Chapter 10
Banishing Ultrafilters from Our Consciousness

Domenico Cantone, Eugenio G. Omodeo and Alberto Policriti

> *The reader who remembers these key points will do well in what follows. In particular, it is now quite all right to entirely forget how the nonstandard universe was defined and to banish ultrafilters from our consciousness.*
>
> (Martin Davis, *Applied Nonstandard Analysis,* 1977)

Abstract The way in which Martin Davis conceived the first chapter of his book "*Applied nonstandard analysis*" is a brilliant example of information hiding as a guiding principle for the design of widely applicable constructions and methods of proof. We discuss here a common trait that we see between that book and another writing of the year 1977, "*Metamathematical extensibility for theorem provers and proof-checkers*", which Martin coauthored with Jacob T. Schwartz. To tie the said part of Martin's study on nonstandard analysis to proof technology, we undertake a verification, by means of a proof-checker based on set theory, of key results of the non-standard approach to analysis.

Keywords Proof checking · Proof engineering · Nonstandard analysis · Foundations of infinitesimal calculus

D. Cantone (✉)
DMI, Università di Catania, Viale A. Doria 6, 95125 Catania, Italy
e-mail: cantone@dmi.unict.it

E.G. Omodeo
DMG/DMI, Università di Trieste, Via Valerio 12/1, 34127 Trieste, Italy
e-mail: eomodeo@units.it

A. Policriti
DMIF, Università di Udine, Via Delle Scienze 206, 33100 Udine, Italy
e-mail: alberto.policriti@uniud.it

E.G. Omodeo and A. Policriti (eds.), *Martin Davis on Computability, Computational Logic, and Mathematical Foundations*,
Outstanding Contributions to Logic 10, DOI 10.1007/978-3-319-41842-1_10

10.1 Introduction

Year 1977: Martin Davis appears in print with "*Applied nonstandard analysis*" [14], whose subject is less close to computability and computational logic than the various areas to which Martin has contributed before. Nevertheless we will tie that book—appropriately, we believe—to another publication of the same year, "*Metamathematical extensibility for theorem provers and proof-checkers*" [23, pp. 120–146], jointly authored by Martin and his friend and colleague "Jack" (namely Jacob T. Schwartz).[1] We aim at unveiling an affinity between some of the matter which that book treats in preparation for analysis proper and the field of automated reasoning of which Martin has been a trailblazer since its early days,[2] and at taking advantage of that link for a proof-checking undertaking which we see as promising.

Martin's book is dedicated to the memory of Abraham Robinson, the creator of nonstandard analysis. At the Summer Institute for Symbolic Logic held at Cornell University, a scientific gathering that both had attended in 1957, Robinson gave a talk in which he "made the provocative remark that the auxiliary points, lines, or circles 'constructed' as part of the solution to a geometry problem can be thought of as being elements of what is now called the Herbrand universe for the problem" [15, pp. 7–8].[3] At the same meeting Martin reported on his own implementation, three years earlier on a JOHNNIAC machine, of Presburger's decision procedure for elementary additive number theory [12]. This proximity of interests between the two distinguished scholars about automating proofs was, presumably, coincidental.

In 1977, on the other hand, disappointment is beginning to take place in the automated deduction community (see [5]), as researchers experience the combinatorial explosion plaguing the automatic search for mathematical proofs even if pruned by the best available techniques. More emphasis is now placed on comfortable interaction between man and computerized proof assistants, and on proof checkers (see, e.g., [38]) as opposed to fully automatic theorem provers. Specific knowledge pertaining to diverse branches of mathematics begins to be perceived as essential for an advancement of the proof techniques; Ballantyne and Bledsoe [3] (see also [2]) succeed in automating the proofs of hard theorems in analysis using methods which rely on the nonstandard viewpoint.

The new context brings to the fore issues related to correct-program technology and proof engineering. An emblem of the times is the Clear specification

[1] See [22] and, therein, the enjoyable [16]; see also [21] and [1, pp. 478–480]. The above-cited [23] led to the sole joint publication by Martin and Jack, namely [24].

[2] Landmark contributions of Martin to automatic theorem-proving in 1st-order predicate logic have been [10, 13, 19, 20, 25], historically occurring between Paul C. Gilmore's and Dag Prawitz' methods, on the one hand, and J. Alan Robinson's resolution principle on the other. Concerning the *linked conjunct* method then proposed by Martin and his team at Bell Labs, see [29, 39].

[3] The term '*Herbrand universe*', today widely used, appeared for the first time in the influential paper [13] (reviewed in [34]); but [17, p. 432] contends that it would be more historically correct to credit the construction of that universe to Thoralf Skolem.

language [7], paving the way to the OBJ family of languages, which will integrate specification, prototyping, and verification into a system with a single underlying logic: theorem-proving is now aimed at providing mechanical assistance for proofs that are needed in the development of software and hardware. This is the scene encountered by the joint work [24] of Martin and Jack, at the dawn of large-scale proof technology.

Here is the issue they raised. "For use of mechanized proof verifier systems to remain comfortable over a wide range of applications, ··· it should be possible to augment the system by adding new symbols, schemes of notation, and extended rules of inference of various kinds" [24, p. 217]. A stringent requirement is that the envisioned changes to a system do not disrupt its soundness: a proof verifier should, therefore, be furnished with the metamathematical capability of justifying its progressive augmentations.

Use of a metamathematical extension mechanism, [24] points out, leads to the common acceptance of algebraic calculations in lieu of detailed predicate calculus proofs. Although recourse to the methods of nonstandard analysis in lieu of the ε-δ methods is not mentioned in that paper, we see that less familiar but expedient detour as being in accord with the matter under discussion.

