meant to provide a calculus of infinitesimals, so as to lower the logical complexity of theorems of classical analysis, and make their proofs more smooth and readable. Perhaps the most fascinating feature of this system of *constructive* nonstandard arithmetic is that not only it does that - with overspill inducing a version of LLPO, and underspill one of Markov's principle, it also allows for a share of *classical* modes of reasoning, which were previously unavailable, with the guarantee that they can be eliminated at the end.

Example 2.21. Many theorems of classical analysis hold constructively only in an *approximate* form; so while the former may state that a certain function f has a zero in a certain interval, the latter will state that f gets *arbitrarily close* to zero in that same interval. Plenty of examples can be found in [8].

So suppose we were able to formalise such a proof in Heyting arithmetic in all finite types; that is, making an informal use of syntax,

$$\text{E-HA}^{\omega *} + \Delta_{\text{int}} \vdash \forall n \, \exists x \, |f(x)| < \frac{1}{n} \; ,$$

where Δ_{int} are the hypotheses of the theorem. Then

$$\text{E-HA}^{\omega *}_{\text{st}} + \Delta_{\text{int}} \vdash \forall^{\text{st}} s \, \exists x \, \forall n \in s \, |f(x)| < \frac{1}{n} \; ,$$

and, utilising the fact that OS^* is equivalent to idealisation,

$$\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}^* + \Delta_{\text{int}} \vdash \exists x \, \forall^{\text{st}} n \, |f(x)| < \frac{1}{n} \; ,$$

which, with the notation of Example 1.37, we can write

$$\text{E-HA}^{\omega *}_{\text{st}} + \mathsf{OS}^* + \Delta_{\text{int}} \vdash \exists x \, |f(x)| \simeq 0 \; .$$

Therefore, at the price of substituting $\simeq$ for $=$, we can retrieve the theorem in a pseudo-classical form, the function having an "ideal" zero; in our proofs, we can

> *"work directly with the ideal objects, and* [...] *on this basis develop a calculus with constructive content which is intermediate between constructive and classical analysis".* [46, p.235]

To what extent can these principles simplify proofs of constructive analysis? In a series of articles [45, 46, 47, 50], Palmgren provided constructive nonstandard proofs, very close to classical ones, for several fundamental theorems of analysis, up to the Implicit Function theorem, by working in the "filter topos" $\mathcal{N}$ - which will be the focus of Chapter 3. As we will see, under few metatheoretical assumptions, first order logic in $\mathcal{N}$ corresponds to D_{st} logic, which is encouraging.

As far as we know, Palmgren did not know of all our characteristic principles that they hold in $\mathcal{N}$; whether they provide further simplification, we would like to know, but did not have time to investigate.

With this, we can wrap up the proof-theoretic part of this thesis. The original models of nonstandard analysis exploited the dangerous nonconstructive nonprincipal *ultrafilters*; in the next chapter, we will venture into topos theory, and try to figure out what our functional interpretations have to do with humble, constructive, ordinary filters of sets.

Chapter 3

Categorical models

In this chapter:

▷ We provide a concise introduction to categorical logic, up to the interpretation of first order logic in a Heyting category, and to the inductive definition of *sheaf semantics* in a Grothendieck topos. - *Section 3.1*

▷ We describe the *filter construction*, building from a category $\mathbf{C}$ a new category $\mathfrak{F}\mathbf{C}$, all the while preserving enough of the original structure; we then define two categories of sheaves over $\mathfrak{F}\mathbf{Set}$, the *filter topoi* $\mathcal{U}$ and $\mathcal{N}$, and give a topos-theoretic characterisation of their relation. - *Section 3.2*

▷ We state and prove a theorem by Moerdijk, linking the internal first order logic of $\mathcal{N}$ to "external" semantics, together with an adaptation for $\mathcal{U}$; these imply a full transfer theorem for $\mathcal{N}$, and a restricted one for $\mathcal{U}$. - *Section 3.3*

▷ We prove that, under few metatheoretical assumptions, the characteristic principles of nonstandard Dialectica hold in $\mathcal{N}$, and those of uniform Diller-Nahm hold in $\mathcal{U}$. - *Section 3.4*

▷ Finally, we survey an entirely different class of elementary, non-Grothendieck topoi that have been considered as models of functional interpretations, including Diller-Nahm and nonstandard Dialectica, and leave a few questions open about their relation to the filter topoi. - *Section 3.5*

3.1 Logic inside a category

In [36], Moerdijk presented a constructive model of nonstandard arithmetic with a full transfer theorem. Although, in principle, the first order part of the model could be described through an explicit forcing relation, he worked in the general context of *Grothendieck topoi* and *Kripke-Joyal semantics*. We will follow the same approach, which allows us to use known results from topos theory in order to prove properties of this model.

We will take the basic notions of category theory - simple limits and colimits, adjoint functors, etc. - for granted; the "Categorical Preliminaries " section of [34] should be sufficient for our needs, and we will use the same book as the standard reference for logical aspects of topoi. We will require *all* categories to be *small*: they will have a "set" of objects, and, for all pairs of objects X, Y, a "set" of morphisms from X to Y - whatever the metatheoretical notion of set is.

Topos theory is an extremely multifaceted subject, as suggested by a list of thirteen possible descriptions of what a topos is, that opens Johnstone's [24], the ongoing *summa* of the field. The one that is perhaps most relevant to a logician is number (iii),

> "A topos is (the embodiment of) an intuitionistic higher-order theory".

One can formalise basically all of mathematics in intuitionistic higher order logic, which justifies the *mantra* that topoi are "universes in which to do mathematics".

Surprisingly, all the structure needed to support a rich enough logic follows from a short list of requirements, appearing in Lawvere's definition of an *elementary topos.*

Definition 3.1. An *elementary topos* is a category $\mathcal{E}$ which

1. is finitely complete,
2. is cartesian closed, and
3. has a subobject classifier.

> **Notation.** We write $\mathrm{Hom}_{\mathbf{C}}(X, Y)$ for the (external) set of morphisms from the object X to the object Y of a category $\mathbf{C}$.
>
> For cartesian closed categories, we write Y^X, or $X \to Y$ when there is no confusion, for the *internal* Hom, the object of morphisms from X to Y.

From these, one can progressively derive that an elementary topos supports more and more expressive logical theories. Recall that the *subobject poset* $\mathrm{Sub}(X)$ of an object X in a category $\mathbf{C}$ is the poset reflection of the set of all monomorphisms $A \rightarrowtail X$ in $\mathbf{C}$ with codomain X, preordered with the relation

> $A \leq B$ if and only if there exists a monomorphism $A \rightarrowtail B$ in $\mathbf{C}$.

Definition 3.2. A finitely complete category $\mathbf{C}$ is *regular* if it has co-equalisers of kernel pairs, and they are stable under pullback.

A regular category $\mathbf{C}$ is *coherent* if, for all objects X of $\mathbf{C}$, the poset $\mathrm{Sub}(X)$ of subobjects of X has all finite unions, and, for all morphisms $f : Y \to X$, these are stable under the change of base functor $f^* : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$.

A coherent category $\mathbf{H}$ is a *Heyting category* if, for all morphisms $f : Y \to X$, the change of base functor $f^* : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ has a right adjoint $\forall_f : \mathrm{Sub}(Y) \to \mathrm{Sub}(X)$.

An elementary topos is a Heyting category (*a fortiori*, it is regular and coherent), which is sufficient for the interpretation of first order logic. In fact, one has the following, fundamental result.

Theorem 3.3. *For all objects X of a Heyting category $\mathbf{H}$, the poset $\mathrm{Sub}(X)$ of subobjects of X has the structure of a Heyting algebra, and, for all morphisms $f : Y \to X$, the change of base functor $f^* : \mathrm{Sub}(X) \to \mathrm{Sub}(Y)$ is a homomorphism of Heyting algebras.*

Proof. We show how to construct the Heyting algebra structure, and refer to [34, Theorem 8.1] for a full proof.

The operation $\wedge$ corresponds to taking the *pullback* of a subobject A along another subobject B:

$$\begin{tikzcd} A \wedge B \arrow[r, tail] \arrow[d, tail] & B \arrow[d, tail] \\ A \arrow[r, tail] & X \end{tikzcd} .$$

The operation $\vee$ corresponds to taking the image of the coproduct $A+B$ in X, which is always possible in regular categories, since they admit a notion of image factorisation:

$$\begin{tikzcd} A \arrow[r, tail] \arrow[ddr, tail, bend right] & A+B \arrow[d, two heads] \arrow[dd, dashed, bend left] & B \arrow[l, tail] \arrow[ddl, tail, bend left] \\ & A \vee B \arrow[d, tail] & \\ & X & \end{tikzcd} .$$

The constants $\top$, $\bot$ correspond, respectively, to the maximal and minimal subobjects $X = X$, $0 \rightarrowtail X$.

Finally, Heyting implication $\rightarrow$ is defined by

$$A \rightarrow B := \forall_m (A \wedge B) ,$$

where $m : A \rightarrowtail X$ is a monomorphism of A into X, and $\forall_m$ is the right adjoint to m^* given by the definition of Heyting category. We can easily check that this works: for subobjects A, B, C, with $m : A \rightarrowtail X$,

$$C \leq A \rightarrow B$$

if and only if, by $\forall_m$ being a right adjoint,

$$m^* C \leq A \wedge B ;$$

but $m^* C = A \wedge C$, and clearly $A \wedge C \leq A \wedge B$ if and only if $A \wedge C \leq B$. $\square$

The left and right adjoints to the change of base functors, which exist, respectively, in all regular categories and in all Heyting categories, enable us to also interpret existential and universal quantifiers in a Heyting category, in a way that we will later make more precise.

What an elementary topos $\mathcal{E}$ has, that a Heyting category does not, is the *subobject classifier*: a monomorphism $\mathrm{true} : 1 \rightarrowtail \Omega$, where 1 is the terminal object of $\mathcal{E}$, such that for all subobjects $m : A \rightarrowtail X$ in $\mathcal{E}$, there exists a unique morphism $\mathrm{char}A : X \rightarrow \Omega$, the *characteristic morphism* of A, such that

$$\begin{tikzcd} A \arrow[r] \arrow[d, tail, "m"] & 1 \arrow[d, tail, "\mathrm{true}"] \\ X \arrow[r, "\mathrm{char}A"'] & \Omega \end{tikzcd}$$

is a pullback diagram.

With a subobject classifier, one can construct *power objects*, by defining for all objects X of $\mathcal{E}$

$$PX := \Omega^X \ .$$

Then, by the definition of exponentials in a cartesian closed category, *global elements* of PX, i.e. morphisms $1 \to PX$, are in one-to-one correspondance with morphisms $1 \times X \simeq X \to \Omega$; hence, by definition of the subobject classifier, with subobjects of X. So power objects generalise *powersets* to arbitrary topoi, and allow one to interpret higher order logic within the latter.

For a model of the functional interpretations of Chapter 2, we actually only need first order logic; but there are other reasons why we may want to give up the full "elementary topos" structure for our models.

In the last century, Zermelo-Fraenkel set theory, with or without the axiom of choice, has been the predominant "universe of mathematics", so reasonably **Set**, a category of sets and functions, is usually taken to be the archetypal topos. Yet there is a considerable plurality of views on how to treat this category, in which two main approaches can be distinguished.

- One may want **Set** to mimick as closely as possible the universe of ZFC, even when working in a constructive metatheory; so **Set** should be *required* to be Boolean, satisfy the axiom of choice, etc. For instance, one may expect **Set** to be a model of Lawvere's *Elementary Theory of the Category of Sets* (ETCS) [31], a categorical axiomatisation of classical set theory.

- One, instead, may want **Set** to be a "small" copy of their preferred metatheory - usually, one where a universe has been fixed; as some form of sets and functions is a basic part of all but the most exotic foundations of mathematics, it makes sense to require that a category of sets and functions should behave in agreement with our informal understanding.

We will follow the latter, "pluralist" approach. *Predicative* foundations of mathematics, however, reject the *powerset axiom* of set theory - in categorical terms, the existence of a (strict) subobject classifier; so, if the second approach is taken, it may happen that **Set** fails to be an elementary topos (although it will, of course, still be a Heyting category). Notions of *predicative topos*, closed under the most important constructions of topoi from topoi, have been proposed, e.g. in [61], but these are not of particular relevance here.

What is relevant, though, is that none of the definitions we give, and of the properties we prove concerning the "filter topoi", need nonconstructive, or controversial assumptions about the metatheory - *except for one case*, where it seems preferable to have the axiom of choice at hand.

Now, much of the theory of Moerdijk's topos for nonstandard arithmetic has been developed by Palmgren in the constructive, predicative framework of *Martin-Löf dependent type theory* (for which [41] is a good reference); and a type-theoretic version of the axiom of choice is a *theorem* of the latter. Thus, although we do not want to impose a single metatheory, it is true that this may be a good choice; but it does entail that **Set** is no longer an elementary topos.

Whatever your preference about **Set** is, the notion of Grothendieck topos can be formulated relative to it, and this is all we need for now.

Definition 3.4. Let $\mathbf{C}$ be a category. A *presheaf* over $\mathbf{C}$ is an object of $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$, the category of contravariant functors from $\mathbf{C}$ to $\mathbf{Set}$, and natural transformations.

> **Notation.** As is customary, if a presheaf F over $\mathbf{C}$ is clear from the context, for all objects X of $\mathbf{C}$, $x \in FX$, and $g : Y \to X$, we write $x \cdot g$ for the element $Fg(x) \in FY$.

For all objects X of $\mathbf{C}$, we write $\mathbf{y}X$ for the *representable presheaf* $\mathrm{Hom}_{\mathbf{C}}(-, X)$, which acts as precomposition on morphisms. This defines an embedding - the *Yoneda embedding* - of $\mathbf{C}$ into $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$. That this is actually an embedding, i.e. a functor that is full, faithful and injective on objects, is guaranteed by the fundamental *Yoneda lemma*, stating that for all presheaves F and objects X there is a natural isomorphism

$$FX \simeq \mathrm{Hom}_{[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]}(\mathbf{y}X, F) \ .$$

Categories of presheaves are the simplest example of Grothendieck topoi; all the other Grothendieck topoi are obtained by singling out certain presheaves as *sheaves*, requiring that they behave properly with respect to a certain notion of *covering* of an object of $\mathbf{C}$. By analogy with sheaves on a topological space, this notion has been called a *topology* on $\mathbf{C}$. Here we will work with *bases* for topologies, in a version inspired by [44].

Definition 3.5. Let $\mathbf{C}$ be a category with pullbacks. A (basis for a) *Grothendieck topology* on $\mathbf{C}$ is a relation K between objects X of $\mathbf{C}$, and sets of morphisms $\{f_i : X_i \to X\}_{i \in I}$ with codomain X, indexed by a set I, to be read

"$\{f_i : X_i \to X\}_{i \in I}$ K-covers X",

such that:

1. if $f : X' \to X$ is an isomorphism, then the singleton $\{f : X' \to X\}$ K-covers X;
2. if $\{f_i : X_i \to X\}_{i \in I}$ K-covers X, and $g : Y \to X$ is any morphism to X, then the set of the $\{\pi_i^1 : Y \times_X X_i \to Y\}_{i \in I}$, as in

$$\begin{tikzcd} Y \times_X X_i \arrow[r, "\pi_i^2"] \arrow[d, "\pi_i^1"'] & X_i \arrow[d, "f_i"] \\ Y \arrow[r, "g"] & X \end{tikzcd} ,$$

   K-covers Y;
3. if $\{f_i : X_i \to X\}_{i \in I}$ K-covers X, and, for all $i \in I$, $\{g_{ij} : X_{ij} \to X_i\}_{j \in J_i}$ K-covers X_i, then the set of composites $\{f_i g_{ij} : X_{ij} \to X\}_{i \in I, j \in J_i}$ K-covers X.

A pair $(\mathbf{C}, K)$ of a category and a Grothendieck topology on it is called a *site*.

The proper notion of "Grothendieck topology" is usually formulated in terms of covering *sieves*. A sieve S on an object X in $\mathbf{C}$ is a family of maps with codomain X that is closed under precomposition: if $f : Y \to X$ is in S, then, for all $g : Z \to Y$, fg is in S. Equivalently, S is a subpresheaf of $\mathbf{y}X$. Every covering family generates a covering sieve, by closing under this condition, so every basis K for a Grothendieck topology generates a Grothendieck topology J.

Definition 3.6. Let $\{f_i : X_i \to X\}_{i\in I}$ be a K-cover of X, and F a presheaf over $\mathbf{C}$. A set $\{x_i \mid x_i \in FX_i\}_{i\in I}$ is a *matching family* of elements of F for the cover if, for all $i, j \in I$, given pullback projections π^1_{ij}, π^2_{ij} as in

$$\begin{tikzcd} X_i \times_X X_j \arrow[r, "\pi^2_{ij}"] \arrow[d, "\pi^1_{ij}"] & X_j \arrow[d, "f_j"] \\ X_i \arrow[r, "f_i"] & X \end{tikzcd},$$

it holds that $x_i \cdot \pi^1_{ij} = x_j \cdot \pi^2_{ij}$; informally, the elements "coincide on intersections".

If $\{x_i \mid x_i \in FX_i\}_{i\in I}$ is a matching family of elements of F, and $x \in FX$, we say that x is an *amalgamation* for the matching family if, for all $i \in I$, $x_i = x \cdot f_i$.

A presheaf F over $\mathbf{C}$ is a K-*sheaf* if, for all objects X of $\mathbf{C}$, and all K-covers $\{f_i : X_i \to X\}_{i\in I}$ of X, every matching family $\{x_i \mid x_i \in FX_i\}_{i\in I}$ of elements of F has a *unique* amalgamation in X.

We write $\mathrm{Sh}(\mathbf{C}, K)$ for the subcategory of the K-sheaves in $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$; such a category is called a *Grothendieck topos*. Notice that, with respect to the minimal Grothendieck topology where only single isomorphisms are covering, all presheaves are sheaves; so this definition does include all presheaf categories.

It is a fundamental result of sheaf theory that the inclusion of Grothendieck topoi $i : \mathrm{Sh}(\mathbf{C}, K) \to [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$ has a left adjoint $\mathbf{a} : [\mathbf{C}^{\mathrm{op}}, \mathbf{Set}] \to \mathrm{Sh}(\mathbf{C}, K)$, the so called *sheafification* functor.

Every Grothendieck topos is a Heyting category, and also an elementary topos, provided that **Set** is. We refer again to [34] for the construction of limits, colimits, exponentials, and the subobject classifier.

3.1.1 The interpretation of first order logic

We can now show how to interpret first order logic in a Heyting category; we will use capital Latin letters for types, in order to differentiate the current, more general context from the proof-theoretic perspective of Chapters 1 and 2. What follows is based on [65, Chapter 4].

Let $\mathcal{L}$ be a many sorted first order language with equality, with types $S, T, \ldots$; denumerable variables $x, y, z, \ldots : S$ of each type; possibly, constants $c, d, \ldots : S$, relation symbols $R \subseteq S_1, \ldots, S_n$, and function symbols $f : S_1, \ldots, S_n \to S$.

The class of terms of $\mathcal{L}$ is defined inductively as follows:

▷ a constant $c : S$ or a variable $x : S$ is a term of type S;

▷ if $t_1 : S_1, \ldots, t_n : S_n$ are terms, and $f : S_1, \ldots, S_n \to S$ is a function symbol, then $f(t_1, \ldots, t_n)$ is a term of type S.

The formulae of $\mathcal{L}$ are generated by the clauses

▷ $\perp$ is a formula;

▷ for all types S and terms $s, t : S$, $s = t$ is a formula;

▷ for all relation symbols $R \subseteq S_1, \ldots, S_n$, and terms $t_1 : S_1, \ldots, t_n : S_n$, $R(t_1, \ldots, t_n)$ is a formula;

▷ if φ, ψ are formulae, $\varphi \wedge \psi$, $\varphi \vee \psi$, $\varphi \to \psi$, $\exists x : S\, \varphi$, and $\forall x : S\, \varphi$ are formulae.