As an arena for experimenting with this circle of ideas, we have undertaken a merciless formal remake of [14, Chap. 1] with Jack's proof checker Ref, see [35, Chap. 4], which embodies a variant of the Zermelo-Fraenkel set theory. This task, which has hardly anything to do with analysis *per se*, is an essential prerequisite if we are to bring the methods of nonstandard analysis within the scope of Ref. As a result of the "mathematical simplicity, elegance, and beauty of these methods"—and of "enthusiasm ··· not unrelated to the well-known pleasures of the illicit"—, we expect to eventually get the reward of "their far-reaching applications" (see [14, p. viii]).

Our effort will also suggest changes to Ref's current implementation which can improve its metamathematical extensibility.

We have set up substantial ground for specifying and proving, by means of the Ref verifier (very succinctly described in Sect. 10.6), the two consequences of Łoś's theorem which we need (namely, Theorems 10.1 and 10.2 in Sect. 10.3): once we will have fully achieved those goals, we will move on to work on Robinson's concurrence theorem and on a few other crucial propositions (Theorems 10.3, 10.4, 10.5, and 10.6 of Sect. 10.4). To complete our job we must then introduce "schemes of notation and extended rules of inference of various kinds" that properly assist Ref's users in exploiting nonstandard methods.

In order to reach the goals of our experiment, we must express in set-theoretic terms metalevel notions such as the evaluation of a sentence in a universe; another not entirely trivial task concerns the representation of individuals (thought of as 'non-sets') within a formal system which deals with sets whose construction ultimately relies on nothing but the null set $\emptyset$. For these two matters, to be discussed in Sect. 10.9 and in Sects. 10.7 and 10.8 respectively, our experiment is innovative, at least as regards the Ref proof checker. In other respects, we can benefit from work previously

done: among other things we found, already adequately formalized, a theory of ordinal numbers conceived *à la* Raphael M. Robinson, and the ultrafilter theorem obtained using Zorn's lemma.

Concerning proof checkers, issues of reuse have an even greater relevance than for theorem provers. Such issues pertain more to proof engineering than to computational logic:[4] rather than going through the same proof pattern several times, one should abstract a common method to be recalled over and over again, with all the conveniences offered by technology.

Reuse is supported in Ref by a construct named 'THEORY' (see [31] and [35, pp. 19–25]), similar to—although of a less algebraic nature—a mechanism for parameterized specifications of the aforementioned Clear specification language. This paper will discuss how to organize THEORYs that enable one to tackle without reiteration of techniques the foundations of nonstandard analysis; hopefully, it will stimulate reflections on good "proof hiding" practices, of the kind which Martin's passage [14, p. 42], quoted in the epigraph to this paper, seems eager to suggest.

10.2 Basic Construction for Nonstandard Analysis

> **Why nonstandard analysis?** Nonstandard analysis is a technique rather than a subject ⋯ The subject can be claimed to be of importance insofar as it leads to simpler, more accessible expositions, or (more important) to mathematical discoveries. [14, p. 1]

The initial part of [14] dwells on how to enlarge a standard universe into a nonstandard one. While taking stock at the end of the first chapter, Martin stresses that much of the machinery developed up to there is not used in the remainder of the book; then, in recapitulating which key points the reader should remember, he underlines the three main tools of nonstandard analysis: *transfer principle*, *concurrence*, and *internality*.

We will now give a quick account of the elaborate ultrapower construction whose details Martin deems "quite all right", after that turning point, "to banish from our consciousness". We thereby undertake a formal recasting of that construction with Ref, in order to encapsulate it within Ref's THEORYs.

The STANDARD UNIVERSE is the SUPERSTRUCTURE

$$\widehat{\mathsf{S}} = \overbrace{\underbrace{\mathsf{S} \cup \mathscr{P}(\mathsf{S})}_{\mathsf{S}_1} \cup \mathscr{P}(\mathsf{S} \cup \mathscr{P}(\mathsf{S}))}^{\mathsf{S}_2} \cup \mathscr{P}(\mathsf{S} \cup \mathscr{P}(\mathsf{S}) \cup \mathscr{P}(\mathsf{S} \cup \mathscr{P}(\mathsf{S}))) \cup \cdots}^{\mathsf{S}_3}$$

[4] See [8, pp. 5–6]. In a recent personal web-page, David Aspinall (Univ. of Edinburgh) defines *Proof Engineering* to mean the activity on construction, maintenance, documentation and presentation of large formal proof developments. Within Proof Engineering, according to Aspinall, "Software Engineering provides the techniques to develop large, structured and well-specified repositories of computer code; proof checking provides the mechanisms to provide a complete semantics and formal correctness as an absolute quality criterion.".

built on a set $\mathsf{s} = \mathsf{s}_0$, whose $(n+1)$-st stage is $\mathsf{s}_{n+1} = \mathsf{s}_n \cup \mathscr{P}(\mathsf{s}_n)$ for each $n \in \mathbb{N} = \{0, 1, 2, \dots\}$ (as customary, $\mathscr{P}$ designates the powerset operator). It is essential that s consists of *individuals*; namely that $\emptyset \notin \mathsf{s}$ and that no element of any element of s pops up at any stage, viz., $\widehat{\mathsf{s}} \cap \bigcup \mathsf{s} = \emptyset$. Every set w of individuals generates a superstructure $\widehat{w}$, much as we have just indicated for s; e.g., $\widehat{\emptyset}$ consists of the entities known as *hereditarily finite sets*.

The superstructure $\widehat{\mathsf{s}}$ gets embedded into another one, $\widehat{\mathsf{w}}$, built on a specific set $\mathsf{w} \supset \mathsf{s}$ of individuals, by means of a function $x \mapsto {}^*x$; in particular ${}^*\mathsf{s} = \mathsf{w}$. A set $\widetilde{\mathsf{w}} \subset \widehat{\mathsf{w}}$ is cut out of the wider superstructure: this $\widetilde{\mathsf{w}}$, satisfying the revealing equality $\widetilde{\mathsf{w}} = \bigcup_{i \in \mathbb{N}} {}^*\mathsf{s}_i$, will be the NONSTANDARD UNIVERSE paired with $\widehat{\mathsf{s}}$.

Such companions $\widehat{\mathsf{s}}$, $\widetilde{\mathsf{w}}$ will—in a sense—have the same properties. An unrestrained formulation of this principle would have paradoxical consequences, though, and we must postpone to Sect. 10.3 the precise formulation of criteria enabling the transferability of properties. In a major instance studied in [14, Chap. 2], s includes an Archimedean ordered field D, e.g., the field $\mathbb{Q}$ of rational numbers or the field $\mathbb{R}$ of real numbers; then ${}^*\mathsf{D}$, included in w, will still satisfy the laws of an ordered field but will violate the Archimedean property which—roughly speaking—rejects infinitely large or infinitely small elements.[5]

Before showing how to construct $\widetilde{\mathsf{w}}$, let us make it clear which are the sets which qualify as universes:[6]

Definition 10.1 A set $\mathscr{U}$ is called a UNIVERSE if $\emptyset \in \mathscr{U}$ and the following properties hold for all x, y:

Upward closure: If $x, y \in \mathscr{U}$, then $\{x, y\} \in \mathscr{U}$.
Downward closure: If $x \in \mathscr{U}$ and $x \cap \mathscr{U} \neq \emptyset$, then $x \subseteq \mathscr{U}$
(this says that each element x of $\mathscr{U}$ is either an individual, hence has no element in $\mathscr{U}$, or is included in $\mathscr{U}$). ⊣

The upward closure property readily yields that a universe $\mathscr{U}$ is always closed with respect to the Kuratowski ordered pair formation $\langle x, y \rangle =_{\text{Def}} \{\{x\}, \{x, y\}\}$; by also exploiting downward closure we get, for each function $g \in \mathscr{U}$ such that $g \cup \mathsf{dom}(g) \subseteq \mathscr{U}$, that the result $g\lceil x$ of applying g to a set x belongs to $\mathscr{U}$. (By function we mean here a single-valued set of ordered pairs; moreover, when g fails to be a function or x does not belong to its domain, $g\lceil x$ is meant to designate $\emptyset$.) Every superstructure based on a set of individuals is a universe, so it is closed with respect to pair formation and to function application.

[5] In particular, when $\mathsf{D} = \mathbb{R}$, we get a field, ${}^*\mathbb{R}$, of entities called *hyperreal* numbers. In ${}^*\mathbb{R}$ there are positive numbers lying infinitely close to zero; the reciprocals of such infinitesimals must, of course, exceed any positive integer.

[6] Our definition of universe marginally differs from the one given in [14, p. 15] in that we are not assuming individuals to be given beforehand. Certain proper classes can also be regarded as universes, according to a plain generalization of this definition to be seen in Fig. 10.5.

The construction of $\widetilde{\mathsf{w}}$ relies on a pair $\mathfrak{æ}$, $\mathfrak{i}$ such that

(1) $\mathfrak{æ} \subseteq \mathscr{P}(\mathfrak{i}) \setminus \{\emptyset\}$;
(2) $x \cap y \in \mathfrak{æ}$ for all $x, y \in \mathfrak{æ}$;
(3) $y \in \mathfrak{æ}$ whenever $x \in \mathfrak{æ}$ and $x \subseteq y \subseteq \mathfrak{i}$;
(4) no strict superset of $\mathfrak{æ}$ meets the *filter* conditions (1)–(3).

(Consequently, see [14, p. 10], $\{x \cap \mathfrak{i}\,,\ \mathfrak{i} \setminus x\} \cap \mathfrak{æ} \neq \emptyset$ holds for every set x.) By well-established terminology, $\mathfrak{æ}$ is an *ultrafilter*[7] over the *index set* $\mathfrak{i} = \bigcup \mathfrak{æ}$. Momentarily we do not commit our choice of $\mathfrak{æ}$ and $\mathfrak{i}$ in any way; this choice is most relevant, though, for the applicability of the nonstandard techniques.

We say that a property $C(j)$ of elements of $\mathfrak{i}$ holds *a.e.* ('almost everywhere') if $\{j \in \mathfrak{i} \mid C(j)\} \in \mathfrak{æ}$, that is, if the indices satisfying C form a set which belongs to $\mathfrak{æ}$. Thus, for example, the condition $gj = hj$ *a.e.* defines an equivalence relation over $\mathsf{s}^{\mathfrak{i}}$, the set of all functions from the index set into standard individuals; we can then pick a representative element ρg out of each equivalence class $\left\{h \in \mathsf{s}^{\mathfrak{i}} \mid gj = hj \text{ a.e.}\right\}$, and finally get the set

$$\mathsf{w} = \left\{ \rho\, g : \ g \in \mathsf{s}^{\mathfrak{i}} \right\}$$

of *nonstandard individuals*. This is an enlargement of s, whose elements can in fact be put in natural correspondence with the representatives of *a.e.* constant functions (the injection of s into w is $x \mapsto \rho\, g_x$, where $g_x \in \{x\}^{\mathfrak{i}}$, i.e. $g_x = \mathfrak{i} \times \{x\}$). We will manage to enforce the strict inclusion $\mathsf{w} \supsetneq \mathsf{s}$ in Sect. 10.4; our present assumptions only suffice to ensure that $\mathsf{w} \supseteq \mathsf{s}$.