As before, we treat negation and coimplication as defined connectives.

An interpretation of $\mathcal{L}$ in a Heyting category **H** is obtained by choosing

(i) for each type S, an object $[\![S]\!]$ of **H**;

(ii) for each constant $c : S$, a global element $[\![c]\!] : 1 \to [\![S]\!]$;

(iii) for each function symbol $f : S_1, \ldots, S_n \to S$, a morphism $[\![f]\!] : [\![S_1]\!] \times \ldots \times [\![S_n]\!] \to [\![S]\!]$;

(iv) for each relation symbol $R \subseteq S_1, \ldots, S_n$, a subobject $[\![R]\!] \rightarrowtail [\![S_1]\!] \times \ldots \times [\![S_n]\!]$.

We will now extend this to an interpretation of all terms and formulae of $\mathcal{L}$. Suppose a term $t : S$, or a formula φ, have free variables $x_1 : S_1, \ldots, x_n : S_n$; then define $\mathrm{FV}(t) \equiv \mathrm{FV}(\varphi) := [\![S_1]\!] \times \ldots \times [\![S_n]\!]$. The idea is that t should be interpreted as a *morphism* $[\![t]\!] : \mathrm{FV}(t) \to [\![S]\!]$; while φ should be interpreted as a *subobject* $[\![\varphi]\!] \rightarrowtail \mathrm{FV}(\varphi)$ - in **Set**, it would be the subset $\{(x_1, \ldots, x_n) \mid \varphi(x_1, \ldots, x_n)\}$ of $[\![S_1]\!] \times \ldots \times [\![S_n]\!]$.

Definition 3.7. The interpretation of a term $t : S$ of $\mathcal{L}$ in **H** is defined by the inductive clauses:

▷ if t is a variable $x : S$, $[\![x]\!]$ is the identity of $[\![S]\!]$;

▷ given terms $t_1 : S_1, \ldots, t_n : S_n$, and a function symbol $f : S_1, \ldots, S_n \to S$, $[\![f(t_1, \ldots, t_n)]\!]$ is the composite morphism

$$\mathrm{FV}(f(t_1, \ldots, t_n)) \xrightarrow{\langle \pi_1, \ldots, \pi_n \rangle} \prod_{i=1}^n \mathrm{FV}(t_i) \xrightarrow{[\![t_1]\!] \times \ldots \times [\![t_n]\!]} \prod_{i=1}^n [\![S_i]\!] \xrightarrow{\quad [\![f]\!] \quad} [\![S]\!]\,,$$

where π_i is the projection $\mathrm{FV}(f(t_1, \ldots, t_n)) \to \mathrm{FV}(t_i)$, for $i = 1, \ldots, n$.

Definition 3.8. The interpretation of a formula φ of $\mathcal{L}$ in **H** is defined by the inductive clauses:

▷ $[\![\bot]\!]$ is the minimal subobject $0 \rightarrowtail 1$;

▷ for $s, t : S$, $[\![s = t]\!] \rightarrowtail \mathrm{FV}(s = t)$ is the equaliser of the diagram

$$\begin{tikzcd}[row sep=small] & \mathrm{FV}(s) \arrow[dr, "{[\![s]\!]}"] & \\ \mathrm{FV}(s = t) \arrow[ur] \arrow[dr] & & {[\![S]\!]}\,, \\ & \mathrm{FV}(t) \arrow[ur, "{[\![t]\!]}"'] & \end{tikzcd}$$

where the left side morphisms are product projections;

▷ for a relation symbol $R \subseteq S_1, \ldots, S_n$, and terms $t_1 : S_1, \ldots, t_n : S_n$, $[\![R(t_1, \ldots, t_n)]\!]$ is the pullback of $[\![R]\!] \rightarrowtail [\![S_1]\!] \times \ldots \times [\![S_n]\!]$ along the morphism

$$\mathrm{FV}(R(t_1, \ldots, t_n)) \xrightarrow{\langle \pi_1, \ldots, \pi_n \rangle} \prod_{i=1}^n \mathrm{FV}(t_i) \xrightarrow{[\![t_1]\!] \times \ldots \times [\![t_n]\!]} \prod_{i=1}^n [\![S_i]\!]\,;$$

- $\triangleright$ given formulae φ, ψ, with interpretations $[\![\varphi]\!] \rightarrowtail \mathrm{FV}(\varphi)$, $[\![\psi]\!] \rightarrowtail \mathrm{FV}(\psi)$, and product projections

$$\begin{tikzcd} & \mathrm{FV}(\varphi \wedge \psi) \arrow[dl, "\pi_1"'] \arrow[dr, "\pi_2"] & \\ \mathrm{FV}(\varphi) & & \mathrm{FV}(\psi) \end{tikzcd} ,$$

we define

$$[\![\varphi \wedge \psi]\!] := \pi_1^*[\![\varphi]\!] \wedge \pi_2^*[\![\psi]\!] ,$$
$$[\![\varphi \vee \psi]\!] := \pi_1^*[\![\varphi]\!] \vee \pi_2^*[\![\psi]\!] ,$$
$$[\![\varphi \to \psi]\!] := \pi_1^*[\![\varphi]\!] \to \pi_2^*[\![\psi]\!] ,$$

where the operations are performed in $\mathrm{Sub}(\mathrm{FV}(\varphi \wedge \psi))$;

- $\triangleright$ given a formula φ with interpretation $[\![\varphi]\!] \rightarrowtail \mathrm{FV}(\varphi)$, let π' be the composite projection

$$\mathrm{FV}(\varphi \wedge (x = x)) \xrightarrow{\pi} \mathrm{FV}(\varphi) \longrightarrow \mathrm{FV}(\exists x : S\, \varphi) ;$$

then we define

$$[\![\exists x : S\, \varphi]\!] := \exists_{\pi'} \pi^*[\![\varphi]\!] ,$$
$$[\![\forall x : S\, \varphi]\!] := \forall_{\pi'} \pi^*[\![\varphi]\!] .$$

At last, we can define a notion of (local) truth of a formula in a category.

Definition 3.9. Let φ be a formula of $\mathcal{L}$, and $a : X \to \mathrm{FV}(\varphi)$ a morphism of $\mathbf{H}$ (a *generalised element*). We say that X *forces* $\varphi(a)$, in symbols

$$X \Vdash \varphi(a) ,$$

if the pullback $a^*[\![\varphi]\!]$ of $[\![\varphi]\!] \rightarrowtail \mathrm{FV}(\varphi)$ along a is the maximal subobject of X.

We just write $\Vdash \varphi(a)$ for $1 \Vdash \varphi(a)$. If φ is a *sentence* of $\mathcal{L}$, we say that φ is *true* in $\mathbf{H}$ if $\Vdash \varphi$.

The forcing relation has the following, important properties, which follow immediately from the definition.

Monotonicity. If $X \Vdash \varphi(a)$, then for all morphisms $f : Y \to X$ in $\mathbf{H}$, also $Y \Vdash \varphi(af)$.

Local character. If $f : Y \twoheadrightarrow X$ is an epimorphism and $Y \Vdash \varphi(af)$, then also $X \Vdash \varphi(a)$.

It is possible to give an inductive definition of the forcing relation - the so-called *Kripke-Joyal semantics* of first order logic; a far-reaching generalisation of the Kripke semantics for intuitionistic logic, which are indeed the special case where the underlying category is a poset.

The fundamental fact about these semantics is that they are *sound* and *complete* for intuitionistic first order logic; so a purely intuitionistic proof is a valid proof in any Heyting category. In fact, if one considers the *internal language* of the category - whose types correspond to the objects of $\mathbf{H}$, the function symbols correspond to the morphisms, etc. - and interprets it in in the obvious way, one can use intuitionistic reasoning, combined with the Kripke-Joyal semantics, to prove "logically" facts about the category itself.

We will specialise to the case of Grothendieck topoi, where the Kripke-Joyal semantics can be simplified. So let $\mathbf{H} = \mathrm{Sh}(\mathbf{C}, K)$ for some site $(\mathbf{C}, K)$. The advantage of Grothendieck topoi, here, is that one can force only with *representables*, i.e. objects of the form $\mathbf{ay}C$, for C an object of $\mathbf{C}$; this is due to their being *separating* for the category, but this is not a notion we need to make explicit.

So suppose we have fixed an interpretation of $\mathcal{L}$ in $\mathbf{H}$, and let $a : \mathbf{ay}C \to \textsc{fv}(\varphi)$ be a generalised element; by the adjunction $\mathbf{a} \dashv i$, this corresponds to a unique $a : \mathbf{y}C \to \textsc{fv}(\varphi)$ in $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$, which in turn, by the Yoneda lemma, corresponds to a unique element $a \in \textsc{fv}(\varphi)C$.

Moreover, $\mathbf{ay}C \Vdash \varphi(a)$ if and only if a factors through $[\![\varphi]\!]$; but, again by the Yoneda lemma, this happens if and only if the element $a \in \textsc{fv}(\varphi)C$ constructed earlier is in fact in $[\![\varphi]\!]C$. Therefore, we can redefine the forcing relation as follows:

$$C \Vdash \varphi(a) \qquad \text{if and only if} \qquad a \in [\![\varphi]\!]C\ .$$

Monotonicity and local character take the following form.

> **Monotonicity.** If $C \Vdash \varphi(a)$, then for all morphisms $f : D \to C$ in $\mathbf{C}$, also $D \Vdash \varphi(a \cdot f)$.
>
> **Local character.** If $\{f_i : C_i \to C\}_{i \in I}$ is a K-cover of C, and, for all $i \in I$, $C_i \Vdash \varphi(a \cdot f_i)$, then also $C \Vdash \varphi(a)$.

We are now able to give an inductive formulation of forcing in a Grothendieck topos, the so-called *sheaf semantics*.

Theorem 3.10. *Suppose that we have fixed an interpretation of $\mathcal{L}$ in the Grothendieck topos $\mathrm{Sh}(\mathbf{C}, K)$. Let C be an object of $\mathbf{C}$, φ, ψ formulae of $\mathcal{L}$, and $a \in \textsc{fv}(\varphi \wedge \psi)C$. Then*

(i) $C \Vdash \bot(a)$ *if and only if the empty family is a K-cover of C;*

(ii) $C \Vdash (s = t)(a)$ *if and only if* $[\![s]\!]C(a) = [\![t]\!]C(a)$*;*

(iii) $C \Vdash R(t_1, \ldots, t_n)(a)$ *if and only if* $([\![t_1]\!]C(a), \ldots, [\![t_n]\!]C(a)) \in [\![R]\!]C$*;*

(iv) $C \Vdash \varphi(a) \wedge \psi(a)$ *if and only if* $C \Vdash \varphi(a)$ *and* $C \Vdash \psi(a)$*;*

(v) $C \Vdash \varphi(a) \vee \psi(a)$ *if and only if there exists a K-cover $\{f_i : C_i \to C\}_{i \in I}$ of C such that, for each $i \in I$, either $C_i \Vdash \varphi(a \cdot f_i)$ or $C_i \Vdash \psi(a \cdot f_i)$;*

(vi) $C \Vdash \varphi(a) \to \psi(a)$ *if and only if for all $f : D \to C$ in $\mathbf{C}$, $D \Vdash \varphi(a \cdot f)$ implies $D \Vdash \psi(a \cdot f)$;*

(vii) $C \Vdash \exists y : S\, \varphi(a, y)$ *if and only if there exist a K-cover $\{f_i : C_i \to C\}_{i \in I}$ of C, and elements $b_i \in [\![S]\!]C_i$, $i \in I$, such that, for all $i \in I$, $C_i \Vdash \varphi(a \cdot f_i, b_i)$;*

(viii) $C \Vdash \forall y : S\,\varphi(a, y)$ *if and only if for all* $f : D \to C$ *in* $\mathbf{C}$*, and for all* $b \in [\![S]\!]D$*,* $D \Vdash \varphi(a \cdot f, b)$*.*

In the case where $[\![S]\!] = \mathbf{ay}B$ *for some object* B *of* $\mathbf{C}$*, the last clause can be simplified:*

(viii') $C \Vdash \forall y : S\,\varphi(a, y)$ *if and only if* $C \times B \Vdash \varphi(a \cdot \pi_1, \pi_2)$*.*

Proof. This is [34, Theorem VI.7.1]. □

With this inevitably rushed survey of categorical logic, we can leave the preliminaries aside for a while, and concentrate on the categories of interest.

3.2 The filter construction

In [9], Blass introduced a category of filters of sets and “continuous” maps between them; rediscovered by Moerdijk, it was used as the underlying category of a site, whose sheaves provided a model of nonstandard arithmetic.

This category arises from **Set** as a special case of a general construction - the *filter construction* - whose properties and functoriality were studied by Butz in [10]. When applied on arbitrary categories with finite limits, it can be considered as a completion of the subobject posets under arbitrary meets. We will briefly discuss the general construction, following Butz, before specialising to the case of **Set**.

We start by recalling the definition of filter on a $\wedge$-semilattice, i.e. on a poset with all finite meets.

Definition 3.11. Let S be a $\wedge$-semilattice. A *filter* on S is an inhabited, upwards closed subset of S that is closed under binary meets.

We say that a filter is *proper* if it does not coincide with S; otherwise, it is *non proper*.

Following Palmgren, we will work more often with *filter bases*, indexed by a set I.

Definition 3.12. A *filter base* $\mathcal{F}_I$ on S is an inhabited set $\{\mathcal{F}_i\}_{i \in I}$ of elements of S, such that, for all $i, j \in I$, there exists $k \in I$ such that $\mathcal{F}_k \le \mathcal{F}_i \wedge \mathcal{F}_j$.

A filter base *generates* a filter, as follows: A belongs to the filter if and only if there exists $i \in I$ such that $\mathcal{F}_i \le A$.

Notice that a filter base generates a *non* proper filter if and only if it contains the bottom element.

In every category $\mathbf{C}$ with finite limits, the subobject posets are in fact $\wedge$-semilattices; it is therefore possible to speak of filters of subobjects. That is sufficient to perform the filter construction.

Definition 3.13. Let $\mathbf{C}$ be a finitely complete category. The *filter category* $\mathfrak{F}\mathbf{C}$ over $\mathbf{C}$ is described by the following data.

- Objects are pairs $(C, \mathcal{F}_I)$, where C is an object of $\mathbf{C}$, and $\mathcal{F}_I$ is an I-indexed filter base on $\mathrm{Sub}(C)$.

We will usually write $\mathcal{F}$ for $(C, \mathcal{F}_I)$, when the underlying object and indexing set are not relevant, and just call it a *filter*. We say that the $\mathcal{F}_i$, $i \in I$, are the *base objects* of the filter.

- Morphisms are "germs of continuous morphisms". A continuous morphism $\alpha : (C, \mathcal{F}_I) \to (D, \mathcal{G}_J)$ is a partial morphism

$$\begin{tikzcd} & \mathcal{F}_i \arrow[dl, tail] \arrow[dr, "\alpha"] & \\ C & & D \end{tikzcd}$$

  in $\mathbf{C}$, defined on some base object $\mathcal{F}_i$, such that for all $j \in J$, there exists $i' \in I$ such that $\mathcal{F}_{i'} \leq \alpha^* \mathcal{G}_j$ in $\mathrm{Sub}(C)$.

  We declare two such morphisms $\alpha : \mathcal{F}_i \to D$, $\alpha' : \mathcal{F}_j \to D$ *equivalent* if there exists $k \in I$ such that $\mathcal{F}_k \leq \mathcal{F}_i \wedge \mathcal{F}_j$, and $\alpha\big|_{\mathcal{F}_k} = \alpha'\big|_{\mathcal{F}_k}$; that is, the following pullback square commutes:

$$\begin{tikzcd} \mathcal{F}_k \arrow[r, tail] \arrow[d, tail] & \mathcal{F}_i \arrow[d, "\alpha"] \\ \mathcal{F}_j \arrow[r, "\alpha'"] & D \, . \end{tikzcd}$$

We have an embedding of $\mathbf{C}$ into $\mathfrak{F}\mathbf{C}$, where an object C of $\mathbf{C}$ is identified with the "simple" filter $(C, \{C\})$. We will usually still denote the latter with C.

We will not be overly pedantic about distinguishing between morphisms and their germs, and will write both in the same style.

Lemma 3.14. *The category $\mathfrak{F}\mathbf{C}$ is finitely complete.*

Proof. It is sufficient that $\mathfrak{F}\mathbf{C}$ has a terminal object, binary products and equalisers. We give their construction, and omit the proof of the universal properties.

The terminal object is the filter $(1, \{1\})$. The product of $(C, \mathcal{F}_I)$ and of $(D, \mathcal{G}_J)$ is the filter $(C \times D, (\mathcal{F} \times \mathcal{G})_{I \times J})$, where $\mathcal{F} \times \mathcal{G}_{(i,j)} := \mathcal{F}_i \times \mathcal{G}_j$, for all $i \in I$, $j \in J$.

The equaliser of two morphisms $\alpha, \beta : (C, \mathcal{F}_I) \to (D, \mathcal{G}_J)$, represented by $\alpha : \mathcal{F}_i \to D$ and $\beta : \mathcal{F}_j \to D$, is the inclusion $(C', (\mathcal{F} \wedge C')_I) \rightarrowtail (C, \mathcal{F}_I)$, where C' is the equaliser of α and β in $\mathbf{C}$, and $(\mathcal{F} \wedge C')_i := \mathcal{F}_i \wedge C'$ for all $i \in I$. □

Lemma 3.15. *A morphism $\alpha : \mathcal{F} \to \mathcal{G}$ of $\mathfrak{F}\mathbf{C}$, defined on a base object $\mathcal{F}_i$, is a monomorphism if and only if there exists a base object $\mathcal{F}_j \leq \mathcal{F}_i$ such that $\alpha\big|_{\mathcal{F}_j}$ is a monomorphism in $\mathbf{C}$.*

Proof. See [10, Lemma 2.2]. □

Proposition 3.16. *For all filters $\mathcal{F}$ in $\mathfrak{F}\mathbf{C}$, $\mathrm{Sub}(\mathcal{F})$ is a meet-complete semilattice, and, for all $\alpha : \mathcal{F} \to \mathcal{G}$, the change of base functor α^* preserves all meets.*

Proof. By the previous lemma, if $\alpha : (C, \mathcal{F}_I) \rightarrowtail (D, \mathcal{G}_J)$ is a monomorphism, there is some base object $\mathcal{F}_i$ such that $\alpha\big|_{\mathcal{F}_i} : \mathcal{F}_i \rightarrowtail D$ is a monomorphism in $\mathbf{C}$. Then $(C, \mathcal{F}_I)$ is isomorphic to the filter $(D, (\mathcal{G} \wedge \mathcal{F}_i)_J)$. It follows that subobjects of $(D, \mathcal{G}_J)$ are in one-to-one correspondance to objects $(D, \mathcal{G}'_{J'})$, such that the base $\mathcal{G}'_{J'}$ generates a filter larger than $\mathcal{G}_J$.

Given an arbitrary family of subobjects $\{(D, \mathcal{G}^{(i)}_{J^{(i)}})\}_{i \in I}$, let $\mathcal{H}$ be the filter generated by finite meets of the form

$$\mathcal{G}^{(i_1)}_{j^{(i_1)}} \wedge \ldots \wedge \mathcal{G}^{(i_n)}_{j^{(i_n)}} ,$$

for $(i_1, \ldots, i_n)$ an arbitrary finite sequence in I, and $j^{(i_k)} \in J^{(i_k)}$, $k = 1, \ldots, n$. Then

$$\bigwedge_{i \in I} (D, \mathcal{G}^{(i)}_{J^{(i)}}) \simeq (D, \mathcal{H}) .$$

That this is preserved by change of base can be easily verified by the explicit construction of pullbacks in $\mathfrak{F}\mathbf{C}$. □

An important feature of the filter construction is that it preserves some of the additional properties that $\mathbf{C}$ may have.