The construction at issue continues with the specification of a function, $\bar{\ }$, whose set of values will be the universe $\widetilde{\mathsf{w}}$ we are after and whose domain is layered in a way mimicking the hierarchical organization

$$\widehat{\mathsf{s}} = \bigcup_{n\in\mathbb{N}} \mathsf{s}_n = \mathsf{s}_0 \uplus \biguplus_{n\in\mathbb{N}} \left(\mathscr{P}(\mathsf{s}_n) \setminus \mathsf{s}_n\right)$$

(where $\uplus$ and $\biguplus$ designate disjoint unions) of the standard universe:

$$\bar{\ }: \bigcup_{n\in\mathbb{N}} \left\{ f \in \widehat{\mathsf{s}}^{\mathfrak{i}} \mid fj \in \mathsf{s}_n \text{ a.e.} \right\} \longrightarrow \widehat{\mathsf{w}}\,.$$

For each $f \in \widehat{\mathsf{s}}^{\mathfrak{i}}$ such that $fj \in \mathsf{s}_0$ *a.e.*, we put $\bar{f} = \rho g$, where $g \in \mathsf{s}^{\mathfrak{i}}$ is such that $fj = gj$ *a.e.* Next, for successive numbers $n \in \mathbb{N}$, we define *à la* Mostowski the image $\bar{f}$ of each function f such that $fj \in \mathscr{P}(\mathsf{s}_n) \setminus \mathsf{s}_n$ *a.e.*, by putting

$$\bar{f} = \left\{ \bar{g} : \ g \in \widehat{\mathsf{s}}^{\mathfrak{i}} \mid gj \ \in \ \mathsf{s}_n \cap (fj) \ \ a.e. \right\}\,.$$

[7] A slicker characterization of ultrafilters will be shown in Fig. 10.7.

The following facts admit straightforward proofs:

- $\widetilde{\mathsf{W}}$, the set of all images $\bar{f}$, is a universe;
- $\bar{f} \in \bar{g}$ if and only if $fj \in gj$ *a.e.*;
- $\bar{f} = \bar{g}$ if and only if $fj = gj$ *a.e.*

Much as before, there is a natural one-one correspondence between $\widehat{\mathsf{S}}$ and those functions, in the domain of $\bar{}$, which are *a.e.* constant; hence the embedding * of $\widehat{\mathsf{S}}$ into $\widetilde{\mathsf{W}}$ announced at the beginning of this section is plainly induced by $\bar{}$. This function * will soon be extended by bringing into its domain many subsets of $\widehat{\mathsf{S}}$ which do not belong to $\widehat{\mathsf{S}}$.

Before going any further, let us pause to recall that Martin works under the assumption that "we have available some given sufficiently large set $\mathscr{I}$ of true individuals (sometimes called *urelemente*), about which we assume nothing except that they are not sets" [14, p. 11], and he repeatedly stresses that questions as to the true 'nature' of such entities are irrelevant to mathematical practice.[8] Anyway, we will have to face this issue (see Sect. 10.8) while carrying out our formalization task, because our framework will be a set theory devoid of individuals proper: our 'individuals' will simply be sets whose elements are 'inaccessible' from within the superstructure.

10.3 Bounded Formulae and the Transfer Principle

> The link between logic and computing is to a great extent the notion of a formal language, which is the kind of language machines understand. [18, p. 83]
>
> Formulas of $\mathscr{L}_U$ can be used not only to make assertions about U, but also to define subsets of U. [14, p. 23]

In order to make assertions about a universe $\mathscr{U}$ and to introduce its definable subsets, [14, pp. 20–21] specifies a language $\mathscr{L}_{\mathscr{U}}$ endowed with:

T0. constants $\boldsymbol{c}$, which are in one-one correspondence with the elements of $\mathscr{U}$ (each $\boldsymbol{c}$ is meant to designate the corresponding element c of $\mathscr{U}$);
T1. a countable infinitude $\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \ldots$ of variables (each ranging over $\mathscr{U}$);
T2. dyadic function symbols $\langle s, t \rangle$ and $(s \restriction t)$ (which are meant to designate, respectively, ordered pair formation and function application);

F0. dyadic relation symbols $(s = t)$ and $(s \in t)$ (designating $=$ and $\in$);
F1. propositional connectives $\neg$ (monadic) and $\&$ (dyadic);
F2. bounded quantifiers of the form $(\exists \boldsymbol{x}_n \in t)$, where t stands for a term where $\boldsymbol{x}_n$ does not appear.

---

[8] In a similar attitude, [11, p. 54] states that "one possible view is that the integers are atoms and should not be viewed as sets. Even in this case, one might still wish to prevent the existence of unrestricted atoms. In any case, for the 'genuine' sets, Extensionality holds and the other sets are merely harmless curiosities.".

More detailed syntactic rules about terms and formulae of $\mathscr{L}_{\mathscr{U}}$, as well as the semantics of $\mathscr{L}_{\mathscr{U}}$, follow the pattern familiar to anyone who has encountered first-order predicate languages; we leave them as understood for the time being and will belabor this point when arriving at our formalization task (see Sect. 10.9). Anyway, it will best suit our purposes to handle only formulae in *negative normal form*: hence we admit as primitive constructs also the propositional connective $\vee$ and bounded universal quantifiers $(\forall \boldsymbol{x}_n \in t)$; moreover, we confine $\neg$ inside contexts of the forms $\neg(s = t)$ and $\neg(s \in t)$, shortened as usual to $(s \neq t)$ and $(s \notin t)$.

If exactly one variable, say $\boldsymbol{x}_n$, occurs free in a formula α of $\mathscr{L}_{\mathscr{U}}$, then we indicate by $\alpha(\boldsymbol{c})$ the sentence[9] resulting from α when all free occurrences of $\boldsymbol{x}_n$ get replaced by a constant, $\boldsymbol{c}$, that designates some $c \in \mathscr{U}$.

Definition 10.2 A set $d \subseteq \mathscr{U}$ is called DEFINABLE if there is a formula α of $\mathscr{L}_{\mathscr{U}}$ with one free variable such that $d = \{c \in \mathscr{U} \mid \alpha(\boldsymbol{c}) \text{ is true in } \mathscr{U}\}$. ⊣

Consider, now, the languages $\mathscr{L}_{\widehat{\mathsf{S}}}$ and $\mathscr{L}_{\widetilde{\mathsf{W}}}$ of the standard universe and of its nonstandard counterpart. Let the notation $\models \alpha$ express the fact that α, a sentence of $\mathscr{L}_{\widehat{\mathsf{S}}}$, is true in $\widehat{\mathsf{S}}$; similarly, indicate by ${}^*\models\beta$ the fact that β, a sentence of $\mathscr{L}_{\widetilde{\mathsf{W}}}$, is true in $\widetilde{\mathsf{W}}$.

A translation $\lambda \mapsto {}^*\lambda$ of terms and formulae from $\mathscr{L}_{\widehat{\mathsf{S}}}$ into $\mathscr{L}_{\widetilde{\mathsf{W}}}$ can be specified as follows: to get ${}^*\lambda$, replace every constant $\boldsymbol{c}$ occurring in λ by the constant ${}^*\boldsymbol{c}$ that designates the image ${}^*c$ of c.

We can now state two propositions, both easily obtainable from Łoś's theorem, a fundamental result of model theory which we underplay here:

Theorem 10.1 *If α, β are formulae of $\mathscr{L}_{\widehat{\mathsf{S}}}$ where the only free variable is $\boldsymbol{x}_1$ and*

$$\{c \in \widehat{\mathsf{S}} \mid \models \alpha(\boldsymbol{c})\} = \{c \in \widehat{\mathsf{S}} \mid \models \beta(\boldsymbol{c})\}$$

holds, then

$$\{c \in \widetilde{\mathsf{W}} \mid {}^*\models {}^*\alpha(\boldsymbol{c})\} = \{c \in \widetilde{\mathsf{W}} \mid {}^*\models {}^*\beta(\boldsymbol{c})\} \ .$$

Theorem 10.2 (Transfer principle) *For every sentence α of $\mathscr{L}_{\widehat{\mathsf{S}}}$,*

$${}^*\models {}^*\alpha \text{ if and only if } \models \alpha \ .$$

Thanks to Theorem 10.1, we can add to the domain of the function *every definable subset d of $\widehat{\mathsf{S}}$, via the unambiguous stipulation that

$${}^*d = \{c \in \widetilde{\mathsf{W}} \mid {}^*\models {}^*\alpha(\boldsymbol{c})\} \text{ when } d = \{c \in \widehat{\mathsf{S}} \mid \models \alpha(\boldsymbol{c})\} \, .$$

[9] When the need will arise, we will adjust this notation also to terms, indicating by $t(\boldsymbol{c})$ a term devoid of variables resulting from replacement of a variable of t by a constant $\boldsymbol{c}$.

Davis thus briefly conveys the significance of the transfer principle:

> There is a formal language that can be used to make assertions that are ambiguous in that they can refer to either of the two structures. · · · The *transfer principle* roughly states that the same assertions of the formal language are true in the standard universe as in the nonstandard universe. It is typically used by proving a desired result in the nonstandard universe, and then, noting that the result is expressible in the language, concluding that it holds in the standard universe as well. [14, pp. 2–3]

Let us pause again, to observe that the task of formalizing within set theory such model-theoretic propositions as the above Theorems 10.1 and 10.2 presupposes that we encode terms and formulae via sets: we will display a technique for that purpose in Sect. 10.9. Similar tasks arise frequently in logic, when it comes to investigate inside a formal system some meta-theoretical issues regarding the system itself. E.g., in preparation for the proof that an axiomatic theory of sets is essentially undecidable one will encode its formulae, inside $\widehat{\emptyset}$ (see [33]) or even by means of natural numbers. Our encoding cannot be carried out with the same parsimony of means, due to the tight interplay between syntax and intended semantics in our languages (see the formation rule T0 of each $\mathscr{L}_{\mathscr{U}}$); we will manage, nonetheless, to encode the formulae of $\mathscr{L}_{\widehat{\mathsf{s}}}$ inside $\widehat{\mathsf{s}}$ and the ones of $\mathscr{L}_{\widehat{\mathsf{w}}}$ inside $\widehat{\mathsf{w}}$.

10.4 A Kind of 'All-at-Once Compactification'

> Another technique is *concurrence*. This is a logical technique that guarantees that the extended structure contains all possible completions, compactifications and so forth. [14, p. 3]

Suppose that s is infinite. If $\ddot{\imath}$ is also infinite and an injection g of $\ddot{\imath}$ into s exists, it will suffice to require that no finite set belongs to the ultrafilter $æ$ in order that $gj \neq x$ *a.e.* for any $x \in \mathsf{s}$; thus g must differ from any function h from $\ddot{\imath}$ to s which is *a.e.* constant, and *nonstandard individuals exist*! This is one way of making the nonstandard enlargement non-trivial (see [28, p. 52]).

Preliminary to the construction of a much richer nonstandard universe, [14, p. 34] defines concurrence. In our own, slightly readjusted terms:

Definition 10.3 Relative to a universe $\mathscr{U}$, a dyadic relation r such that $r \in \mathscr{U}$ and $r \cup \mathsf{dom}(r) \subseteq \mathscr{U}$ is said to be CONCURRENT if to every *finite* $d \subseteq \mathsf{dom}(r)$ there corresponds some $b \in \mathscr{U}$ s.t. $d \times \{b\} \subseteq r$. ⊣

Now let $\ddot{\imath}$ be the set of functions ϕ such that $\mathsf{dom}(\phi)$ is the set of all concurrent relations $r \in \widehat{\mathsf{s}}$ and $\phi\, r$ is a finite subset of $\mathsf{dom}(r)$ for each such r. The ultrafilter $æ$ will then be chosen so that $\ddot{\imath} = \bigcup æ$ holds and the membership relation

$$\{\phi \in \ddot{\imath} \mid \psi\, r \subseteq \phi\, r \text{ for each concurrent } r \in \widehat{\mathsf{s}}\,\} \in æ$$

also holds, for each $\psi \in \ddot{\imath}$. Here comes a key theorem, due to Abraham Robinson:

Theorem 10.3 (Concurrence theorem) *To every concurrent relation* $r \in \widehat{\mathsf{s}}$ *there corresponds some* $\ell \in \widetilde{\mathsf{w}}$ *such that* $\{{}^*a : a \in \mathsf{dom}(r)\} \times \{\ell\} \subseteq {}^*r$.