Proposition 3.17. *Let* $\mathbf{C}$ *be a finitely complete category.*

(a) If $\mathbf{C}$ *is regular, then* $\mathfrak{F}\mathbf{C}$ *is also regular.*

(b) If $\mathbf{C}$ *is coherent, then* $\mathfrak{F}\mathbf{C}$ *is also coherent.*

Proof. See [10, Proposition 3.1] and [10, Proposition 3.2], respectively. □

Moreover, if $\mathbf{C}$ has all finite coproducts, then $\mathfrak{F}\mathbf{C}$ has them too. In this case, the initial object of $\mathfrak{F}\mathbf{C}$ is the simple filter $(0, \{0\})$; this is isomorphic to any non proper filter $(C, \mathcal{F}_I)$, where $\mathcal{F}_i = 0$ for some $i \in I$. Given two filters $(C, \mathcal{F}_I)$ and $(D, \mathcal{G}_J)$, their coproduct in $\mathfrak{F}\mathbf{C}$ is the filter $(C + D, (\mathcal{F} + \mathcal{G})_{I \times J})$, where $\mathcal{F} + \mathcal{G}_{(i,j)} := \mathcal{F}_i + \mathcal{G}_j$, for all $i \in I$, $j \in J$.

It is *not*, however, the case that $\mathfrak{F}\mathbf{H}$ is necessarily a Heyting category, when $\mathbf{H}$ is. But this is not a problem, since we really only need $\mathfrak{F}\mathbf{Set}$ to be a coherent category.

As it happens, coherent categories admit a "natural" Grothendieck topology, sometimes called the *precanonical* topology: for all objects C of $\mathbf{C}$, a K-cover of C is a *finite* family $\{f_i : C_i \to C\}_{i=1}^n$, such that the union of the images of the f_i is the whole of C.

As shown in [24, Example C2.1.12.(d)], K is *subcanonical*; that is, representable presheaves, of the form $\mathbf{y}C$, for C an object of $\mathbf{C}$, are K-sheaves.

Explicitly, for a filter category $\mathfrak{F}\mathbf{C}$, that $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^n$ is a K-cover means that, for all choices of base objects $\mathcal{G}_{k,j_k}$ of $\mathcal{G}_k$, $k = 1, \ldots, n$, there exists a base object $\mathcal{F}_i$ of $\mathcal{F}$ such that

$$\mathcal{F}_i \leq \beta_1 \mathcal{G}_{1,j_1} \vee \ldots \vee \mathcal{G}_{n,j_n} .$$

Besides K, we will also consider the smaller topology K_1, where covers of C are *single* covering morphisms $\{f : D \twoheadrightarrow C\}$. *A fortiori*, K_1 is also subcanonical for coherent categories.

Finally, everything is set for the definition of our sheaf models.

Definition 3.18. We call $\mathcal{U}$ the Grothendieck topos $\mathrm{Sh}(\mathfrak{F}\mathbf{Set}, K_1)$, and $\mathcal{N}$ its subtopos $\mathrm{Sh}(\mathfrak{F}\mathbf{Set}, K)$.

$\mathcal{N}$, for *nonstandard* universe, is the name used by Palmgren for Moerdijk's topos; and we used the letter $\mathcal{U}$ to remind of the *uniform* Diller-Nahm interpretation.

As for all Grothendieck topoi, the *global sections* functors

$$\Gamma_1 : \mathcal{U} \to \mathbf{Set} \ ,$$
$$\Gamma : \mathcal{N} \to \mathbf{Set} \ ,$$

sending a sheaf F to the set $\mathrm{Hom}(1, F)$, have left adjoints $\Delta_1 : \mathbf{Set} \to \mathcal{U}$, and $\Delta : \mathbf{Set} \to \mathcal{N}$, respectively - the *constant objects* functors. These can be explicitly characterised as follows: for all sets S, at all filters $\mathcal{F}$ of $\mathfrak{F}\mathbf{Set}$,

$$(\Delta_1 S)\mathcal{F} = \{\alpha : \mathcal{F} \to S \mid \alpha \text{ is constant}\} \ ,$$
$$(\Delta S)\mathcal{F} = \{\alpha : \mathcal{F} \to S \mid \alpha \text{ takes a finite number of values}\} \ .$$

Here, S is identified with the simple filter $(S, \{S\})$. Since Δ_1 and Δ are left adjoints, they preserve coproducts; in particular, if $2 := 1 + 1$ in the relevant topos,

$$2 \simeq \Delta_1 2 \qquad \text{in } \mathcal{U} \ ,$$
$$2 \simeq \Delta 2 \equiv \mathbf{y}2 \qquad \text{in } \mathcal{N} \ .$$

It follows that the Yoneda embedding in $\mathcal{N}$ preserves all *finite* coproducts of copies of 1, and no coproducts in $\mathcal{U}$. Thus, in $\mathcal{U}$, there is a proper monomorphism $m : 2 \rightarrowtail \mathbf{y}2$; moreover, since the sheafification functor $\mathbf{a}$ of $\mathcal{N}$ is itself left adjoint to the inclusion of $\mathcal{N}$ in $\mathcal{U}$, we have that $\mathbf{a}\Delta_1 2 \simeq \mathbf{y}2$. We say that m is a *K-dense* morphism.

Indeed, this fact alone characterises the topology of $\mathcal{N}$ with respect to $\mathcal{U}$. To make this rigorous, we will temporarily assume that our Grothendieck topoi have a subobject classifier, i.e. are elementary topoi. In this way, we can use a more general notion of "topology" on a topos, due to Lawvere and Tierney.

Definition 3.19. Let $\mathcal{E}$ be an elementary topos. A *local operator* in $\mathcal{E}$ is a morphism $j : \Omega \to \Omega$, where Ω is the subobject classifier of $\mathcal{E}$, such that the following diagrams commute:

$$\begin{tikzcd} & 1 \arrow[dl, "\mathrm{true}"'] \arrow[d, tail, "\mathrm{true}"] \\ \Omega \arrow[r, "j"] & \Omega \ , \end{tikzcd} \qquad \begin{tikzcd} & \Omega \arrow[dl, "j"'] \arrow[d, "j"] \\ \Omega \arrow[r, "j"] & \Omega \ , \end{tikzcd} \qquad \begin{tikzcd} \Omega \times \Omega \arrow[r, "\wedge"] \arrow[d, "j \times j"] & \Omega \arrow[d, "j"] \\ \Omega \times \Omega \arrow[r, "\wedge"] & \Omega \ . \end{tikzcd} \tag{3.1}$$

Let X be an object of $\mathcal{E}$. A monomorphism $A \rightarrowtail X$ of X is *j-dense* if the diagram

$$\begin{tikzcd} & & 1 \arrow[d, tail, "\mathrm{true}"] \\ X \arrow[r, "\mathrm{char} A"] \arrow[urr, bend left, "!"] & \Omega \arrow[r, "j"] & \Omega \end{tikzcd}$$

commutes, for $\mathrm{char}A$ the characteristic morphism of A; so $j\,\mathrm{char}A$ classifies the maximal subobject of X.

Let F be an object of $\mathcal{E}$. We say that F is a *j-sheaf* if, for all j-dense monomorphisms $A \rightarrowtail X$, the induced morphism

$$\mathrm{Hom}_{\mathcal{E}}(X, F) \to \mathrm{Hom}_{\mathcal{E}}(A, F)$$

is an isomorphism. We write $\mathrm{sh}_j(\mathcal{E})$ for the subcategory of j-sheaves of $\mathcal{E}$.

In a Grothendieck topos, the subobject classifier admits the following description.

Definition 3.20. Given a Grothendieck topology J on $\mathbf{C}$, we say that a sieve S on an object C of $\mathbf{C}$ is *J-closed* if, for all $f : D \to C$, if the pullback sieve $f^*S := \{g : E \to D \mid fg \in S\}$ on D is J-covering, then $f \in S$.

The subobject classifier in $\mathrm{Sh}(\mathbf{C}, J)$ is defined by

$$\Omega C := \{S \mid S \text{ is a } J\text{-closed sieve on } C\} ,$$

for all objects C of $\mathbf{C}$. A local operator j characterises a subobject of Ω, hence selects for all objects a family of sieves; the conditions (3.1) guarantee that this defines a Grothendieck topology on $\mathbf{C}$.

In fact, the two notions are equivalent in the case of presheaf topoi: a Grothendieck topology J on $\mathbf{C}$ uniquely defines a local operator j on $[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$, and vice versa, so that $\mathrm{Sh}(\mathbf{C}, J) = \mathrm{sh}_j[\mathbf{C}^{\mathrm{op}}, \mathbf{Set}]$.

Local operators on a topos $\mathcal{E}$ form a lattice, induced by the lattice structure of Ω; so it makes sense to compare local operators, and speak of a smaller, or larger one relative to another.

Proposition 3.21. *Let $m : A \rightarrowtail X$ be a monomorphism in a topos $\mathcal{E}$. Then there exists a smallest local operator j on $\mathcal{E}$ such that m is j-dense.*

Proof. See [24, Example A4.5.14.(b)]. □

Such a local operator can be constructed explicitly in two steps, as described by [24, Proposition A4.5.12].

1. First, we take the image in the subobject classifier Ω of the characteristic map of m. This yields a monomorphism $D \rightarrowtail \Omega$.

2. Then, we apply the consecutive transformations $D \mapsto D^r \mapsto D^{rl}$, described by the formulae

$$D^r(q) := \forall p : \Omega \left((D(p) \wedge p \to q) \to q\right) ,$$
$$D^{rl}(p) := \forall q : \Omega \left((D^r(q) \wedge p \to q) \to q\right)$$

in the internal language of the topos.

One checks that there is an inclusion $D \rightarrowtail D^{rl}$, and that the characteristic map of D^{rl} is the requested local operator.

We can now provide a characterisation of $\mathcal{N}$ with respect to $\mathcal{U}$.

Proposition 3.22. *Let j be the smallest local operator on $\mathcal{U}$ such that $m : 2 \rightarrowtail \mathbf{y}2$ is j-dense. Then $\mathrm{sh}_j(\mathcal{U}) \simeq \mathcal{N}$.*

Proof. By the construction of characteristic maps in Grothendieck topoi [34, Proposition III.7.3], we have at all filters $\mathcal{F}$, for all $\alpha : \mathcal{F} \to 2$,

$$\mathrm{char}2\,\mathcal{F}(\alpha) = \{\beta : \mathcal{G} \to \mathcal{F} \mid \alpha\beta \text{ is constant}\}.$$

Clearly, $\mathrm{char}2\,\mathcal{F}(\alpha)$ is a sieve on $\mathcal{F}$, closed for the topology K_1.

Let $\mathcal{F}_1 := \alpha^*0$, $\mathcal{F}_2 := \alpha^*1$ be the pullbacks along α of the two global elements of 2 in $\mathfrak{F}\mathbf{Set}$. Then $\mathcal{F} \simeq \mathcal{F}_1 + \mathcal{F}_2$, and $\beta \in \mathrm{char}2\,\mathcal{F}(\alpha)$ if and only if β factors either through $\mathcal{F}_1$ or through $\mathcal{F}_2$. Thus, there is a bijection

$$\mathrm{char}2\,\mathcal{F}(\alpha) \simeq \mathbf{y}\mathcal{F}_1 + \mathbf{y}\mathcal{F}_2 \ ,$$

provided that, if $\mathcal{F}_1$ or $\mathcal{F}_2$ is the initial filter, it is not counted in the sum.

Now, every decomposition $\mathcal{F} \simeq \mathcal{F}_1 + \mathcal{F}_2$ gives a morphism $\alpha : \mathcal{F} \to 2$ such that $\mathcal{F}_1 \simeq \alpha^*0$ and $\mathcal{F}_2 \simeq \alpha^*1$, in the obvious way. It follows that

$$D\mathcal{F} := \{\mathbf{y}\mathcal{F}_1 + \mathbf{y}\mathcal{F}_2 \mid \mathcal{F} \simeq \mathcal{F}_1 + \mathcal{F}_2\} \simeq \mathrm{im}(\mathrm{char}2)\,\mathcal{F} \ .$$

By sheaf semantics, for every K_1-closed sieve R on $\mathcal{F}$, $R \in D^r\,\mathcal{F}$ if and only if

$$\begin{aligned}&\text{for all } \alpha : \mathcal{G} \to \mathcal{F} \text{ and } S \in \Omega\mathcal{G},\ \text{ if } S \in D\mathcal{G} \text{ and, for all } \beta : \mathcal{H} \to \mathcal{G}\ ,\\ &\beta^*S \subseteq (\alpha\beta)^*R\ ,\ \text{ then } \alpha \in R\ .\end{aligned}$$

Since $S \subseteq \alpha^*R$ implies that, for all β, $\beta^*S \subseteq (\alpha\beta)^*R$, and by definition of D, this is equivalent to the simpler condition

$$\text{for all } \alpha : \mathcal{G} \to \mathcal{F},\ \mathcal{G} \simeq \mathcal{G}_1 + \mathcal{G}_2\ ,\ \text{ if } \alpha^*R \supseteq \mathbf{y}\mathcal{G}_1 + \mathbf{y}\mathcal{G}_2\ ,\ \text{ then } \alpha \in R\ .$$

Informally, this states that, if any pullback of R can be decomposed as a sum $\mathbf{y}\mathcal{G}_1 + \mathbf{y}\mathcal{G}_2$, then one of the factors $\mathcal{G}_1$, $\mathcal{G}_2$ is 0, so the pullback sieve is actually the maximal sieve. We say that R is *indecomposable* if it satisfies this condition.

Finally, we have that for all $S \in \Omega\mathcal{F}$, $S \in D^{rl}\mathcal{F}$ if and only if

$$\text{for all } \alpha : \mathcal{G} \to \mathcal{F} \text{ and } R \in \Omega\mathcal{G},\ \text{ if } R \in D^r\mathcal{G} \text{ and } \alpha^*S \subseteq R\ ,\ \text{ then } R \text{ is maximal}\ .$$

Again, this is just ordinary sheaf semantics, and we have already performed the possible simplifications. This states that the only K_1-closed, indecomposable sieves containing a pullback of S are the maximal sieves.

Let j be the characteristic map of D^{rl}. We will prove that all j-sheaves are also K-sheaves; by minimality of j, and the fact that m is K-dense, it will follow that $\mathrm{sh}_j(\mathcal{U}) \simeq \mathcal{N}$.

Clearly, the sieves on $\mathcal{F}$ generated by a coproduct pair $\{\mathcal{F}_i \to \mathcal{F}_1 + \mathcal{F}_2\}_{i=1}^2$, $\mathcal{F} \simeq \mathcal{F}_1 + \mathcal{F}_2$, are j-covering. For suppose that, for some $\alpha : \mathcal{G} \to \mathcal{F}$, $\alpha^*(\mathbf{y}\mathcal{F}_1 + \mathbf{y}\mathcal{F}_2) \simeq \mathbf{y}\alpha^*\mathcal{F}_1 + \mathbf{y}\alpha^*\mathcal{F}_2$ is contained in an indecomposable closed sieve R. Since $\mathfrak{F}\mathbf{Set}$ is coherent, finite coproducts are stable under pullback, and $\alpha^*\mathcal{F}_1 + \alpha^*\mathcal{F}_2 \simeq \alpha^*(\mathcal{F}_1 + \mathcal{F}_2) \simeq \mathcal{G}$; by definition, then, R contains $\mathrm{id}_{\mathcal{G}}$, hence is maximal.

Now, let X be a j-sheaf, and $\{x_k \in X\mathcal{G}_k\}_{k=1}^n$ a matching family of elements of X for a K-cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^n$. By the previous remark, there exists a unique amalgamation $x_{12} \in X(\mathcal{G}_1 + \mathcal{G}_2)$ of $\{x_1, x_2\}$.

Repeat with $\{x_{12}, x_3\}$, which is obviously matching for the j-covering $\{\mathcal{G}_1 + \mathcal{G}_2, \mathcal{G}_3\}$ of $\mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3$; this gives a uniquely determined $x_{123} \in X(\mathcal{G}_1 + \mathcal{G}_2 + \mathcal{G}_3)$. After $n-1$ steps, we obtain an element $x_{1\ldots n}$ in $X(\mathcal{G}_1 + \ldots + \mathcal{G}_n)$.

Since $\beta_1 + \ldots + \beta_n : \mathcal{G}_1 + \ldots + \mathcal{G}_n \twoheadrightarrow \mathcal{F}$ is a covering map, and X is also a K_1-sheaf, we can find a unique $x \in X\mathcal{F}$ such that $X(\beta_1 + \ldots + \beta_n)(x) = x_{1\ldots n}$. This is a necessarily unique amalgamation of the x_k, $k = 1, \ldots, n$. It follows that X is a K-sheaf; therefore, $\mathrm{sh}_j(\mathcal{U})$ is a subtopos of $\mathcal{N}$. The reverse inclusion is obvious, so this proves the proposition. □

Since in the last part of the proof we only used covering sieves that were already in D, and D^{rl} is the smallest Grothendieck topology containing D, we could have avoided its explicit construction; we only included it as an example of the use of sheaf semantics in proofs about topoi.

Notice also that the proof goes through for $\mathrm{Sh}(\mathfrak{F}\mathbf{C}, K_1)$ and $\mathrm{Sh}(\mathfrak{F}\mathbf{C}, K)$, given any coherent category $\mathbf{C}$.

So, we have defined the Grothendieck topoi $\mathcal{U}$ and $\mathcal{N}$, and established their topos-theoretic relation. In the next section, we will study their properties side by side, and see that they mirror those of the uniform Diller-Nahm and nonstandard Dialectica interpretations, respectively; so that the transition from single covering morphisms to finite families can be seen as the categorical analogue of herbrandisation.

3.3 Sheaf semantics of the filter topoi

Let $\mathcal{L}$ be a many sorted first order language, and suppose we have fixed an interpretation of $\mathcal{L}$ in **Set**. We call formulae of $\mathcal{L}$ *internal*, and denote them with small Greek letters. We also want the types of $\mathcal{L}$ to be closed under the clause

▷ if S is a type, then S^* is a type,

where S^* is meant to denote the type of finite sequences of elements of type S. We will borrow all the notation from Chapter 1 in handling finite sequences.

We will identify types, function and relation symbols of $\mathcal{L}$ with their interpretation in **Set**, and reserve the square bracket notation for the derived interpretations that we are now going to define. We will take advantage of this semantic overload, and say, for instance, that the type S is inhabited, or that it is infinite, if its interpretation in **Set** is; and also that a formula φ is *true*, if its interpretation is true in **Set**.

Let $\mathcal{L}_{\mathrm{st}}$ be the extension of $\mathcal{L}$ with a unary predicate symbol $\mathrm{st}_S \subseteq S$ for each type S. We denote formulae of $\mathcal{L}_{\mathrm{st}}$ with *capital* Greek letters. We will use abbreviations

$$\forall^{\mathrm{st}} x : S\, \Phi(x) := \forall x : S\, (\mathrm{st}_S(x) \to \Phi(x))\ ,$$
$$\exists^{\mathrm{st}} x : S\, \Phi(x) := \exists x : S\, (\mathrm{st}_S(x) \wedge \Phi(x))\ ,$$

as well as the defined predicate

$$\mathrm{hyper}_S(s) := \forall^{\mathrm{st}} x : S\, (x \in s)\ ,$$

for $s : S^*$, with the relative quantifiers

$$\forall^{\mathrm{hyp}} s : S^*\, \Phi(s) := \forall s : S^*\, (\mathrm{hyper}_S(s) \to \Phi(s))\ ,$$
$$\exists^{\mathrm{hyp}} s : S^*\, \Phi(s) := \exists s : S^*\, (\mathrm{hyper}_S(s) \wedge \Phi(s))\ .$$

We will often drop the subscript, and just write $\mathrm{st}(x)$, or $\mathrm{hyper}(s)$.