From this claim, [14, p. 36] draws the conclusion that nonstandard individuals exist: for, assuming $\mathbb{N} \subseteq \mathsf{s}$ in order to slightly simplify the argument, one such is the 'limit' element ℓ corresponding to the concurrent relation

$$\{\langle n, m\rangle : n \in \mathbb{N}, m \in \mathbb{N} \mid n < m\} ;$$

in fact, $\ell \in {}^*\mathbb{N} \setminus \mathsf{s}$.

> The third technique is *internality*. A set s of elements of the nonstandard universe is *internal* if s itself is an element of the nonstandard universe; otherwise, s is *external*. A surprisingly useful method of proof is one by reductio ad absurdum in which the contradiction is that some set one knows to be *external* would in fact be *internal* under the assumption being refuted. [14, p. 3]

Definition 10.4 We call

EXTERNAL SET: every element of $\widehat{\mathsf{w}} \setminus \widetilde{\mathsf{w}}$;
INTERNAL SET: every element of $\widetilde{\mathsf{w}} \setminus \mathsf{w}$. ⊣

After showing, with the aid of the transfer principle, that ${}^*\mathbb{N} \setminus \mathbb{N}$ is an external set, [14, pp. 39–41] provides criteria for demonstrating the internality of specific sets:

Theorem 10.4 (Internality theorem) *If* $d \subseteq \widetilde{\mathsf{w}}$ *is definable in* $\widetilde{\mathsf{w}}$ *and* a *is an internal set, then* $a \cap d$ *is an internal set.*

Theorem 10.5 *If* a *and* b *are internal sets, then so is* $a \times b$.

Theorem 10.6 (Internal function theorem) *If* $f \in b^a$, *where* a *and* b *are internal sets, and for a suitable term* t *of* $\mathscr{L}_{\widetilde{\mathsf{w}}}$ *involving one free variable*

$$f\,c \;\; \textit{is the value of} \;\; t(\boldsymbol{c}) \;\; \textit{in}\; \widetilde{\mathsf{w}}\,, \;\; \textit{for each} \;\; c \in a\,,$$

then f *is internal.*

Along the way, [14, pp. 39–41] shows $\mathbb{N}$ to be an external set.

10.5 Key Application of the Nonstandard Methods

In [14, Chap. 2] the construction of the nonstandard universe is used twice: first to obtain $\mathbb{R}$, the field of real numbers, from the field $\mathbb{Q}$ of the rationals; on second application, to work out the structure of ${}^*\mathbb{R}$ from $\mathbb{R}$. The first use can supersede such classical constructions as the ones devised by George Cantor and Richard Dedekind.

The second use brings infinitesimals into play, along with their inverses, which are infinite numbers: one is thus led into the realm of HYPERREAL numbers.

To briefly see how these embeddings work, consider first an *ordered field* D (in the customary sense). For any such field, we can assume w.l.o.g. that $\mathbb{Q} \subseteq D$.

Definition 10.5 Put

$$F = \bigcup_{n\in\mathbb{N}} \{\, x \in D \mid 0 \le x \le n \ \vee\ 0 < -x \le n \,\},$$
$$I = \{\, x \in D \mid x = 0 \ \vee\ (1/x) \in D \setminus F \,\}.$$

An element x of D is said to be FINITE, INFINITE, or INFINITESIMAL, depending as whether $x \in F$, $x \in D \setminus F$, or $x \in I$. For x, y in D, we say that x IS NEAR y if $x - y \in I$; if so, we write $x \approx y$.

D is called ARCHIMEDEAN if $F = D$; otherwise stated, if $I = \{0\}$. ⊣

As is plain, I is an ideal in the subring F of D; moreover, $\approx$ is an equivalence relation on D, whose restriction to F equals the equivalence relation induced by I. Consequently, the quotient $F/\approx = F/I$ is a ring; actually, it is an Archimedean ordered field.

Suppose next that D is an *Archimedean* ordered field and that $\mathsf{D} \subseteq \mathsf{s}$, where s is as in Sects. 10.2, 10.3, and 10.4. By virtue of the transfer principle, the ${}^*\mathsf{D}$ resulting from D through the ultrapower construction is, in its turn, an ordered field (of which D is a subfield). It is no longer Archimedean, though; for, its nonnull subset ${}^*\mathbb{N} \setminus \mathsf{s}$ consists of elements which are infinite. If we now designate by F and I the set of all finite, respectively infinitesimal, elements of ${}^*\mathsf{D}$, then it readily turns out that the canonical homomorphism $^\circ$ of F onto F/I acts as a monomorphism of D into F/I. After so embedding D in the Archimedean field F/I, [14, p. 51] goes on to prove that F, D, I, and ${}^*\mathsf{D} \setminus \mathsf{F}$ are all external subsets of ${}^*\mathsf{D}$; then, by resorting to the concurrence theorem, [14] obtains the following:

Theorem 10.7 (Dedekind's Theorem) *If A, B are nonnull subsets of D such that $a < b$ holds for all $a \in A$ and $b \in B$, then there is a $c \in \mathsf{F}/\mathsf{I}$ such that $a \le c \le b$ holds for all $a \in A$ and $b \in B$.*

From this, [14] gets that

Theorem 10.8 *F/I is a complete ordered field,*

after noting that between two elements x, y of an Archimedean ordered field such that $x < y$ there always lies a $q \in \mathbb{Q}$ such that $x < q < y$. Archimedean ordered fields exist (one such is, of course, $\mathbb{Q}$); therefore, a complete ordered field exists as well. Up to isomorphism, this must be *unique* (owing, in particular, to the fact that any complete ordered field is Archimedean): by definition, $\mathbb{R}$ is taken to be this field.

If we go over the same construction again, now taking $\mathsf{D} = \mathbb{R} \subseteq \mathsf{s}$, we can naturally identify F/I with $\mathbb{R}$ and, accordingly, think of $^\circ$ as being the field homomorphism that sends each finite hyperreal number to its *standard part*, namely to the sole real number which lies near it. It can also be shown (see [14, pp. 53, 56]) that infinitesimally near each real number there is a $q \in {}^*\mathbb{Q}$.

Typical notions of elementary real analysis can be captured in new terms from the nonstandard viewpoint, after which classical theorems can be obtained by nonstandard methods. Various illustrations of this are provided in [14, pp. 56–74], e.g.:

Theorem 10.9 *Consider a sequence* $\{ s_n : n \in \mathbb{N} \setminus \{0\} \}$ *of real numbers* s_n *and a real number* ℓ. *Then*

- *the sequence converges to* ℓ *if and only if* $({}^*s)_n \approx \ell$ *holds for* all *infinite* $n \in {}^*\mathbb{N}$;
- $({}^*s)_n \approx \ell$ *holds for* some *infinite* $n \in {}^*\mathbb{N}$ *if and only if, for each* $\varepsilon > 0$ *in* $\mathbb{R}$, *the inequality* $|s_n - \ell| < \varepsilon$ *is satisfied for infinitely many* $n \in \mathbb{N}$.

Theorem 10.10 *Let* f *be a real-valued function on the closed interval* $[a, b] =_{Def} \{ x \in \mathbb{R} \mid a \leq x \leq b \}$, *where* $a, b \in \mathbb{R}$ *and* $a < b$. *Then* f *is continuous at* $x_0 \in [a, b]$ *if and only if, for all* $x \in {}^*[a, b]$, $\quad x \approx x_0$ *implies* ${}^*f(x) \approx {}^*f(x_0)$.

Theorem 10.11 *Let* f *be a continuous real-valued function on the closed interval* $[a, b]$. *If* $f(a) < 0 < f(b)$, *then* $f(c) = 0$ *holds for some* $c \in [a, b]$.

Proof 1 (Sketch) Consider the function $t : \mathbb{N} \times \mathbb{N} \longrightarrow \mathbb{R}$ defined as follows:

$$t(n, i) = \begin{cases} a + i\,(b-a)/n & \text{if } n \in \mathbb{N} \setminus \{0\} \text{ and } 0 \leq i \leq n\,, \\ 0 & \text{otherwise,} \end{cases}$$

so that ${}^*t : {}^*\mathbb{N} \times {}^*\mathbb{N} \longrightarrow {}^*\mathbb{R}$ meets an analogous condition, by the transfer principle.

Choose $\nu \in {}^*\mathbb{N} \setminus \mathbb{N}$. Since $L = \{ i \in {}^*\mathbb{N} \mid f({}^*t(\nu, i)) > 0 \text{ and } i \leq \nu \}$ is a definable subset of ${}^*\mathsf{s}$, L is also internal by Theorem 10.4; and since $\nu \in L$, there is a least element $j > 0$ in L. If we take c to be the standard part of ${}^*t(\nu, j)$, it turns out that $c \approx {}^*t(\nu, j) \approx {}^*t(\nu, j-1)$; therefore $f(c) \approx f({}^*t(\nu, j)) \approx f({}^*t(\nu, j-1))$, and hence $f(c) = {}^\circ(f({}^*t(\nu, j))) = {}^\circ(f({}^*t(\nu, j-1)))$, where the inequalities ${}^\circ(f({}^*t(\nu, j))) \geq 0$ and ${}^\circ(f({}^*t(\nu, j-1))) \leq 0$ hold. We conclude that $f(c) = 0$, as desired. □

10.6 Basic Features of Our Proof Checker

Our proof-checker Ref, a.k.a. ÆtnaNova or Referee, processes script files, named SCENARIOS, which consist of definitions, theorems, and detailed proofs of the theorems. After checking a scenario for syntactic validity, Ref verifies that the proofs are compliant with the version of set theory built into it. The language in which scenarios are written extends the usual language of first-order predicate logic with constructs reflecting the theory which underlies Ref: we can for example, as shown by most of the abbreviating definitions in Fig. 10.1,[10] exploit a very flexible set abstraction construct of the form

[10] About Ref's built-in operator ***arb*** (X) that occurs thrice in Fig. 10.1, suffice it to say for the time being that it selects an element of its operand X when $X \neq \emptyset$, and that ***arb*** $(\emptyset) = \emptyset$.