We simultaneously define interpretations of $\mathcal{L}_{\mathrm{st}}$ in $\mathcal{U}$ and in $\mathcal{N}$, as follows:

(i) for each type S, $[\![S]\!] := \mathbf{y}S$;

(ii) for each constant $c : S$, $[\![c]\!] := \mathbf{y}c : 1 \to \mathbf{y}S$;

(iii) for each function symbol $f : S_1, \ldots, S_n \to S$, $[\![f]\!] := \mathbf{y}f : \mathbf{y}(S_1 \times \ldots \times S_n) \to \mathbf{y}S$;

(iv) for each relation symbol $R \subseteq S_1, \ldots, S_n$ of $\mathcal{L}$, $[\![R]\!] := \mathbf{y}R \rightarrowtail \mathbf{y}(S_1 \times \ldots \times S_n)$;

(v) for each type S, $[\![\mathrm{st}_S]\!] := \Delta_1 S$ in $\mathcal{U}$, and $[\![\mathrm{st}_S]\!] := \Delta S$ in $\mathcal{N}$.

In particular, $[\![\mathrm{st}_{\mathbb{N}}]\!]$ is the *natural numbers object* both in $\mathcal{U}$ and in $\mathcal{N}$; in the latter, the larger sheaf $[\![\mathbb{N}]\!]$ is a nonstandard model of arithmetic.

The following, fundamental theorem connects the forcing semantics of internal formulae in $\mathcal{U}$ and in $\mathcal{N}$ with truth in the metatheory. It is an adaptation of [46, Theorem 1], which is itself an extension of [36, Lemma 2.1].

We write $\Vdash_{\mathcal{U}}$, $\Vdash_{\mathcal{N}}$ for the forcing relation in $\mathcal{U}$, and $\mathcal{N}$ respectively; and just $\Vdash$ for statements that are true of both.

Theorem 3.23. *Let $\varphi(x)$ be an internal formula, with free variable x of type S, and $(C, \mathcal{F}_I)$ a filter. For all $\alpha \in [\![S]\!]\mathcal{F}$,*

$$\mathcal{F} \Vdash_{\mathcal{N}} \varphi(\alpha)$$

if and only if there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$, it holds that $\varphi(\alpha(u))$.

If φ is also $\vee$-free, then the same condition is equivalent to

$$\mathcal{F} \Vdash_{\mathcal{U}} \varphi(\alpha) .$$

Proof. We proceed by induction on the structure of φ, as in Theorem 3.10; we will do the case of disjunction last, since it only works in $\mathcal{N}$. The clauses (i) and (ii) are trivial.

(iii) Suppose $\mathcal{F} \Vdash R(a)$, where we assumed, without loss of generality, that R depends explicitly on a single variable $x : S$. By sheaf semantics, this means that $a \in [\![R]\!]\mathcal{F} \subseteq [\![S]\!]\mathcal{F}$; by the interpretation we chose, this says that a, as a morphism $\mathcal{F} \to S$ in $\mathfrak{F}\mathbf{Set}$, factors through R.

But this is precisely the fact that, for some $i \in I$, the image of the base set $\mathcal{F}_i$ is contained in R, which was to be proved.

(iv) Suppose $\mathcal{F} \Vdash \varphi(\alpha) \wedge \psi(\alpha)$. By sheaf semantics, this is equivalent to $\mathcal{F} \Vdash \varphi(\alpha)$ and $\mathcal{F} \Vdash \psi(\alpha)$; by the induction hypothesis, there exist $i, j \in I$ such that, for all $u \in \mathcal{F}_i$, and all $v \in \mathcal{F}_j$,

$$\varphi(\alpha(u)) \wedge \psi(\alpha(v)) .$$

It now suffices to pick $k \in I$ such that $\mathcal{F}_k \subseteq \mathcal{F}_i \cap \mathcal{F}_j$.

(vi) Suppose first that there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$, $\varphi(\alpha(u))$ implies $\psi(\alpha(u))$. Take any morphism $\beta : (D, \mathcal{G}_J) \to \mathcal{F}$, and suppose $\mathcal{G} \Vdash \varphi(\alpha\beta)$. By the induction hypothesis, there exists $j \in J$ such that, for all $v \in \mathcal{G}_j$, $\varphi(\alpha\beta(v))$.

By continuity of β, there exists $j' \in J$ such that $\mathcal{G}_{j'} \subseteq \beta^*\mathcal{F}_i$. Taking $\mathcal{G}_k \subseteq \mathcal{G}_j \cap \mathcal{G}_{j'}$, we obtain, by the assumption, that for all $v \in \mathcal{G}_k$, $\psi(\alpha\beta(v))$. Again, by the induction hypothesis, $\mathcal{G} \Vdash \psi(\alpha\beta)$.

Conversely, suppose $\mathcal{F} \Vdash \varphi(\alpha) \to \psi(\alpha)$. Define a filter $(C, \mathcal{G}_I)$, with $\mathcal{G}_i := \{u \in \mathcal{F}_i \,|\, \varphi(\alpha(u))\}$; the inclusion $\imath : \mathcal{G} \rightarrowtail \mathcal{F}$ is continuous. By construction, $\mathcal{G} \Vdash \varphi(\alpha\imath)$; hence, by the assumption, $\mathcal{G} \Vdash \psi(\alpha\imath)$.

By the induction hypothesis, this means that there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$ such that $\varphi(\alpha(u))$ holds, $\psi(\alpha(u))$ holds. But this is precisely what we wanted to prove.

(vii) Suppose that $\mathcal{F} \Vdash_{\mathcal{N}} \exists y : T\, \varphi(\alpha, y)$; then, there exist a K-cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^n$, and elements $\zeta_k \in [\![T]\!]\mathcal{G}_k$, $k = 1, \ldots, n$, such that

$$\mathcal{G}_k \Vdash \varphi(\alpha\beta_k, \zeta_k)\ , \qquad k = 1, \ldots, n\ .$$

By the induction hypothesis, there exist base sets $\mathcal{G}_{k,j_k}$ of $\mathcal{G}_k$, $k = 1, \ldots, n$, such that, for all $v \in \mathcal{G}_{k,j_k}$, $\varphi(\alpha\beta_k(v), \zeta_k(v))$; implying $\exists y \in T\, \varphi(\alpha\beta_k(v), y)$.

Now, by the cover condition, there exists $i \in I$ such that

$$\mathcal{F}_i \subseteq \beta_1\mathcal{G}_{1,j_1} \cup \ldots \cup \beta_n\mathcal{G}_{n,j_n}\ ;$$

hence, for all $u \in \mathcal{F}_i$, $\exists y \in T\, \varphi(\alpha(u), y)$. For the same proof in $\mathcal{U}$, just take $n = 1$ in the cover.

Conversely, suppose that there is some $i \in I$ such that, for all $u \in \mathcal{F}_i$,

$$\exists y \in T\, \varphi(\alpha(u), y)\ .$$

Define a filter $(C, \mathcal{G}_I)$, with $\mathcal{G}_i := \{(u, y) \in \mathcal{F}_i \times T \mid \varphi(\alpha(u), y)\}$, for all $i \in I$. Then the first projection $\pi_1 : \mathcal{G} \twoheadrightarrow \mathcal{F}$ is covering, and, by the induction hypothesis,

$$\mathcal{G} \Vdash \varphi(\alpha\pi_1, \pi_2)\ .$$

By sheaf semantics, it follows that $\mathcal{F} \Vdash \exists y : T\, \varphi(\alpha, y)$.

(viii) Assume $\mathcal{F} \Vdash \forall y : T\, \varphi(\alpha, y)$; since $[\![T]\!]$ is representable, this is equivalent to

$$\mathcal{F} \times T \Vdash \varphi(\alpha\pi_1, \pi_2)\ .$$

Since the base sets of $\mathcal{F} \times T$ are of the form $\mathcal{F}_i \times T$, $i \in I$, by the induction hypothesis, there exists $i \in I$ such that, and for all $u \in \mathcal{F}_i$, for all $y \in T$, $\varphi(\alpha(u), y)$.

Conversely, assume there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$,

$$\forall y \in T\, \varphi(\alpha(u), y)\ .$$

Let $\beta : (D, \mathcal{G}_J) \to \mathcal{F}$, and $\zeta \in [\![T]\!]\mathcal{G}$. By continuity of β, there exists some $j \in J$ such that $\mathcal{G}_j \subseteq \beta^*\mathcal{F}_i$. Take $k \in J$, such that $\mathcal{G}_k \subseteq \mathcal{G}_j$, and ζ is defined on $\mathcal{G}_k$ as a morphism $\mathcal{G} \to T$. Then, for all $v \in \mathcal{G}_k$, $\varphi(\alpha\beta(v), \zeta(v))$. By the induction hypothesis,

$$\mathcal{G} \Vdash \varphi(\alpha\beta, \zeta)\ ;$$

and by arbitrariety of β and ζ it follows that $\mathcal{F} \Vdash \forall y : T\, \varphi(\alpha, y)$.

This concludes the proof for $\mathcal{U}$, and $\vee$-free formulae. In $\mathcal{N}$, we can also handle the case of disjunction.

(v) Suppose $\mathcal{F} \Vdash_{\mathcal{N}} \varphi(\alpha) \vee \psi(\alpha)$. By sheaf semantics, there exists a K-cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^n$ such that, for each $k = 1, \ldots, n$, either $\mathcal{G}_k \Vdash_{\mathcal{N}} \varphi(\alpha\beta_k)$ or $\mathcal{G}_k \Vdash_{\mathcal{N}} \psi(\alpha\beta_k)$. Using the induction hypothesis on the disjuncts that hold, we find base sets $\mathcal{G}_{k,j_k}$ of $\mathcal{G}_k$, $k = 1, \ldots, n$, such that for all k, and for all $v \in \mathcal{G}_{k,j_k}$,

$$\varphi(\alpha\beta_k(v)) \vee \psi(\alpha\beta_k(v))\ .$$

By the cover condition, there exists $i \in I$ such that

$$\mathcal{F}_i \subseteq \beta_1 \mathcal{G}_{1,j_1} \cup \ldots \cup \beta_n \mathcal{G}_{n,j_n} \ ;$$

therefore, for all $u \in \mathcal{F}_i$, $\varphi(\alpha(u)) \vee \psi(\alpha(u))$.

Conversely, assume there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$, $\varphi(\alpha(u)) \vee \psi(\alpha(u))$. We define filters $(C, \mathcal{G}_{1,I})$, $(C, \mathcal{G}_{2,I})$, such that, for all $i \in I$,

$$\mathcal{G}_{1,i} := \{u \in \mathcal{F}_i \,|\, \varphi(\alpha(u))\} \ ,$$
$$\mathcal{G}_{2,i} := \{u \in \mathcal{F}_i \,|\, \psi(\alpha(u))\} \ .$$

By construction, $\mathcal{G}_1 \Vdash_{\mathcal{N}} \varphi(\alpha)$ and $\mathcal{G}_2 \Vdash_{\mathcal{N}} \psi(\alpha)$; and, by the assumption, $\{\mathcal{G}_k \rightarrowtail \mathcal{F}\}_{k=1}^{2}$ is a K-cover. Hence, by sheaf semantics, $\mathcal{F} \Vdash_{\mathcal{N}} \varphi(\alpha) \vee \psi(\alpha)$.

With this, the proof is completed. □

Corollary 3.24 (Transfer theorem)**.** *Let φ be an internal sentence. Then φ is true if and only if $\Vdash_{\mathcal{N}} \varphi$. If φ is also $\vee$-free, then φ is true if and only if $\Vdash_{\mathcal{U}} \varphi$.*

The transfer theorem should be compared with Corollary 2.8 and Corollary 2.17: there, it was herbrandisation that allowed to extend conservativity to formulae with disjunctions; here, it is the passage from singleton to finite covers.

Theorem 3.23 says everything there is to know about internal formulae; we move on to the semantics of the standardness predicate.

Lemma 3.25. *Let $\mathcal{F}$ be a filter, S a type of $\mathcal{L}$, and $\alpha \in [\![S]\!]\mathcal{F}$. Then:*

(a) $\mathcal{F} \Vdash_{\mathcal{N}} \mathrm{st}_S(\alpha)$ if and only if there exist a K-cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^{n}$, and elements $x_1, \ldots, x_n \in S$, such that the diagrams

$$\begin{array}{ccc} \mathcal{G}_k & \xrightarrow{\beta_k} & \mathcal{F} \\ \Big\downarrow{\scriptstyle !} & & \Big\downarrow{\scriptstyle \alpha} \\ 1 & \xrightarrow{x_k} & S \end{array}$$

commute in $\mathfrak{F}\mathbf{Set}$, i.e. $\alpha\beta_k = x_k!$, for $k = 1, \ldots, n$.

(b) $\mathcal{F} \Vdash_{\mathcal{U}} \mathrm{st}_S(\alpha)$ if and only if there exist a covering map $\beta : \mathcal{G} \twoheadrightarrow \mathcal{F}$, and an element $x \in S$, such that $\alpha\beta = x!$ in $\mathfrak{F}\mathbf{Set}$.

Proof. Follows immediately from the interpretation chosen for the standardness predicate, and the description of $\Delta_1 S$ and ΔS. □

Lemma 3.26. *Let $\Phi(x, y)$ be an external formula, with free variables $x : S$ and $y : T$, $\mathcal{F}$ a filter, and $\alpha \in [\![S]\!]\mathcal{F}$. Then:*

(a) $\mathcal{F} \Vdash \forall^{\mathrm{st}} y : T\, \Phi(\alpha, y)$ if and only if, for all $y \in T$, $\mathcal{F} \Vdash \Phi(\alpha, y!)$;

(b) $\mathcal{F} \Vdash_{\mathcal{N}} \exists^{\mathrm{st}} y : T\, \Phi(\alpha, y)$ if and only if there exist a K-cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^{n}$, and elements $y_1, \ldots, y_n \in T$, such that

$$\mathcal{G}_k \Vdash_{\mathcal{N}} \Phi(\alpha\beta_k, y_k!) \ , \qquad k = 1, \ldots, n \ ,$$

or, equivalently, there exists $t \in T^*$ such that $\mathcal{F} \Vdash \exists y \in t!\, \Phi(\alpha, y)$;

(c) $\mathcal{F} \Vdash_{\mathcal{U}} \exists^{\mathrm{st}} y : T\, \Phi(\alpha, y)$ *if and only if there exists* $y \in T$ *such that* $\mathcal{F} \Vdash_{\mathcal{U}} \Phi(\alpha, y!)$.

Proof. For item (a): the direction from left to right is immediate. Conversely, let $\beta : \mathcal{G} \to \mathcal{F}$, and $\zeta \in [\![T]\!]\mathcal{G}$ such that $\mathcal{G} \Vdash \mathrm{st}(\zeta)$. By the previous lemma, there exist a cover $\{\gamma_k : \mathcal{H}_k \to \mathcal{G}\}_{k=1}^n$, and elements $y_1, \ldots, y_n \in T$ ($n = 1$ in case we are in $\mathcal{U}$), such that $\zeta\gamma_k = y_k!$, for $k = 1, \ldots, n$. By the assumption, and monotonicity of the forcing relation,

$$\mathcal{H}_k \Vdash \Phi(\alpha\beta\gamma_k, \zeta\gamma_k) \ , \qquad k = 1, \ldots, n \ .$$

It follows by local character that $\mathcal{G} \Vdash \Phi(\alpha\beta, \zeta)$, which was to be proved.

For item (b): the direction from right to left is immediate. From left to right, suppose $\mathcal{F} \Vdash_{\mathcal{N}} \exists^{\mathrm{st}} y : T\, \Phi(\alpha, y)$; then there exist a cover $\{\beta_k : \mathcal{G}_k \to \mathcal{F}\}_{k=1}^n$, and elements $\zeta_k \in [\![T]\!]\mathcal{G}_k$, such that

$$\mathcal{G}_k \Vdash_{\mathcal{N}} \mathrm{st}(\zeta_k) \wedge \Phi(\alpha\beta_k, \zeta_k) \ , \qquad k = 1, \ldots, n \ .$$

By the previous lemma, then, for all $k = 1, \ldots, n$, there exist a cover $\{\gamma_{k\ell} : \mathcal{H}_{k\ell} \to \mathcal{G}_k\}_{\ell=1}^{m_k}$, and elements $y_{k1}, \ldots, y_{km_k} \in T$ such that $\zeta_k\gamma_{k\ell} = y_{k\ell}!$, $k = 1, \ldots, n$, $\ell = 1, \ldots, m_k$. The thesis follows by taking the composition

$$\{\beta_k\gamma_{k\ell} : \mathcal{H}_{k\ell} \to \mathcal{G}_k \to \mathcal{F}\}_{\ell=1,\ldots,m_k}^{k=1,\ldots,n}$$

of all covers, which is still a cover.

For item (c): again, the direction from right to left is immediate. From left to right, suppose $\mathcal{F} \Vdash_{\mathcal{U}} \exists^{\mathrm{st}} y : T\, \Phi(\alpha, y)$; then there is a covering map $\beta : \mathcal{G} \twoheadrightarrow \mathcal{F}$, and an element $\zeta \in [\![T]\!]\mathcal{G}$, such that

$$\mathcal{G} \Vdash_{\mathcal{U}} \mathrm{st}(\zeta) \wedge \Phi(\alpha\beta, \zeta) \ .$$

By the previous lemma, there exist another covering map $\gamma : \mathcal{H} \twoheadrightarrow \mathcal{G}$, and an element $y \in T$, such that $\zeta\gamma = y!$. Then

$$\mathcal{H} \Vdash_{\mathcal{U}} \Phi(\alpha\beta\gamma, y!) \ ,$$

and, by local character, $\mathcal{F} \Vdash_{\mathcal{U}} \Phi(\alpha, y!)$. □

With these lemmata, we are able to prove that the simple axioms that we imposed on the standardness predicate in Chapter 1 hold both in $\mathcal{N}$ and in $\mathcal{U}$. That the predicate respects equality is immediate; that closed terms are standard amounts, in this context, to the fact that, for all types S and elements $\alpha \in [\![S]\!]1$,

$$\Vdash \mathrm{st}_S(\alpha) \ ,$$

as any morphism $\alpha : 1 \to S$ is obviously constant.

Proposition 3.27. *For all types* S, T, *the following statement is true in* $\mathcal{N}$ *and* $\mathcal{U}$*:*

$$\forall^{\mathrm{st}} f : S \to T\, \forall^{\mathrm{st}} x : S\, \mathrm{st}_T(f(x)) \ .$$

Proof. By Lemma 3.26, $\Vdash \forall^{\mathrm{st}} f : S \to T\, \forall^{\mathrm{st}} x : S\, \mathrm{st}_T(f(x))$ if and only if, for all $f \in (S \to T)$, and $x \in S$,

$$\Vdash \mathrm{st}_T(f!(x!)) \ .$$

But $[\![f!(x!)]\!] = [\![f(x)!]\!]$, and the latter is clearly standard. □

Proposition 3.28. *The external induction schema* $\mathsf{IA}^{\mathrm{st}}$ *holds in* $\mathcal{N}$ *and* $\mathcal{U}$.

Proof. Let $\Phi(x, n)$ be an external formula, with $x : S$ and $n : \mathbb{N}$, $\mathcal{F}$ a filter, and $\alpha \in [\![S]\!]\mathcal{F}$. Suppose

$$\mathcal{F} \Vdash \Phi(\alpha, 0!) \wedge \forall^{\mathrm{st}} n : \mathbb{N} \, (\Phi(\alpha, n) \to \Phi(\alpha, \mathrm{S}n)) \; .$$

Then, by Lemma 3.26, we have that $\mathcal{F} \Vdash \Phi(\alpha, 0!)$ and that, for all $n \in \mathbb{N}$, $\mathcal{F} \Vdash \Phi(\alpha, n!)$ implies $\mathcal{F} \Vdash \Phi(\alpha, \mathrm{S}n!)$. By induction in the metatheory, we obtain that, for all $n \in \mathbb{N}$,

$$\mathcal{F} \Vdash \Phi(\alpha, n!) \; ,$$

so, again by the semantics of the external quantifiers, $\mathcal{F} \Vdash \forall^{\mathrm{st}} n : \mathbb{N} \, \Phi(\alpha, n)$. □

Lemma 3.26 has also the following easy consequence.

Corollary 3.29. *Let* φ *be an internal formula.*

(a) $\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} x : S \, \exists^{\mathrm{st}} y : T \, \varphi(x, y)$ *if and only if it is true that* $\forall x \in S \, \exists y \in T \, \varphi(x, y)$.

(b) If φ *is also* $\vee$*-free,*

$$\Vdash_{\mathcal{U}} \exists^{\mathrm{st}} x_1 : S_1 \, \forall^{\mathrm{st}} y_1 : T_1 \, \ldots \exists^{\mathrm{st}} x_n : S_n \, \forall^{\mathrm{st}} y_n : T_n \, \varphi(x_1, y_1, \ldots, x_n, y_n)$$

if and only if it is true that

$$\exists x_1 \in S_1 \, \forall y_1 \in T_1 \, \ldots \exists x_n \in S_n \, \forall y_n \in T_n \, \varphi(x_1, y_1, \ldots, x_n, y_n) \; .$$

Finally, we give the following result without proof. If φ is an internal formula, let φ^{st} be the formula of $\mathcal{L}_{\mathrm{st}}$ where all quantifiers have been restricted to standard elements.

Proposition 3.30. *Suppose* φ *is an internal formula, which is built up from quantifier-free formulae using only* $\wedge, \vee, \exists$ *and* $\forall$. *Then, if* φ *is true,* $\Vdash_{\mathcal{N}} \varphi^{\mathrm{st}}$. *If* φ *is also* $\vee$*-free, then* $\Vdash_{\mathcal{U}} \varphi^{\mathrm{st}}$.

Proof. See [45, Theorem 3.6]. □

We have, by now, a good picture of the semantics of first order logic in the filter topoi. In the next section, we will start proving that the characteristic principles of uniform Diller-Nahm and nonstandard Dialectica hold in $\mathcal{U}$ and $\mathcal{N}$, respectively.

3.4 Characteristic principles

For the results in this section, we cannot take much credit, since a characterisation of first order logic in the topoi $\mathrm{Sh}(\mathfrak{F}\mathbf{C}, K)$, with $\mathbf{C}$ coherent, has already been provided by Butz [10, Proposition 4.5], albeit with a different aim and formalism. The choice of principles, however, is different, due to our focus on nonstandard arithmetic; and it allows us to see herbrandisation "in action", by providing proofs for $\mathcal{U}$ alongside $\mathcal{N}$.

We start from the truly "nonstandard" principles, sequence overspill and underspill.

Proposition 3.31. *(a) The principle* OS^* *holds in* $\mathcal{N}$.

(b) The principle $\mathsf{OS}^*_{\vee}$ *holds in* $\mathcal{U}$.

Proof. For item (a), let $\varphi(y,s)$ be an internal formula, and, for item (b), an internal, $\vee$-free formula, with variables $y : T$, and $s : S^*$; from here the proof proceeds in the same way for both cases.

Let $(C, \mathcal{F}_I)$ be any filter, $\alpha \in [\![T]\!]\mathcal{F}$, and assume

$$\mathcal{F} \Vdash \forall^{\mathrm{st}} s : S^* \, \varphi(\alpha, s) \ .$$

By Lemma 3.26, for all $s \in S^*$, $\mathcal{F} \Vdash \varphi(\alpha, s!)$; by transfer (Theorem 3.23), for all $s \in S^*$, there exists $i \in I$ such that, for all $u \in \mathcal{F}_i$,

$$\varphi(\alpha(u), s) \ .$$

Define a filter $(C \times S^*, \mathcal{G}_{I \times S^*})$, as follows: for all $i \in I$, $t \in S^*$,

$$\mathcal{G}_{(i,t)} := \{(u,s) \,|\, u \in \mathcal{F}_i \wedge t \subseteq s \wedge \varphi(\alpha(u), s)\} \ .$$

The filter condition is easily checked: given $\mathcal{G}_{(i,t)}$, $\mathcal{G}_{(j,t')}$, pick $k \in I$ such that $\mathcal{F}_k \subseteq \mathcal{F}_i \cap \mathcal{F}_j$, and $t'' := t \cdot t'$; then, $\mathcal{G}_{(k,t'')} \subseteq \mathcal{G}_{(i,t)} \cap \mathcal{G}_{(j,t')}$.

The projections $\pi_1 : \mathcal{G} \to \mathcal{F}$, and $\pi_2 : \mathcal{G} \to S^*$ are clearly continuous. We now check

$$\mathcal{G} \Vdash \mathrm{hyper}(\pi_2) \ .$$

By definition, this means $\mathcal{G} \Vdash \forall^{\mathrm{st}} x : S \, (x \in \pi)$; equivalently, for all $x \in S$, $\mathcal{G} \Vdash x! \in \pi$. By transfer, it suffices to prove that, for all $x \in S$, there exists $(i,t) \in I \times S^*$, such that for all $u \in \mathcal{F}_i$, and $s \supseteq t$, it holds that $x \in s$; so we can take $t := \langle x \rangle$, and $i \in I$ arbitrary.

Furthermore, $\mathcal{G} \Vdash \varphi(\alpha\pi_1, \pi_2)$ holds by construction. Hence, in order to derive that

$$\mathcal{F} \Vdash \exists s : S^* \, (\mathrm{hyper}(s) \wedge \varphi(\alpha, s)) \ ,$$

it remains to show that π_1 is covering. Let $\mathcal{G}_{(i,t)}$ be an arbitrary base set of $\mathcal{G}$. By the assumption, we can find $j \in I$ such that, for all $u \in \mathcal{F}_j$, $\varphi(\alpha(u), t)$; then, if we choose $k \in I$ such that $\mathcal{F}_k \subseteq \mathcal{F}_i \cap \mathcal{F}_j$, we have that

$$\mathcal{F}_k \subseteq \pi_1 \, \mathcal{G}_{(i,t)} \ .$$

This concludes the proof. $\square$

Lemma 3.32. *Let $\Phi(x)$ be an external formula, $x : S$, such that*

$$\Vdash \exists x : S \, \Phi(x) \ . \tag{3.2}$$

Then

$$\forall y : T \, \big(\forall x : S \, (\Phi(x) \to \varphi(y,x)) \to \exists^{\mathrm{st}} x : S \, \varphi(y,x)\big)$$

holds in $\mathcal{N}$ for all internal formulae, and in $\mathcal{U}$ for all internal, $\vee$-free formulae $\varphi(y,x)$, with $y : T$.

Proof. Let $\mathcal{F}$ be any filter, $\varphi(y,x)$ an internal formula ($\vee$-free in $\mathcal{U}$), $y : T$, and $\alpha \in [\![T]\!]\mathcal{F}$. Suppose $\mathcal{F} \Vdash \forall x : S \, (\Phi(x) \to \varphi(\alpha, x))$; equivalently,

$$\mathcal{F} \times S \Vdash \Phi(\pi_2) \to \varphi(\alpha\pi_1, \pi_2) \ . \tag{3.3}$$

Assume (3.2). Then, there exist a cover $\{\mathcal{G}_k \to 1\}_{k=1}^{n}$ ($n = 1$ if we are in $\mathcal{U}$), and elements $\sigma_k \in [\![S]\!]\mathcal{G}_k$, $k = 1, \ldots, n$, such that

$$\mathcal{G}_k \Vdash \Phi(\sigma_k) \;, \quad k = 1, \ldots, n \;.$$

By our interpretation of the type S, the σ_k correspond to morphisms $\sigma_k : \mathcal{G}_k \to S$ in $\mathfrak{F}\mathbf{Set}$; by monotonicity, we obtain

$$\mathcal{F} \times \mathcal{G}_k \Vdash \Phi(\sigma_k \pi_2) \;,$$

which, by the commutativity of the diagrams

$$\begin{array}{ccc} \mathcal{F} \times \mathcal{G}_k & \xrightarrow{\pi_2} & \mathcal{G}_k \\ \big\downarrow{\scriptstyle \mathrm{id} \times \sigma_k} & & \big\downarrow{\scriptstyle \sigma_k} \\ \mathcal{F} \times S & \xrightarrow{\pi_2} & S \;, \end{array}$$

for $k = 1, \ldots, n$, is the same as $\mathcal{F} \times \mathcal{G}_k \Vdash \Phi(\pi_2(\mathrm{id} \times \sigma_k))$.

Therefore, from (3.3), it follows, by monotonicity, that

$$\mathcal{F} \times \mathcal{G}_k \Vdash \varphi(\alpha\pi_1, \sigma_k \pi_2) \;;$$

by transfer, for all $k = 1, \ldots, n$, there exist base sets $\mathcal{F}_{i_k}$ of $\mathcal{F}$, $\mathcal{G}_{k,j_k}$ of $\mathcal{G}_k$, such that for all $u \in \mathcal{F}_{i_k}$, and $v \in \mathcal{G}_{k,j_k}$, it holds that $\varphi(\alpha(u), \sigma_k(v))$.

Now, by the cover condition, $\mathcal{G}_{1,j_1} \cup \ldots \cup \mathcal{G}_{n,j_n}$ contains at least one element, so there exists some $x \in \sigma_1 \mathcal{G}_{1,j_1} \cup \ldots \cup \sigma_n \mathcal{G}_{n,j_n}$. For such an x, taking $\mathcal{F}_i \subseteq \mathcal{F}_{i_1} \cap \ldots \cap \mathcal{F}_{i_n}$, and using transfer,

$$\mathcal{F} \Vdash \varphi(\alpha, x!) \;;$$

hence $\mathcal{F} \Vdash \exists^{\mathrm{st}} x : S\, \varphi(\alpha, x)$, which was to be proved. $\square$

Proposition 3.33. *(a) The principle* US^* *holds in* $\mathcal{N}$.

(b) The principle $\mathsf{US}^*_\vee$ *holds in* $\mathcal{U}$.

Proof. Follows from the previous lemma, by taking $\Phi(s) := \mathrm{hyper}(s)$, and using for condition (3.2) the fact that, by sequence overspill, hyperfinite enumerations of any type exist both in $\mathcal{N}$ and in $\mathcal{U}$. $\square$

Given sequence overspill and underspill, one can adapt the proofs of Chapter 1 to show that other principles, including idealisation and the herbrandised generalised Markov's principle, hold in $\mathcal{N}$ and in $\mathcal{U}$. However, one should pay attention to the fact that, while finite types were all inhabited, and actually had infinite elements, in this context a type S can be finite, or even empty. So, for instance, the implication $\mathsf{OS}^* \to \mathsf{OS}$ only holds for types with infinite elements: by definition of standardness, a finite set has only standard elements.

Next, we deal with the characteristic principles of nonstandard Dialectica, and uniform Diller-Nahm, whose validity in the filter topoi is independent of the metatheory.

Proposition 3.34. *(a) The principle* NCR *holds in* $\mathcal{N}$.

(b) The principle NU *holds in* $\mathcal{U}$.

Proof. Let $\mathcal{F}$ be any filter, $\Phi(z,x,y)$ an external formula, $x : S, y : T, z : U$, and $\alpha \in [\![U]\!]\mathcal{F}$. Assume $\mathcal{F} \Vdash \forall y : T\, \exists^{\mathrm{st}} x : S\, \Phi(\alpha, x, y)$, or, equivalently,

$$\mathcal{F} \times T \Vdash \exists^{\mathrm{st}} x : S\, \Phi(\alpha\pi_1, x, \pi_2)\ .$$

By the semantics of the $\exists^{\mathrm{st}}$ quantifier in $\mathcal{N}$, this means that there exists $s \in S^*$ such that

$$\mathcal{F} \times T \Vdash_{\mathcal{N}} \exists x \in s!\, \Phi(\alpha\pi_1, x, \pi_2)\ ;$$

equivalently, since $s! = s!\pi_1$, $\mathcal{F} \Vdash_{\mathcal{N}} \forall y : T\, \exists x \in s!\, \Phi(\alpha, x, y)$. Therefore,

$$\mathcal{F} \Vdash_{\mathcal{N}} \exists^{\mathrm{st}} s : S^*\, \forall y : T\, \exists x \in s!\, \Phi(\alpha, x, y)\ .$$

The proof for $\mathcal{U}$ is the same, with the semantics of $\exists^{\mathrm{st}}$ providing a single $x \in S$ instead of a finite sequence, and leads to

$$\mathcal{F} \Vdash_{\mathcal{U}} \exists^{\mathrm{st}} x : S\, \forall y : T\, \Phi(\alpha, x, y)\ . \qquad \square$$

The next proof is a variant of one by Butz [10]. It utilises the following, general result about Grothendieck topoi.

Lemma 3.35. *Let $(\mathbf{C}, J)$ be a site. A set $\{f_i : C_i \to C\}_{i\in I}$ of morphisms of $\mathbf{C}$ is J-covering if and only if the set $\{\mathbf{ay}f_i : \mathbf{ay}C_i \to \mathbf{ay}C\}_{i\in I}$ is jointly epimorphic in $\mathrm{Sh}(\mathbf{C}, J)$.*

Proof. See [34, Corollary III.7.7]. $\square$

Proposition 3.36. *(a) The principle $\mathsf{HIP}^{\mathrm{st}}_{\forall}$ holds in $\mathcal{N}$.*

(b) The principle $\mathsf{IP}^{\mathrm{st}}_{\forall\vee}$ holds in $\mathcal{U}$.

Proof. Let $\mathcal{F}$ be any filter, $\varphi(z,x)$ an internal formula ($\vee$-free in $\mathcal{U}$), $\Psi(z,y)$ an external formula, with $x : S$, $y : T$, $z : U$, and $\alpha \in [\![U]\!]\mathcal{F}$. Suppose

$$\mathcal{F} \Vdash \forall^{\mathrm{st}} x : S\, \varphi(\alpha, x) \to \exists^{\mathrm{st}} y : T\, \Psi(\alpha, y)\ .$$

By the semantics of first order logic in a Heyting category, this is equivalent to

$$\alpha^*[\![\forall^{\mathrm{st}} x\, \varphi(z,x)]\!] \le \alpha^*[\![\exists^{\mathrm{st}} y : T\, \Psi(z,y)]\!]$$

in $\mathrm{Sub}(\mathbf{y}\mathcal{F})$. By the semantics of the $\forall^{\mathrm{st}}$ predicate, we can write

$$\alpha^*[\![\forall^{\mathrm{st}} x\, \varphi(z,x)]\!] = \alpha^* \bigwedge_{x\in S} [\![\varphi(z, x!)]\!]\ ;$$

and, by the suitable transfer theorem, for all $x \in S$,

$$[\![\varphi(z, x!)]\!] = \mathbf{y}\{z \in U \mid \varphi(z,x)\}\ .$$

Since the Yoneda embedding preserves and reflects all limits, we obtain

$$\alpha^*[\![\forall^{\mathrm{st}} x\, \varphi(z,x)]\!] = \mathbf{y}\Big(\alpha^* \bigwedge_{x\in S} \{z \in U \mid \varphi(z,x)\}\Big) =: \mathbf{y}\mathcal{H}\ .$$

For the consequence, we have, by the semantics of $\exists^{\mathrm{st}}$ in $\mathcal{N}$ (the case of $\mathcal{U}$ is similar, and easier), that

$$\begin{aligned}\alpha^*[\![\exists^{\mathrm{st}} y : T\, \Psi(z,y)]\!] = \alpha^* \bigvee_{t \in T^*} [\![\exists y \in t!\, \Psi(z,y)]\!] = \\ = \bigvee_{t \in T^*} \alpha^*[\![\exists y \in t!\, \Psi(z,y)]\!] =: \bigvee_{t \in T^*} F_t \ ,\end{aligned}$$

where we also used that unions are stable under pullback. Thus, there is a monomorphism $m : \mathbf{y}\mathcal{H} \rightarrowtail \bigvee_{t \in T^*} F_t$.

Let $\imath_t : F_t \rightarrowtail \bigvee_{t \in T^*} F_y$ be the inclusions of the F_t in their union, for all $t \in T^*$, and consider the pullback diagrams

$$\begin{tikzcd} m^* F_t \arrow[r, tail] \arrow[d, tail] & F_t \arrow[d, tail, "\imath_t"] \\ \mathbf{y}\mathcal{H} \arrow[r, tail, "m"] & \bigvee_{t \in T^*} F_t \ . \end{tikzcd}$$

Now, we use the fact that each $m^* F_t$ can be covered with a family of representable sheaves, to obtain a family $\{f_t : \mathbf{y}\mathcal{G}_t \to \mathbf{y}\mathcal{H}\}_{t \in T^*}$ of morphisms, such that each $m f_t$ factors through a *single* F_t.

Moreover, since the $\{\imath_t : F_t \rightarrowtail \bigvee_{t \in T^*} F_t\}_{y \in T}$ jointly cover $\bigvee_{t \in T^*} F_t$, and in a Heyting category all epimorphisms are stable under pullback [34, Proposition IV.7.3], the family $\{f_t : \mathbf{y}\mathcal{G}_t \to \mathbf{y}\mathcal{H}\}_{t \in T^*}$ is jointly epimorphic over $\mathbf{y}\mathcal{H}$.

By the previous lemma, we can extract from it a family of the form $\{\mathbf{y}\beta_k : \mathbf{y}\mathcal{G}_k \to \mathbf{y}\mathcal{H}\}_{k=1}^n$, where $\{\beta_k : \mathcal{G}_k \to \mathcal{H}\}_{k=1}^n$ is a K-cover in $\mathfrak{F}\mathbf{Set}$ ($n = 1$ in $\mathcal{U}$). Let $t := t_1 \cdot \ldots \cdot t_n$, such that $\mathbf{y}\beta_k$ factors through F_{t_k}, $k = 1, \ldots, n$. Then,

$$\mathbf{y}\mathcal{H} = \alpha^*[\![\forall^{\mathrm{st}} x\, \varphi(z,x)]\!] \leq \alpha^* \bigvee_{k=1}^{n} [\![\exists y \in t_k!\, \Psi(z,y)]\!] = \alpha^*[\![\exists y \in t!\, \Psi(z,y)]\!] \ .$$

Translating back to forcing semantics, this is precisely the statement that

$$\mathcal{F} \Vdash_{\mathcal{N}} \forall^{\mathrm{st}} x : S\, \varphi(\alpha,x) \to \exists y \in t!\, \Psi(\alpha,y) \ ,$$

from which it follows that

$$\mathcal{F} \Vdash_{\mathcal{N}} \exists^{\mathrm{st}} t : T^* \left(\forall^{\mathrm{st}} x : S\, \varphi(\alpha,x) \to \exists^{\mathrm{st}} y \in t\, \Psi(\alpha,y)\right) \ .$$

Similar reasoning about $\mathcal{U}$ leads to

$$\mathcal{F} \Vdash_{\mathcal{U}} \exists^{\mathrm{st}} y : T \left(\forall^{\mathrm{st}} x : S\, \varphi(\alpha,x) \to \Psi(\alpha,y)\right) \ . \qquad \square$$

So far, we used no principles whose constructive status is controversial, neither in the construction of the model, nor in our proofs. However, for our last pair of characteristic principles to hold, we must require that the *axiom of choice* hold in the metatheory.

On one hand, this does not necessarily spoil constructiveness: as we stated earlier, it is possible to build $\mathcal{N}$ and $\mathcal{U}$ within Martin-Löf type theory, which is constructive, predicative, *and* has a version of the axiom of choice. On the other hand, Martin-Löf type theory has intensional equality; in an *extensional* theory, where it is possible to formalise Diaconescu's theorem, the axiom of choice leads to classicality.

The systems of arithmetic that we considered in the first two chapters, on the other hand, were extensional; so there appears to be a slight discrepancy between the proof-theoretic systems, and their proposed models. This is an issue that deserves to be further investigated.

Proposition 3.37. *Suppose that the axiom of choice holds in the metatheory.*

(a) The principle $\mathsf{HAC}^{\mathrm{st}}$ *then holds in* $\mathcal{N}$*.*

(b) The principle $\mathsf{AC}^{\mathrm{st}}$ *then holds in* $\mathcal{U}$*.*

Proof. Let $\mathcal{F}$ be any filter, $\Phi(z, x, y)$ an external formula, $x : S, y : T, z : U$, and $\alpha \in [\![U]\!]\mathcal{F}$. Assume

$$\mathcal{F} \Vdash \forall^{\mathrm{st}} x : S \, \exists^{\mathrm{st}} y : T \, \Phi(\alpha, x, y) \ ;$$

by Lemma 3.26, this means in $\mathcal{N}$ that, for all $x \in S$, there exists $t \in T^*$ such that

$$\mathcal{F} \Vdash_{\mathcal{N}} \exists y \in t! \, \Phi(\alpha, x!, y) \ .$$

With the axiom of choice, we can find a function $f \in S \to T^*$ such that, for all $x \in S$,

$$\mathcal{F} \Vdash_{\mathcal{N}} \exists y \in f(x)! \, \Phi(\alpha, x!, y) \ .$$

Since $[\![f(x)!]\!] = [\![f!(x!)]\!]$, it follows that $\mathcal{F} \Vdash_{\mathcal{N}} \exists^{\mathrm{st}} f : S \to T^* \, \forall^{\mathrm{st}} x : S \, \exists y \in f(x) \, \Phi(\alpha, x, y)$. The proof for $\mathcal{U}$ is completely analogous. □

In fact, a herbrandised version of the axiom of choice would suffice for $\mathcal{N}$; but that would be a strange axiom to have in one's metatheory. The condition is necessary to a certain extent, for $\mathsf{HAC}^{\mathrm{st}}$ implies a herbrandised axiom of choice - call it HAC - in **Set**: suppose

$$\forall x \in S \, \exists y \in T \, \varphi(x, y) \ .$$

By Corollary 3.29, it follows that $\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} x : S \, \exists^{\mathrm{st}} y : T \, \varphi(x, y)$. If $\mathsf{HAC}^{\mathrm{st}}$ holds in $\mathcal{N}$, we can deduce

$$\Vdash_{\mathcal{N}} \exists^{\mathrm{st}} f : S \to T^* \, \forall^{\mathrm{st}} x : S \, \exists y \in f(x) \, \varphi(x, y) \ ;$$

and, applying the transfer theorem again, we obtain

$$\exists f \in S \to T^* \, \forall x \in S \, \exists y \in f(x) \, \varphi(x, y)$$

in **Set**. In the same way, the validity of $\mathsf{AC}^{\mathrm{st}}$ in $\mathcal{U}$ implies the axiom of choice for $\vee$-free formulae, $\mathsf{AC}_{\vee}$, in **Set**.

The relevant link between the internal and external world, here, is the validity of the transfer rules $\mathsf{TR}_{\forall\exists}$, and $\mathsf{TR}_{\forall\exists\vee}$, in $\mathcal{N}$ and $\mathcal{U}$ respectively, ensured by Corollary 3.29. If the systems of Chapter 2 proved to be closed under these rules, we would immediately obtain a proof that E-HA$^{\omega*}$ + HAC, and E-HA$^{\omega*}_{\vee}$ + $\mathsf{AC}_{\vee}$, are conservative extensions of their base systems.

The transfer rules can be used to rule out unconstrained validity of other principles in $\mathcal{N}$ and $\mathcal{U}$, as in the following examples.

Example 3.38. Let $T(s)$ be a binary tree, i.e. an internal formula on binary sequences such that

1. $T(\langle\rangle)$ holds, and
2. $\forall n, m \in \mathbb{N}\, \forall s \in 2^{\mathbb{N}}\, (T(\bar{s}m) \wedge n \leq m) \to T(\bar{s}n)$, where $\bar{s}n := \langle s_0 \rangle \cdot \ldots \cdot \langle s_{n-1} \rangle$.

The *fan theorem* is the statement that for any such T, if, for all sequences $s \in 2^{\mathbb{N}}$, there exists $n \in \mathbb{N}$ such that $\neg\, T(\bar{s}n)$, then there exists some $n \in \mathbb{N}$ that works uniformly for all $s \in 2^{\mathbb{N}}$.

We consider the following, external version of the fan theorem:

$$\mathsf{FAN}^{\mathrm{st}}: \quad \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \exists^{\mathrm{st}} n : \mathbb{N}\, \neg\, T(\bar{s}n) \to \exists^{\mathrm{st}} n : \mathbb{N}\, \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \neg\, T(\bar{s}n)\ .$$

We claim that, if $\mathsf{FAN}^{\mathrm{st}}$ holds in $\mathcal{N}$, then the fan theorem holds in the metatheory. For suppose that, for all $s \in 2^{\mathbb{N}}$, there exists $n \in \mathbb{N}$ such that $\neg\, T(\bar{s}n)$. By transfer,

$$\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \exists^{\mathrm{st}} n : \mathbb{N}\, \neg\, T(\bar{s}n)\ ;$$

and, if $\mathsf{FAN}^{\mathrm{st}}$ holds, we deduce

$$\Vdash_{\mathcal{N}} \exists^{\mathrm{st}} n : \mathbb{N}\, \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \neg\, T(\bar{s}n)\ .$$

This means that there exists a finite sequence t of natural numbers, such that

$$\Vdash_{\mathcal{N}} \exists n \in t!\, \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \neg\, T(\bar{s}n)\ .$$

By condition 2 on binary trees, we have that, if $\neg\, T(\bar{s}n)$ and $m \geq n$, then also $\neg\, T(\bar{s}m)$; therefore, picking $\tilde{n} := \max\{t_0, \ldots, t_{|t|-1}\}$, we are sure that

$$\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} s : 2^{\mathbb{N}}\, \neg\, T(\bar{s}\tilde{n}!)\ .$$

By transfer, for all $s \in 2^{\mathbb{N}}$, $\neg\, T(\bar{s}\tilde{n})$, and we have proved the fan theorem.

Example 3.39. Since US^* holds in $\mathcal{N}$, by Proposition 1.36, $\mathsf{LLPO}^{\mathrm{st}}$ holds as well. Of course, that is not true of $\mathcal{U}$, as the principle is incompatible with the disjunction property.

A stronger principle than LLPO_0 is

$$\mathsf{LPO}_0: \quad \forall n : \mathbb{N}\, \varphi(n) \vee \exists n : \mathbb{N}\, \neg\, \varphi(n)\ ,$$

the *limited principle of omniscience*, where $\varphi(n)$ ranges over *decidable* formulae of arithmetic. It was shown by Palmgren [46, Proposition 14] that if the following, external version of LPO_0,

$$\mathsf{LPO}_0^{\mathrm{st}}: \quad \forall^{\mathrm{st}} n : \mathbb{N}\, \varphi(n) \vee \exists^{\mathrm{st}} n : \mathbb{N}\, \neg\, \varphi(n)\ ,$$

holds in $\mathcal{N}$, then LPO_0 holds in the metatheory. For let $\varphi(n)$ be any internal, decidable formula, and suppose

$$\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} n : \mathbb{N}\, \varphi(n) \vee \exists^{\mathrm{st}} n : \mathbb{N}\, \neg\, \varphi(n)\ .$$

Then, there exist a K-cover $\{\mathcal{G}_k \to 1\}_{k=1}^m$, and elements $n_1, \ldots, n_m \in \mathbb{N}$, such that, for each $k = 1, \ldots, m$, either $\mathcal{G}_k \Vdash_{\mathcal{N}} \forall^{\mathrm{st}} n : \mathbb{N}\, \varphi(n)$, or $\mathcal{G}_k \Vdash_{\mathcal{N}} \neg\, \varphi(n_k!)$.

For all $k = 1, \dots, m$, we can check whether $\varphi(n_k)$ holds. If it holds for all k, then

$$\mathcal{G}_k \Vdash_{\mathcal{N}} \forall^{\mathrm{st}} n : \mathbb{N}\, \varphi(n)\ , \quad k = 1, \dots, m,$$

and, by local character, $\Vdash_{\mathcal{N}} \forall^{\mathrm{st}} n : \mathbb{N}\, \varphi(n)$; by transfer, for all $n \in \mathbb{N}$, $\varphi(n)$. Otherwise, $\neg\, \varphi(n_k)$ for some k, so there exists $n \in \mathbb{N}$ such that $\neg\, \varphi(n)$. This proves the limited principle of omniscience.

Finally, we state, without proof, the following result by Palmgren.

Proposition 3.40. *The principle* CSAT *holds in* $\mathcal{N}$*.*

Proof. Follows from [49, Theorem 3.4], and [50, Theorem 3.1]. □

We also point out that $\mathcal{N}$ has an interesting subcategory of $\neg\neg$-sheaves (that is, sheaves with respect to the local operator $\neg\neg : \Omega \to \Omega$), with an internal logic that is classical, yielding a categorical model of Nelson's IST. This has been studied by Awodey and Eliasson in [3, 13].

With this, we conclude our present analysis of the filter topoi $\mathcal{N}$ and $\mathcal{U}$. Before moving on, we want to remark that, irrespective of any interest in nonstandard arithmetic, and with the *caveat* about $\mathsf{AC}^{\mathrm{st}}$ and the axiom of choice, $\mathcal{U}$ also provides a model for the logic of the "standard" Diller-Nahm translation, under the interpretation $[\![0]\!]_\wedge := \Delta_1\mathbb{N}$, $[\![0 \to 0]\!]_\wedge := \Delta_1(\mathbb{N} \to \mathbb{N})$, etc. In this case, we obtain a weaker transfer theorem for $\vee$-free formulae whose quantifiers are all bounded, i.e. they range over some finite sequence.

In the final section, we will briefly survey an entirely different class of categorical models that have been proposed for functional interpretations, including Diller-Nahm and nonstandard Dialectica; whose relation to the sheaf models, if any, is still obscure.

3.5 Herbrandised realisability topoi

In 1979, Hyland discovered the *effective topos* **Eff** [19] - the prime example of an interesting elementary topos which is *not* a Grothendieck topos, its global sections functor lacking a left adjoint. First order arithmetic in **Eff** matches Kleene's recursive realisability interpretation; and, in fact, it was the internal logic of **Eff** that suggested the correct higher order generalisation of the latter, including the uniformity principle UP that we mentioned in Chapter 1.

Hyland gave a direct construction; later, together with Johnstone and Pitts, he saw that it could be profitably generalised [20]. The resulting procedure - the *tripos-to-topos construction* -, and the new techniques developed around it, mostly by Pitts [51], have been instrumental in the definition of many other non-Grothendieck topoi, which we collectively call *realisability topoi.*

Since then, other, equivalent ways have been found to build realisability topoi, such as *exact completions* [35]; but we will only describe the first, which is more perspicuous in its logical underpinning. We will use [67] as a reference.

In the following definition, a *Heyting prealgebra* is a preorder whose poset reflection is a Heyting algebra; there is a preorder-enriched category **Heytpre** of Heyting prealgebras, and morphism that preserve all the relevant structure. This is a subcategory of **Preord**, the category of preorders and order-preserving maps.

A *pseudofunctor* $F : \mathbf{C} \to \mathbf{D}$ is a functor "up to isomorphism"; that is, for all objects C of $\mathbf{C}$, $F\mathrm{id}_C \simeq \mathrm{id}_{FC}$, and, for all pairs of composable morphisms $f : C \to C'$, $g : C' \to C''$ in $\mathbf{C}$, $F(gf) \simeq FgFf$. Similarly, a pseudo-natural transformation is defined in the same way as a natural transformation, with equalities replaced by isomorphisms.

The definition of a tripos can be given relative to any category with finite products; but since we will only actually use triposes over **Set**, we give a simplified definitions, valid over cartesian closed categories.

Definition 3.41. Let $\mathbf{C}$ be a cartesian closed category. A *tripos* over $\mathbf{C}$ is a pseudofunctor $\mathcal{P} : \mathbf{C}^{\mathrm{op}} \to \mathbf{Heytpre}$, satisfying the following conditions.

1. For all morphisms $f : C \to D$ in $\mathbf{C}$, $\mathcal{P}f : \mathcal{P}D \to \mathcal{P}C$ has both a left adjoint $\exists_f$ and a right adjoint $\forall_f$ in **Preord**, and these satisfy the *Beck-Chevalley condition*: for all pullback squares

$$\begin{tikzcd} E \arrow[r, "h"] \arrow[d, "k"] & D \arrow[d, "f"] \\ D' \arrow[r, "g"] & C \end{tikzcd}$$

   in $\mathbf{C}^{\mathrm{op}}$, it holds that $\forall_h \mathcal{P}k \simeq \mathcal{P}f \forall_g$.

2. There exist an object Σ of $\mathbf{C}$, and an element $\sigma \in \mathcal{P}\Sigma$, such that, for all objects C of $\mathbf{C}$, and all $\varphi \in \mathcal{P}C$, there is a morphism $[\varphi] : C \to \Sigma$ such that $\varphi \simeq \mathcal{P}[\varphi](\sigma)$ in $\mathcal{P}C$.

Such a σ is called a *generic predicate* of $\mathcal{P}$. Note that the Beck-Chevalley condition for the right adjoints implies a similar condition for the left adjoints.

A *transformation* of triposes is a pseudo-natural transformation $f : \mathcal{P} \to \mathcal{Q}$, where $\mathcal{P}$ and $\mathcal{Q}$ are treated as pseudofunctors from $\mathbf{C}^{\mathrm{op}}$ to **Preord** (that is, for all objects C of $\mathbf{C}$, fC is only required to be order-preserving).

Definition 3.42. A tripos $\mathcal{P}$ over **Set** is *canonically presented* if there exists a set Σ such that, for all sets S, $\mathcal{P}S$ has Σ^S as the underlying set, and, for all functions $f : S \to T$, $\mathcal{P}f$ is precomposition with f.

For a canonically presented tripos, one can take $\mathrm{id}_\Sigma \in \Sigma^\Sigma$ as the generic predicate. It can be proved that every tripos over **Set** is "equivalent" to a canonically presented one - meaning that they generate equivalent topoi through the tripos-to-topos construction [20, Proposition 1.9]; and, in fact, we will only define canonically presented triposes.

Suppose $\mathcal{E}$ is an elementary topos. Then we can define a tripos over $\mathcal{E}$ by taking, for all objects X of $\mathcal{E}$, $\mathcal{P}X := \mathrm{Sub}(X)$, or just the preorder of the monomorphisms into X; and, as the generic predicate, the subobject classifier $\mathrm{true} : 1 \rightarrowtail \Omega$.

The idea behind the tripos-to-topos construction is that this procedure can be reversed; that one can define, from a pair $(\mathbf{C}, \mathcal{P})$ of a category and a tripos over it, an elementary topos where the structure of individual subobject lattices, and their mutual relation through change of base functors, is induced by $\mathcal{P}$.

Much like we interpreted a first order language $\mathcal{L}$ in a Heyting category $\mathbf{H}$, it is possible to interpret it in a pair $(\mathbf{C}, \mathcal{P})$ - except for *equality*, which we want to add at a later stage; just proceed as in Section 3.1.1, replacing

- for all objects C of $\mathbf{C}$, the subobject lattices $\mathrm{Sub}(C)$ with the Heyting prealgebras $\mathcal{P}C$;
- for all morphisms $f : C \to D$ in $\mathbf{C}$, the change of base functor $f^* : \mathrm{Sub}(D) \to \mathrm{Sub}(C)$ and its left and right adjoints, with $\mathcal{P}f : \mathcal{P}D \to \mathcal{P}C$ and its left and right adjoints.

Given a sentence φ of $\mathcal{L}$, we write

$$\mathcal{P} \models \varphi$$

if the interpretation of φ in $(\mathbf{C}, \mathcal{P})$ is the top element of $\mathcal{P}1$. As in the case of Heyting categories, it can be proved that this interpretation is sound for intuitionistic first order logic without equality.

In particular, if we take as $\mathcal{L}$ a language whose types are the objects of $\mathbf{C}$, with their obvious interpretation in $(\mathbf{C}, \mathcal{P})$ as themselves, we can use it to speak about properties of the tripos itself, as we do in the following definition.

Definition 3.43 (Tripos-to-topos construction)**.** Let $(\mathbf{C}, \mathcal{P})$ be a pair of a cartesian closed category and a tripos over it. The category $\mathbf{C}[\mathcal{P}]$ is described by the following data.

- Objects are *partial equivalence $\mathcal{P}$-relations*; that is, pairs $(C, \sim)$, where C is an object of $\mathbf{C}$, and $\sim$ is an element of $\mathcal{P}(C \times C)$, such that

$$\begin{aligned} &\mathcal{P} \models \forall x, y : C\,(x \sim y \to y \sim x)\ , \\ &\mathcal{P} \models \forall x, y, z : C\,(x \sim y \wedge y \sim z \to x \sim z)\ . \end{aligned}$$

  $\sim$ is called an *equality predicate* for C. We write $E(x) := x \sim x$, and read it as "*x exists*".

- Morphisms $f : (C, \sim) \to (D, \sim)$ are isomorphism classes of *functional $\mathcal{P}$-relations*; that is, elements F of $\mathcal{P}(C \times D)$ such that

$$\begin{aligned} &\mathcal{P} \models \forall x : C\,\forall y : D\,(F(x,y) \to E(x) \wedge E(y))\ , \\ &\mathcal{P} \models \forall x, x' : C\,\forall y, y' : D\,\big(F(x,y) \wedge x \sim x' \wedge y \sim y' \to F(x', y')\big)\ , \\ &\mathcal{P} \models \forall x : C\,\forall y, y' : D\,(F(x,y) \wedge F(x,y') \to y \sim y')\ , \\ &\mathcal{P} \models \forall x : C\,\big(E(x) \to \exists y : D\,F(x,y)\big)\ , \end{aligned}$$

  expressing the fact that F is, respectively, *strict*, *relational*, *single-valued*, and *total.*

Theorem 3.44. *Let $(\mathbf{C}, \mathcal{P})$ be a pair of a cartesian closed category and a tripos over it. Then $\mathbf{C}[\mathcal{P}]$ is an elementary topos.*

Proof. See [67, Theorem 2.2.1]. □

So a topos is obtained from a tripos by adding a notion of equality, and imposing that morphisms respect it. The interpretation of logic in $(\mathbf{C}, \mathcal{P})$, with tripos semantics, and in $\mathbf{C}[\mathcal{P}]$, with Kripke-Joyal semantics, are still close enough that it is possible to reduce the latter to the former; for which see [67, Section 2.3].

Moreover, we can define certain operations on triposes, which then lift to operations on the generated topoi. Recall the following, fundamental definition.

Definition 3.45. Let $\mathcal{E}, \mathcal{E}'$ be elementary topoi. A *geometric morphism* $f : \mathcal{E} \to \mathcal{E}'$ consists of a pair of functors

$$f^* : \mathcal{E}' \rightleftarrows \mathcal{E} : f_* \ ,$$

such that $f^* \dashv f_*$, and f^* preserves finite limits. The left adjoint f^* is called *inverse image*, and the right adjoint f_* *direct image.*

We say that f is *surjective* if f^* is faithful, and that f is an *inclusion* if f_* is full and faithful.

Geometric morphisms are the adequate notion of functor between elementary topoi, when one is interested in the geometric aspects of the latter, i.e. topoi as a generalisation of sheaves on a topological space; there is a vast theory surrounding them, for which we can only refer to [24]. We now define a corresponding notion for triposes.

Definition 3.46. Let $\mathcal{P}, \mathcal{Q}$ be triposes over $\mathbf{C}$. A *geometric morphism* $f : \mathcal{P} \to \mathcal{Q}$ consists of a pair of transformations of triposes

$$f^* : \mathcal{Q} \rightleftarrows \mathcal{P} : f_* \ ,$$

such that, for all objects C of $\mathbf{C}$, $f^*C \dashv f_*C$ as maps between preorders, and f^*C preserves finite meets.

We say that f is *connected* if f^* is full and faithful, and that f is an *inclusion* if f_* is full and faithful.

Theorem 3.47. *Every geometric morphism $f : \mathcal{P} \to \mathcal{Q}$ of triposes over $\mathbf{C}$ can be lifted to a geometric morphism $\bar{f} : \mathbf{C}[\mathcal{P}] \to \mathbf{C}[\mathcal{Q}]$ of topoi.*

Moreover, if f is connected, then $\bar{f}$ is surjective, and if f is an inclusion of triposes, then $\bar{f}$ is an inclusion.

Proof. See [67, Theorem 2.5.8] for the lifting, [67, Theorem 2.5.11] for inclusions, and [6, Theorem I.1.16] for connected geometric morphisms. $\square$

In [6], tripos theory was applied to the construction of topoi whose internal logic would mirror the Dialectica interpretation, and its variants. We are interested, in particular, in the topos called $\mathbf{DN}_m$ there, which was introduced in [58], and which most closely resembles the Diller-Nahm interpretation. Its definition is based on **Mod**, the *modified realisability* topos [66], whose construction we also provide.

All the next definitions could be given relative to any *partial combinatory algebra* $\mathbb{A}$ [67, Chapter 1]; however, since we will not make use of this extra generality, we will always refer to the partial combinatory algebra $\mathbb{N}$ of (appropriately numbered) partial recursive functions, on which we assume that a coding of finite sequences has been fixed.

> **Notation.** Given $e, n \in \mathbb{N}$, we write $en\downarrow$ if the partial recursive function numbered e is defined on n, and, in that case, we write en for the result of the application.
>
> For all $n_0, \ldots, n_m \in \mathbb{N}$, we write $\langle n_0, \ldots, n_m \rangle$ for the code of the finite sequence s such that $|s| = m + 1$, and $s_k = n_k$, for $k = 0, \ldots, m$.

Given subsets $A, B \subseteq \mathbb{N}$, we introduce standard abbreviations:

$$\begin{aligned} A \otimes B &:= \{\langle a, b\rangle \mid a \in A \wedge b \in B\} , \\ A \oplus B &:= (\{0\} \otimes A) \cup (\{1\} \otimes B) , \\ A \Rightarrow B &:= \{e \in \mathbb{N} \mid \forall a \in A\,(ea\downarrow \wedge\, ea \in B)\} , \\ A^* &:= \{\langle a_0, \ldots, a_m\rangle \mid a_0, \ldots, a_m \in A\} . \end{aligned}$$

We also write $\mathrm{S}A$ for the set $\{n+1 \mid n \in A\}$.

Definition 3.48. The *modified realisability tripos* is a canonically presented tripos $\mathcal{M}od$ over **Set**, such that, for all sets S, $\mathcal{M}od\, S = ((\Sigma_{\mathcal{M}od})^S, \leq_S)$, where

$$\Sigma_{\mathcal{M}od} := \{A = (A_a, A_p) \in P\mathbb{N} \times P\mathbb{N} \mid A_a \subseteq A_p \wedge 0 \in A_p\} .$$

We call the elements of A_p *potential* realisers, and the elements of A_a *actual* realisers.

The logical connectives on $\mathcal{M}od\, S$ are defined pointwise: for all functions $A, B : S \to \Sigma_{\mathcal{M}od}$, and $x \in S$,

$$\begin{aligned} (A \wedge B)(x) &:= (A(x)_a \otimes B(x)_a, A(x)_p \otimes B(x)_p) , \\ (A \vee B)(x) &:= (A(x)_a \oplus B(x)_a, A(x)_p \oplus B(x)_p) , \\ (A \to B)(x) &:= \big((A(x)_a \Rightarrow B(x)_a) \cap (A(x)_p \Rightarrow B(x)_p), A(x)_p \Rightarrow B(x)_p\big) ; \end{aligned}$$

the preorder, on the other hand, is *uniform*:

$$A \leq_S B \quad \text{if and only if} \quad \exists n \in \bigcap_{x \in S} (A \to B)(x)_a .$$

Informally, $A \leq_S B$ if and only if there is a number that works as an actual realiser of $(A \to B)(x)$, uniformly for all $x \in S$.

The top and bottom elements of each $\mathcal{M}od\, S$ are the constant functions with value $(\mathbb{N}, \mathbb{N})$, and $(\emptyset, \mathbb{N})$, respectively.

For a function $f : S \to T$, the left and right adjoint of $\mathcal{M}od\, f$ can be defined, for $A \in \mathcal{M}od\, S$, and $y \in T$, as

$$\begin{aligned} \exists_f A(y) &:= \Big(\mathrm{S}\Big(\bigcup_{y=f(x)} A(x)_a\Big), \{0\} \cup \mathrm{S}\Big(\bigcup_{y=f(x)} A(x)_p\Big)\Big) , \\ \forall_f A(y) &:= \Big(\bigcap_{y=f(x)} (\{0\} \Rightarrow A(x)_a), \bigcap_{y=f(x)} (\{0\} \Rightarrow A(x)_p)\Big) . \end{aligned}$$

The *modified realisability topos* is then the topos $\mathbf{Mod} := \mathbf{Set}[\mathcal{M}od]$.

To understand the definition of $\mathcal{M}od$, one should think of the pairs (A_a, A_p) in $\Sigma_{\mathcal{M}od}$ as representing formulae $\exists n \in A_p\; A_a(n)$; then, implication is defined precisely as implication in modified realisability: $\exists n \in A_p\; A_a(n) \to \exists m \in B_p\; B_a(m)$ becomes

$$\exists e \in (A_p \Rightarrow B_p)\, \forall n \in A_p\; \big(en\downarrow \wedge\, (A_a(n) \to B_a(en))\big) .$$

For the Diller-Nahm interpretation, we want to represent formulae of the form

$$\exists n \in A_+\, \forall m \in A_-\; A_a(n, m) ,$$

and we want their implication to be Diller-Nahm implication; that alone suggests the following definition.

From now on, we avoid explicitly defining the left and right adjoints to $\mathcal{P}f$, for a tripos $\mathcal{P}$ and a function $f : S \to T$; it suffices to know that they exist.

Definition 3.49. The *modified Diller-Nahm tripos* is a canonically presented tripos $\mathcal{DN}_m$ over **Set**, such that, for all sets S, $\mathcal{DN}_m\, S = ((\Sigma_{\mathcal{DN}_m})^S, \leq_S)$, where

$$\Sigma_{\mathcal{DN}_m} := \{A = (A_a, A_+, A_-) \in P(\mathbb{N} \times \mathbb{N}) \times P\mathbb{N} \times P\mathbb{N} \,|\, A_a \subseteq (A_+ \times A_-) \wedge 0 \in A_+\}\ .$$

The logical connectives in $\mathcal{DN}_m\, S$ are defined as follows: for all functions $A, B : S \to \Sigma_{\mathcal{DN}_m}$,

$$\begin{aligned}
&(A \wedge B)(x)_+ := A(x)_+ \otimes B(x)_+\ , \qquad (A \wedge B)(x)_- := A(x)_- \oplus B(x)_-\ ,\\
&(A \wedge B)(x)_a := \{(\langle n, m\rangle, \langle z, k\rangle) \,|\, (z = 0 \wedge A(x)_a(n, k)) \vee (z = 1 \wedge B(x)_a(m, k))\}\ ;\\
&(A \vee B)(x)_+ := A(x)_+ \oplus B(x)_+\ , \qquad (A \vee B)(x)_- := A(x)_- \oplus B(x)_-\ ,\\
&(A \vee B)(x)_a := \{(\langle z, n\rangle, \langle z', m\rangle) \,|\, z = z' \to C_z(x)_a(n, m)\}\ , \qquad C_0 \equiv A,\ C_1 \equiv B\ ;\\
&(A \to B)(x)_+ := (A(x)_+ \Rightarrow B(x)_+) \otimes (A(x)_+ \otimes B(x)_- \Rightarrow A(x)_-^*)\ ,\\
&(A \to B)(x)_- := A(x)_+ \otimes B(x)_-\ ,\\
&(A \to B)(x)_a := \{(\langle e_+, e_-\rangle, \langle n, m\rangle) \,|\, \forall k \in e_-\langle n, m\rangle\ A(x)_a(n, k) \to B(x)_a(e_+ n, m)\}\ .
\end{aligned}$$

Again, the preorder requires uniform realisers of implication: $A \leq_S B$ if and only if there exists

$$n \in \bigcap_{x \in S} (A \to B)(x)_+$$

such that, for all $x \in S$, and for all $m \in (A \to B)(x)_-$, $(n, m) \in (A \to B)(x)_a$.

As top and bottom elements we can choose the constant functions with value $(\emptyset, \mathbb{N}, \emptyset)$, and $(\emptyset, \mathbb{N}, \mathbb{N})$, respectively.

The *modified Diller-Nahm topos* is then the topos $\mathbf{DN}_m := \mathbf{Set}[\mathcal{DN}_m]$.

Proposition 3.50. *There is a connected geometric morphism of triposes $q : \mathcal{DN}_m \to \mathcal{M}od$, hence a surjective geometric morphism of topoi $\bar{q} : \mathbf{DN}_m \to \mathbf{Mod}$.*

Proof. Define a pair of functions $q_* : \Sigma_{\mathcal{DN}_m} \to \Sigma_{\mathcal{M}od}$, and $q^* : \Sigma_{\mathcal{M}od} \to \Sigma_{\mathcal{DN}_m}$, as follows:

$$\begin{aligned}
q_* &: (A_a, A_+, A_-) \mapsto (\{n \in A_+ \,|\, \forall m \in A_-\ A_a(n, m)\}, A_+)\ ,\\
q^* &: (A_a, A_p) \mapsto (A_a \times \{0\}, A_p, \{0\})\ .
\end{aligned}$$

Let $q : \mathcal{DN}_m \to \mathcal{M}od$ be the geometric morphism of triposes induced by postcomposition of functions with q_* and q^*. It is easy to check that q is connected. □

In fact, all geometric morphisms between canonically presented triposes $\mathcal{P}$ and $\mathcal{Q}$ are induced by functions between the sets $\Sigma_{\mathcal{P}}$ and $\Sigma_{\mathcal{Q}}$ [6, Proposition I.1.17].

It is intuitive to think of the function q_* as mapping a formula $\exists n \in A_+\, \forall m \in A_-\ A_a(n, m)$ to itself; only, the universal quantifier over A_- is considered a part of the "actual realisers" predicate in $\mathcal{M}od$, instead of a structural element as in $\mathcal{DN}_m$. Similarly, q^* introduces a vacuous quantifier, transforming $\exists n \in A_p\ A_a(n)$ into

$$\exists n \in A_p\, \forall m \in \{0\}\ A_a(n)\ .$$

In [62] and [63], van den Berg introduced *herbrandised* versions of the modified realisability, and modified Diller-Nahm triposes, with the intent of providing a semantic counterpart to the *Herbrand realisability* interpretation [64] - a herbrandised variant of modified realisability -, and, respectively, to nonstandard Dialectica; which, as we saw, is basically herbrandised Diller-Nahm.

Definition 3.51. The *Herbrand tripos* is a canonically presented tripos $\mathcal{H}er$ over **Set**, such that, for all sets S, $\mathcal{H}er\,S = ((\Sigma_{\mathcal{H}er})^S, \leq_S)$, where

$$\Sigma_{\mathcal{H}er} := \{A = (A_a, A_p) \in P\mathbb{N} \times P\mathbb{N} \,|\, A_a \subseteq A_p^* \,, A_a \text{ upwards closed in } A_p^*\} \,.$$

There is an isomorphism $(A \oplus B)^* \simeq A^* \otimes B^*$. The connectives are thus defined as follows: for all functions $A, B : S \to \Sigma_{\mathcal{H}er}$,

$$\begin{aligned}
(A \wedge B)(x) &:= (A(x)_a \otimes B(x)_a, A(x)_p \oplus B(x)_p) \,, \\
(A \vee B)(x) &:= (\{\langle n, m\rangle \,|\, n \in A(x)_a \vee m \in B(x)_a\}, A(x)_p \oplus B(x)_p) \,, \\
(A \to B)(x)_p &:= A(x)_p^* \Rightarrow B(x)_p^* \,, \\
(A \to B)(x)_a &:= \{s \in (A(x)_p^* \Rightarrow B(x)_p^*)^* \,|\, \exists n \in s \left(n \in (A(x)_a \Rightarrow B(x)_a)\right)\} \,.
\end{aligned}$$

Notice that disjunction is treated in the same way as conjunction, compatibly with the loss of its constructive meaning due to herbrandisation. The preorder is defined as in $\mathcal{M}od$:

$$A \leq_S B \quad \text{if and only if} \quad \exists n \in \bigcap_{x \in S} (A \to B)(x)_a \,.$$

The top and bottom element can be taken to be the constant functions with value $(\mathbb{N}^*, \mathbb{N})$, and $(\emptyset, \mathbb{N})$, respectively.

The *Herbrand topos* is then the topos $\mathbf{Her} := \mathbf{Set}[\mathcal{H}er]$.

The nonstandard Dialectica tripos has the same relation to the Herbrand tripos, as $\mathcal{DN}_m$ has to $\mathcal{M}od$; and the same relation to $\mathcal{DN}_m$, as $\mathcal{H}er$ has to $\mathcal{M}od$.

Definition 3.52. The *nonstandard Dialectica tripos* is a canonically presented tripos $\mathcal{D}st$ over **Set**, such that, for all sets S, $\mathcal{D}st\,S = ((\Sigma_{\mathcal{D}st})^S, \leq_S)$, where

$$\begin{aligned}
\Sigma_{\mathcal{D}st} := \{A = (A_a, A_+, A_-) \in P(\mathbb{N} \times \mathbb{N}) \times P\mathbb{N} \times P\mathbb{N} \,|\, A_a \subseteq (A_+^* \times A_-) \,, \\
A_a \text{ upwards closed in } A_+^*\} \,.
\end{aligned}$$

The logical connectives in $\mathcal{D}st\,S$ are defined as follows: for all functions $A, B : S \to \Sigma_{\mathcal{D}st}$,

$$\begin{aligned}
&(A \wedge B)(x)_+ := A(x)_+ \oplus B(x)_+ \,, \qquad (A \wedge B)(x)_- := A(x)_- \oplus B(x)_- \,, \\
&(A \wedge B)(x)_a := \{(\langle n, m\rangle, \langle z, k\rangle) \,|\, (z = 0 \wedge A(x)_a(n, k)) \vee (z = 1 \wedge B(x)_a(m, k))\} \,; \\
&(A \vee B)(x)_+ := A(x)_+ \oplus B(x)_+ \,, \qquad (A \vee B)(x)_- := A(x)_- \otimes B(x)_- \,, \\
&(A \vee B)(x)_a := \{(\langle n, m\rangle, \langle j, k\rangle) \,|\, A(x)_a(n, j) \vee B(x)_a(m, k)\} \,; \\
&(A \to B)(x)_+ := (A(x)_+^* \Rightarrow B(x)_+^*) \oplus (A(x)_+^* \otimes B(x)_- \Rightarrow A(x)_-^*) \,, \\
&(A \to B)(x)_- := A(x)_+^* \otimes B(x)_- \,, \\
&(A \to B)(x)_a := \{(\langle e_+, e_-\rangle, \langle n, m\rangle) \,|\, \forall k \in e_-[\langle n, m\rangle]\; A(x)_a(n, k) \to B(x)_a(e_+[n], m)\} \,;
\end{aligned}$$

where, for $e \in (A \to B^*)^*$, and $n \in A$, we define $e[n] := e_0 n \cdot \ldots \cdot e_{|e|-1} n$, in analogy with the monotone sequence application we defined in Chapter 1.

The preorder is defined much like in $\mathcal{DN}_m$: $A \leq_S B$ if and only if there exists

$$n \in \bigcap_{x \in S} (A \to B)(x)^*_+$$

such that, for all $x \in S$, and for all $m \in (A \to B)(x)_-$, $(n, m) \in (A \to B)(x)_a$. Finally, we can choose the constant functions with value $(\emptyset, \emptyset, \emptyset)$, and $(\emptyset, \emptyset, \mathbb{N})$, as the top and the bottom element, respectively.

The *nonstandard Dialectica topos* is then the topos $\mathbf{Dst} := \mathbf{Set}[\mathcal{D}st]$.

Again, a good way to think of the elements of $\Sigma_{\mathcal{H}er}$ is as formulae

$$\exists s : A^*_p \; A_a(s) \; ,$$

and of the elements of $\Sigma_{\mathcal{D}st}$ as formulae

$$\exists s : A^*_+ \; \forall m : A_- \; A_a(s, m) \; ,$$

with A_a upwards closed in s. We will use these analogies in the following proof, which answers an open problem from [63].

Proposition 3.53. *There is a pullback square*

$$\begin{array}{ccc} \mathcal{D}st & \xrightarrow{d} & \mathcal{DN}_m \\ \big\downarrow{\scriptstyle p} & & \big\downarrow{\scriptstyle q} \\ \mathcal{H}er & \xrightarrow{u} & \mathcal{M}od \end{array}$$

of geometric morphisms of triposes, such that d and u are inclusions, and p and q are connected geometric morphisms.

Proof. Since all the triposes are canonically presented, such a square will arise from a diagram

$$\begin{array}{ccc} \Sigma_{\mathcal{D}st} & \underset{d^*}{\overset{d_*}{\rightleftarrows}} & \Sigma_{\mathcal{DN}_m} \\ {\scriptstyle p_*}\big\downarrow\big\uparrow{\scriptstyle p^*} & & {\scriptstyle q_*}\big\downarrow\big\uparrow{\scriptstyle q^*} \\ \Sigma_{\mathcal{H}er} & \underset{u^*}{\overset{u_*}{\rightleftarrows}} & \Sigma_{\mathcal{M}od} \end{array}$$

of functions, such that the direct images form a pullback diagram in **Set**.

We have already constructed the connected geometric morphism $d : \mathcal{DN}_m \to \mathcal{M}od$. For the others, instead of giving the direct definition in terms of elements of the Σ_- sets, we will describe their action on the associated formulae.

For $u : \mathcal{H}er \to \mathcal{M}od$, the direct image maps $\exists s \in A^*_p \; A_a(s)$ to itself; the inverse image maps $\exists n \in A_p \; A_a(n)$ to

$$\exists s \in A^*_p \; \exists n \in s \; A_a(n) \; .$$

Similarly, for $d : \mathcal{D}st \to \mathcal{DN}_m$: the direct image maps $\exists s \in A^*_+ \; \forall m \in A_- \; A_a(s, m)$ to itself, while the inverse image maps $\exists n \in A_+ \; \forall m \in A_- \; A_a(n, m)$ to

$$\exists s \in A^*_+ \; \forall m \in A_- \; \exists n \in s \; A_a(n, m) \; .$$

Finally, for $p : \mathcal{D}st \to \mathcal{H}er$, the direct image maps $\exists s \in A_+^* \; \forall m \in A_- \; A_a(s,m)$ to itself, while the inverse image maps $\exists s \in A_p^* \; A_a(s)$ to

$$\exists s \in A_p^* \; \forall m \in \{0\} \; A_a(s) \; .$$

It is a simple verification that these are well defined, that the direct images form a pullback diagram, and that the geometric morphisms induced by d and u are inclusions, while the one induced by p is connected. □

It follows that there is also a pullback square

$$\begin{tikzcd} \mathbf{Dst} \arrow[r, tail, "d"] \arrow[d, "p"] & \mathbf{DN}_m \arrow[d, "q"] \\ \mathbf{Her} \arrow[r, tail, "u"] & \mathbf{Mod} \end{tikzcd}$$

of topoi and geometric morphisms, where p and q are surjective, and d and u are geometric inclusions. Recently, a universal characterisation of the inclusion $\mathbf{Her} \rightarrowtail \mathbf{Mod}$ has been found by Johnstone [25], the Herbrand topos being the *Gleason cover* [23] of the modified realisability topos; yet no such characterisation is known for the inclusion $\mathbf{Dst} \rightarrowtail \mathbf{DN}_m$, and, to date, the properties of both topoi remain quite obscure.

Nevertheless, our analysis of the Grothendieck topoi $\mathcal{N}$ and $\mathcal{U}$ raises a number of suspicions. There, too, we had a geometric inclusion $\mathcal{N} \rightarrowtail \mathcal{U}$, with inverse image the sheafification functor.

Moreover, as discussed in [62], there are embeddings $\nabla : \mathbf{Set} \rightarrowtail \mathbf{Mod}$, and $\nabla' : \mathbf{Set} \rightarrowtail \mathbf{Her}$, both right adjoint to the global sections functor of their target topos, with the property that ∇ does not preserve any coproducts, while ∇' preserves first order logic (and, in particular, coproducts of a finite number of copies of 1), but not the natural numbers object. Hence, $\nabla'\mathbb{N}$ is a nonstandard model of arithmetic in $\mathbf{Her}$.

Since p^* and q^* are full and faithful, composing with them, we obtain embeddings $q^*\nabla : \mathbf{Set} \rightarrowtail \mathbf{DN}_m$, and $p^*\nabla' : \mathbf{Set} \rightarrowtail \mathbf{Dst}$; and since they are inverse images of geometric morphisms, hence preserve finite limits and all colimits, what we said about ∇, ∇' is still true of $q^*\nabla$, $p^*\nabla'$. In particular, $p^*\nabla'\mathbb{N}$ is a nonstandard model of arithmetic in $\mathbf{Dst}$.

In the case of the filter topoi, by first embedding $\mathbf{Set}$ into $\mathfrak{F}\mathbf{Set}$, with the identification of a set S with the simple filter $(S, \{S\})$, and then by applying the appropriate Yoneda embedding, we obtained functors $\mathbf{y} : \mathbf{Set} \rightarrowtail \mathcal{U}$ and $\mathbf{y}' : \mathbf{Set} \rightarrowtail \mathcal{N}$. And there, too, $\mathbf{y}'$ preserved first order logic, but not the natural numbers object.

So naturally, we wonder: is there an analogue of Proposition 3.22 for the inclusion $\mathbf{Dst} \rightarrowtail \mathbf{DN}_m$? Explicitly, is $\mathbf{Dst}$ the topos $\mathrm{sh}_j(\mathbf{DN}_m)$, where j is the smallest local operator on $\mathbf{DN}_m$ such that the monomorphism $m : 2 \rightarrowtail q^*\nabla 2$ is j-dense?

Furthermore, applying the filter construction to $\mathbf{Set}$, and then taking K_1-sheaves, respectively K-sheaves, leads to topoi whose internal logic mirrors the (uniform) Diller-Nahm, respectively the nonstandard Dialectica interpretation. In this case, we have geometric morphisms from $\mathbf{DN}_m$ and from $\mathbf{Dst}$ to $\mathbf{Mod}$. Could it be that these topoi arise from a similar construction, performed relative to the modified realisability topos, instead of $\mathbf{Set}$?

We had no time, or were unable to answer these questions; with them, we close this chapter.

Conclusions

If we are to sum up, in brief, what we believe are the main *conceptual* achievements of this thesis - those that an interested reader should not overlook - it may just come down to the following.

1. *The significance of $\mathcal{U}$.* The filter topos $\mathcal{U}$ doubles as a model of the logic of the Diller-Nahm interpretation, and as a cue to its extension with uniform quantifiers - the uniform Diller-Nahm interpretation. This was previously unknown, and might lead to an improved understanding of the underlying, *geometric* structure of Diller-Nahm logic.

2. *A better view on herbrandisation.* The comparison of $\mathcal{N}$ with $\mathcal{U}$ provided a categorical counterpart to herbrandisation, and allowed for a refined analysis of its effects. This includes the re-contextualisation of NCR as a herbrandised uniformity principle; on the contrary, the role of finite sequences in $\mathsf{HGMP}^{\mathrm{st}}$ appeared to be a byproduct of their role in US^*, rather than the result of herbrandisation.

Singling these two out, of course, is the expression of a personal penchant: if you are more into nonstandard arithmetic, for instance, the definition and examination of sequence overspill and underspill may be more interesting.

Of equal, if not superior, importance are the new questions, that new results raise. We want to conclude this thesis with a review of these, together with a few, more speculative directions for potential future research.

Chapter 1. The main question left unsolved concerns the constructive acceptability of the principle $\mathsf{TR}_{\forall\exists}$, which would also imply the conservativity of HAC over E-HA$^{\omega*}$.

Moreover, we would like to know how independent the principles OS^* and US^* are. We know that the Herbrand realisability interpretation vacuously accepts the former, yet does *not* have a realiser for the latter [64]; but it would be desirable to actually construct a model of intuitionistic nonstandard arithmetic with overspill, and no underspill.

In addition, we did not address the question from [64] about the interpretation of the general saturation principle SAT in a constructive setting; this, too, deserves to be clarified.

Chapter 2. We defined a new functional interpretation, but ignore, so far, how useful it is for applications. Its similarity to light Dialectica is encouraging; on the other hand, the use of functional interpretations has been most successful in program extraction from *classical* proofs, and we have not investigated yet how well uniform Diller-Nahm composes with negative translations, such as Kuroda's [30].

In light of the results of Chapter 3, Palmgren's work on the topos $\mathcal{N}$ indicates that the characteristic principles of nonstandard Dialectica lead to a useful calculus for nonstandard analysis. We conjectured that the characteristic principles of uniform Diller-Nahm may be a good axiomatisation of Lifschitz's calculability arithmetic [32]; is this correct, and could this also be a useful calculus by itself?

On a more speculative note, Oliva provided in [42] a unified view of the Dialectica, Diller-Nahm, and modified realisability interpretations, through *linear logic.* Is there an equivalent of herbrandisation in linear logic - connected, perhaps, to the bang (!) modaliser - such that nonstandard Dialectica and Herbrand realisability, too, would be amenable to such a treatment?

Chapter 3. We already put our main open questions at the end of the chapter: they concern the relation between the inclusion $\mathcal{N} \rightarrowtail \mathcal{U}$ of the filter topoi, and the inclusion $\mathbf{Dst} \rightarrowtail \mathbf{DN}_m$ of the matching realisability topoi.

On a different subject, we pointed out that, for $\mathsf{HAC}^{\mathrm{st}}$ and $\mathsf{AC}^{\mathrm{st}}$ to hold in $\mathcal{N}$ and in $\mathcal{U}$, respectively, we need some choice principles in the metatheory; which seems to indicate that an intensional metatheory is preferable. However, the nonstandard Dialectica and uniform Diller-Nahm interpretations were formulated in *extensional* Heyting arithmetic; so there is seemingly a clash, which we would like to figure out.

We hope that these, and related questions can be answered in future work, by us or by you.

Acknowledgments

First of all, I am grateful to my supervisor, Benno van den Berg; for I feel I have learnt a lot, during the last few months, and that wouldn't have happened, if it weren't for his insightful indications and comments. Our weekly meetings have always been a pleasure, and I honestly hope that this collaboration may be the first of many.

Then, I am grateful to Ludovico Pernazza, who kindly agreed to supervise a thesis on a subject that is not his primary, and read my drafts with much more care than it would have sufficed.

On a more personal level, thanks to you all, *Peltlaaners* - Bill, Ket, Kelly, Filipe, Inu, Tommaso, and Makis - for being flatmates with a Premium Seal of Quality; I'll miss our dinners, and all the rest. I still have to go canoeing with you on the Oudegracht, so wait - you'll hear from me again!

Then, my friends and mates from Collegio Ghislieri. How does one thank *anybody* for five, inestimable years? Whether we did well - *si fecimus bonum* - I do not know; but it's been a great ride, all along. With every single one of you I'd have something to share, yet that would require the size of another thesis (and nobody would give me a degree for that, either); so I'll keep it to the ones who've been somewhat closer to me. Thank you Alarico, for conversations ranging from feelings to transhumanism; thank you Proxy, for pushing me into some great adventurous trips, and Amen, for spectacularly arranging their logistics. Thank you Marghe, for just being great in some way I can't define, and Kent, for all the laughs and the junk food. Thank you Pirans - you're the best actor I've directed -, thank you Tarro for the Stupendous Europe Foundation; thank you Paola, for coffee and... **wink**, and Crema, who's by far the best news agency on the market. Thank you Sara, for many shared interests, classes, and notes; thank you Sacco for being a dean that's like a rose, and Salvo - well, I had to thank you right afterwards - for being what you are. Thank you Ruiz, I *really* missed the latest two barbecues; thank you Oxy, for being a true believer in gender parity, and Alice, for nighttime teas that sadly became too sporadic. Thank you Bracco, for being truly out of the ordinary, and thank you, Manetta and Appy, for never really leaving us. I wish well to all of you, and to the others I did not mention explicitly.

Then, the friendships that lasted, even if I've been elsewhere for five years. Thank you Tommaso, thank you Francesco, for always being there; we are different, we have different lives, yet I feel there is some foundation to our friendship that never fails. I trust it'll keep being that way, even as we're scattered farther and farther across Europe. Thanks also to Laura, Luca, and other friends from Milan; whom I haven't heard regularly, but when it happened, it was always nice.

Thanks to my parents - in fact, thanks to everyone in my family - for their everlasting support, and the independence that I learned from them.

Thanks, Alessandra - so many names I can call you - for all the happiness.

Bibliography

[1] J. Avigad and S. Feferman. Chapter V: Gödel's functional ("Dialectica") interpretation. *Studies in Logic and the Foundations of Mathematics*, 137:337–405, 1998.

[2] J. Avigad and J. Helzner. Transfer principles in nonstandard intuitionistic arithmetic. *Archive for Mathematical Logic*, 41(6):581–602, 2002.

[3] S. Awodey and J. Eliasson. Ultrasheaves and double negation. *Notre Dame Journal of Formal Logic*, 45(4):235–245, 2004.

[4] U. Berger. Uniform Heyting arithmetic. *Annals of Pure and Applied Logic*, 133(1):125–148, 2005.

[5] A. R. Bernstein and A. Robinson. Solution of an invariant subspace problem of K. T. Smith and P. R. Halmos. *Pacific J. Math*, 16:421–431, 1966.

[6] B. Biering. *Dialectica Interpretations: A Categorical Analysis.* PhD thesis, IT University of Copenhagen, 2008.

[7] E. Bishop. Review of Elementary Calculus by H.J. Keisler. *Bulletin of the American Mathematical Society*, 83(2):205–208, 1977.

[8] E. Bishop and D. S. Bridges. *Constructive analysis*, volume 278. Springer Berlin, 1985.

[9] A. Blass. Two closed categories of filters. *Fundamenta Mathematicae*, 94(2):129–143, 1977.

[10] C. Butz. Saturated models of intuitionistic theories. *Annals of Pure and Applied Logic*, 129(1):245–275, 2004.

[11] R. Diaconescu. Axiom of choice and complementation. *Proceedings of the American Mathematical Society*, 51(1):176–178, 1975.

[12] J. Diller. Eine Variante zur Dialectica-Interpretation der Heyting-Arithmetik endlicher Typen. *Archive for Mathematical Logic*, 16(1):49–66, 1974.

[13] J. Eliasson. *Ultrasheaves.* PhD thesis, Uppsala University, 2003.

[14] S. Feferman. Gödel's Dialectica interpretation and its two-way stretch. In *Computational logic and proof theory*, pages 23–40. Springer, 1993.

[15] K. Gödel. Über eine bisher noch nicht benützte Erweiterung des finiten Standpunktes. *Dialectica*, 12(3-4):280–287, 1958.

[16] K. Gödel. Remark on nonstandard analysis. In *Collected works, volume II: Publications 1938–1974*. Oxford University Press, New York, NY, 1990.

[17] M.-D. Hernest. Light functional interpretation. In *Computer science logic*, pages 477–492. Springer, 2005.

[18] W. A. Howard. Hereditarily majorizable functionals of finite type. *Metamathematical investigation of intuitionistic arithmetic and analysis*, 344:454–461, 1973.

[19] J. M. E. Hyland. The effective topos. *Studies in Logic and the Foundations of Mathematics*, 110:165–216, 1982.

[20] J. M. E. Hyland, P. T. Johnstone, and A. M. Pitts. Tripos theory. In *Math. Proc. Camb. Phil. Soc*, volume 88, pages 205–232. Cambridge Univ Press, 1980.

[21] H. Ishihara. Reverse mathematics in Bishop's constructive mathematics. *Philosophia Scientiæ. Travaux d'histoire et de philosophie des sciences*, 6(CS):43–59, 2006.

[22] T. Jech. *Set theory.* Springer Monographs in Mathematics. Springer-Verlag, Berlin, 2003. The third millennium edition, revised and expanded.

[23] P. T. Johnstone. The Gleason cover of a topos, I. *Journal of Pure and Applied Algebra*, 19:171–192, 1980.

[24] P. T. Johnstone. *Sketches of an elephant: a topos theory compendium*, volume 1. Oxford University Press, 2002.

[25] P. T. Johnstone. The Gleason cover of a realizability topos. Unpublished manuscript, 2013.

[26] U. Kohlenbach. *Applied proof theory: proof interpretations and their use in mathematics.* Springer, 2008.

[27] U. Kohlenbach and P. Oliva. *Proof mining: a systematic way of analysing proofs in mathematics.* BRICS, Computer Science Department, University of Aarhus, 2002.

[28] G. Kreisel. On the interpretation of non-finitist proofs - Part I. *The Journal of Symbolic Logic*, 16(4):241–267, 1951.

[29] G. Kreisel. Axiomatizations of nonstandard analysis that are conservative extensions of formal systems for classical standard analysis. *Applications of model theory to algebra, analysis and probability*, pages 93–106, 1969.

[30] S. Kuroda. Intuitionistische Untersuchungen der formalistischen Logik. *Nagoya Mathematical Journal*, 2:35–47, 1951.

[31] F. W. Lawvere. An elementary theory of the category of sets (long version) with commentary. *Repr. Theory Appl. Categ*, 11:1–35, 2005.

[32] V. Lifschitz. Calculable natural numbers. *Studies in Logic and the Foundations of Mathematics*, 113:173–190, 1985.

[33] P. A. Loeb. Conversion from nonstandard to standard measure spaces and applications in probability theory. *Transactions of the American Mathematical society*, 211:113–122, 1975.

[34] S. Mac Lane and I. Moerdijk. *Sheaves in geometry and logic: A first introduction to topos theory.* Springer Verlag, 1992.

[35] M. Menni. *Exact completions and toposes.* PhD thesis, University of Edinburgh, 2000.

[36] I. Moerdijk. A model for intuitionistic non-standard arithmetic. *Annals of Pure and Applied Logic*, 73(1):37–51, 1995.

[37] I. Moerdijk and E. Palmgren. Minimal models of Heyting arithmetic. *The Journal of Symbolic Logic*, 62(4):1448–1460, 1997.

[38] E. Nelson. Internal set theory: a new approach to nonstandard analysis. *Bulletin of the American Mathematical Society*, 83(6):1165–1198, 1977.

[39] E. Nelson. *Radically Elementary Probability Theory.* Number 117 in Annals of mathematics studies. Princeton University Press, 1987.

[40] E. Nelson. The syntax of nonstandard analysis. *Annals of pure and applied logic*, 38(2):123–134, 1988.

[41] B. Nordström, K. Petersson, and J. M. Smith. *Programming in Martin-Löf's type theory*, volume 200. Oxford University Press Oxford, 1990.

[42] P. Oliva. Unifying functional interpretations. *Notre Dame Journal of Formal Logic*, 47(2):263–290, 2006.

[43] P. Oliva. An analysis of Gödel's Dialectica interpretation via linear logic. *Dialectica*, 62(2):269–290, 2008.

[44] E. Palmgren. Constructive sheaf semantics. *Mathematical Logic Quarterly*, 43(3):321–327, 1997.

[45] E. Palmgren. A sheaf-theoretic foundation for nonstandard analysis. *Annals of Pure and Applied Logic*, 85(1):69–86, 1997.

[46] E. Palmgren. Developments in constructive nonstandard analysis. *Bulletin of Symbolic Logic*, pages 233–272, 1998.

[47] E. Palmgren. Constructive nonstandard representations of generalized functions. *Indagationes Mathematicae*, 11(1):129–138, 2000.

[48] E. Palmgren. Overspill and countable saturation implies LLPO. Note available on author's personal web page, 2000.

[49] E. Palmgren. Real numbers in the topos of sheaves over the category of filters. *Journal of Pure and Applied Algebra*, 160(2):275–284, 2001.

[50] E. Palmgren. Unifying constructive and nonstandard analysis. *Reuniting the Antipodes-Constructive and Nonstandard Views of the Continuum*, 306:167, 2002.

[51] A. M. Pitts. *The theory of triposes.* PhD thesis, University of Cambridge, 1982.

[52] A. M. Robert. *Nonstandard analysis.* Dover Publications Inc., Mineola, NY, 1988.

[53] A. Robinson. *Non-standard analysis.* Princeton Landmarks in Mathematics. Princeton University Press, Princeton, NJ, 1996. Reprint of the second (1974) edition, with a foreword by W. A. J. Luxemburg.

[54] H. Rogers Jr. *Theory of recursive functions and effective computability.* MIT press, 1987.

[55] S. Sanders. Reuniting the antipodes: bringing together nonstandard analysis and constructive analysis. In *SAS Opening Workshop "The Mathematical Legacy of Alan Turing" (Spitalfields Day)*, 2013.

[56] C. Schmieden and D. Laugwitz. Eine erweiterung der infinitesimalrechnung. *Mathematische Zeitschrift*, 69(1):1–39, 1958.

[57] T. Skolem. Über die Nicht-charakterisierbarkeit der Zahlenreihe mittels endlich oder abzählbar unendlich vieler Aussagen mit ausschliesslich Zahlenvariablen. *Fundamenta mathematicae*, 23(1):150–161, 1934.

[58] T. Streicher. A semantic version of the Diller-Nahm variant of Gödels Dialectica interpretation. Note available on author's personal web page, 2006.

[59] A. S. Troelstra. *Metamathematical investigation of intuitionistic arithmetic and analysis*, volume 344. Springer, 1973.

[60] A. S. Troelstra. Chapter VI: Realizability. *Studies in Logic and the Foundations of Mathematics*, 137:407–473, 1998.

[61] B. van den Berg. Predicative toposes. *arXiv preprint arXiv:1207.0959*, 2012.

[62] B. van den Berg. The Herbrand topos. *Math. Proc. Camb. Phil. Soc*, 155(2):361–374, 2013.

[63] B. van den Berg. A topos for a nonstandard functional interpretation. *arXiv preprint arXiv:1301.3679*, 2013.

[64] B. van den Berg, E. Briseid, and P. Safarik. A functional interpretation for nonstandard arithmetic. *Annals of Pure and Applied Logic*, 2012.

[65] J. van Oosten. *Basic category theory.* Computer Science Department, 1995.

[66] J. van Oosten. The modified realizability topos. *Journal of pure and applied algebra*, 116(1):273–289, 1997.

[67] J. van Oosten. *Realizability: an introduction to its categorical side*, volume 152 of *Studies in Logic and the Foundations of Mathematics.* Elsevier B. V., Amsterdam, 2008.

[68] F. Wattenberg. Nonstandard analysis and constructivism? *Studia Logica*, 47(3):303–309, 1988.

List of axioms and rules

Index