DEF $\bigcup$: [El'ts of el'ts]	$\bigcup \mathsf{S}$	$=_{\text{Def}}$	$\{\mathsf{u} : \mathsf{v} \in \mathsf{S}, \mathsf{u} \in \mathsf{v}\}$
DEF $\mathscr{P}$: [All subsets]	$\mathscr{P}(\mathsf{S})$	$=_{\text{Def}}$	$\{\mathsf{x} : \mathsf{x} \subseteq \mathsf{S}\}$
DEF pair_0: [Ord'd pair]	$\langle \mathsf{X}, \mathsf{Y} \rangle$	$=_{\text{Def}}$	$\{\{\mathsf{X}\}, \{\mathsf{X}, \mathsf{Y}\}\}$
DEF pair_1: [Left proj.]	$\mathsf{Q}^{[1]}$	$=_{\text{Def}}$	$\boldsymbol{arb}(\{\mathsf{x} : \mathsf{s} \in \mathsf{Q}, \mathsf{x} \in \mathsf{s} \mid \mathsf{s} = \{\mathsf{x}\}\})$
DEF pair_2: [Right proj.]	$\mathsf{Q}^{[2]}$	$=_{\text{Def}}$	$\boldsymbol{arb}(\{\mathsf{y} : \mathsf{d} \in \mathsf{Q}, \mathsf{y} \in \mathsf{d} \mid \mathsf{Q} = \{\{\mathsf{y}\}\} \vee \mathsf{d} \setminus \{\mathsf{y}\} \in \mathsf{Q}\})$
DEF map_1: [Domain]	$\mathbf{dom}(\mathsf{F})$	$=_{\text{Def}}$	$\left\{\mathsf{p}^{[1]} : \mathsf{p} \in \mathsf{F}\right\}$
DEF map_2: [Restriction]	$\mathsf{F}_{\mid \mathsf{A}}$	$=_{\text{Def}}$	$\left\{\mathsf{p} \in \mathsf{F} \mid \mathsf{p}^{[1]} \in \mathsf{A}\right\}$
DEF map_3: [Image]	$\mathsf{F}^{\mid}\mathsf{x}$	$=_{\text{Def}}$	$\boldsymbol{arb}\left(\mathsf{F}_{\mid\{\mathsf{x}\}}\right)^{[2]}$
DEF map_4: [Is a map]	$\mathsf{Is\_map}(\mathsf{F})$	$\leftrightarrow_{\text{Def}}$	$\left\langle \forall \mathsf{p} \in \mathsf{F} \mid \mathsf{p} = \langle \mathsf{p}^{[1]}, \mathsf{p}^{[2]} \rangle \right\rangle$
DEF Fin: [Finitude]	$\mathsf{Finite}(\mathsf{F})$	$\leftrightarrow_{\text{Def}}$	$\left\langle \forall \mathsf{g} \in \mathscr{P}(\mathscr{P}(\mathsf{F})) \setminus \{\emptyset\}, \exists \mathsf{m} \mid \mathsf{g} \cap \mathscr{P}(\mathsf{m}) = \{\mathsf{m}\} \right\rangle$

Fig. 10.1 A few basic operations over sets and maps; two special properties

$$\left\{ term : \ iterators \mid condition \right\}$$

to specify many familiar operations and relations over sets.

Ref's second-order construct named THEORY enables one to package definitions and theorems into reusable proofware components. Besides providing theorems of which it holds the proofs, a THEORY has the ability to bring into a mathematical discourse decisive clues.[11] Like procedures of a programming language, Ref's THEORYs have input formal parameters, in exchange for whose actualization they supply useful information. Actual input parameters must satisfy a conjunction of statements, called the ASSUMPTIONS of the THEORY. A THEORY usually encapsulates the definitions of entities related to the input parameters and it supplies, along with some consequences of the assumptions, theorems talking about those internally defined entities that the THEORY returns as output parameters. After having been derived by the user once and for all inside the THEORY, the consequences of the assumptions, as well as the claims involving the output parameters, are available to be exploited repeatedly.

Two THEORY interfaces are shown in Fig. 10.2. The THEORY finiteImage awaits as input parameters a set f_0, assumed to be finite, and a *global* function g, namely one that sends every set x to a value $\mathsf{g}\,x$; whenever applied to fitting actual parameters, this finiteImage will simply produce a claim of the form $\mathsf{Finite}\big(\{\mathsf{g}\,x : x \in \mathsf{f}_0\}\big)$. The other one, reachGlob,[12] only expects a global function g; it will return the global function glob_Θ sending every set b to the smallest superset $\{b, \mathsf{g}\,b, \mathsf{g}(\mathsf{g}\,b), \ldots\}$ of

[11] In a passage echoing Abraham Robinson's 'provocative remark' which we have recalled in the Introduction through Martin's words, Jack says about this ability of THEORYs [35, p. 9]: "· · · definitions serve to 'instantiate', that is, to introduce the objects whose special properties are crucial to an intended argument. Like the selection of crucial lines, points, and circles from the infinity of geometric elements that might be considered in a Euclidean argument, definitions of this kind often carry a proof's most vital ideas". A typical case of this kind is, in arithmetic, the selection of the least natural number that meets some key property.

[12] This is a specialized variant of the THEORY reachability presented in [35, Sect. 7.3]. As seen here, the formal output parameters of a THEORY always carry a subscript Θ.

THEORY $\mathsf{finiteImage}(\mathsf{f}_0, \mathsf{g}(\mathsf{X}))$
$\quad \mathsf{Finite}(\mathsf{f}_0)$
$\Rightarrow$
$\quad \mathsf{Finite}(\{\mathsf{g}(\mathsf{x}) : \mathsf{x} \in \mathsf{f}_0\})$
END finiteImage

THEORY $\mathsf{reachGlob}(\mathsf{g}(\mathsf{X}))$
$\Rightarrow (\mathsf{glob}_\Theta)$
$\quad \langle \forall \mathsf{y}, \mathsf{x}, \mathsf{z} \mid \mathsf{y} \in \mathsf{glob}_\Theta(\mathsf{x}) \,\&\, \mathsf{z} \in \mathsf{glob}_\Theta(\mathsf{y}) \rightarrow \mathsf{z} \in \mathsf{glob}_\Theta(\mathsf{x}) \rangle$
$\quad \langle \forall \mathsf{b}, \mathsf{x}, \mathsf{y} \mid \mathsf{b} \in \mathsf{glob}_\Theta(\mathsf{b}) \,\&\, (\mathsf{x} \in \mathsf{glob}_\Theta(\mathsf{b}) \,\&\, \mathsf{y} = \mathsf{g}(\mathsf{x}) \rightarrow \mathsf{y} \in \mathsf{glob}_\Theta(\mathsf{b})) \rangle$
$\quad \langle \forall \mathsf{b}, \mathsf{t} \mid \mathsf{b} \in \mathsf{t} \,\&\, \langle \forall \mathsf{x} \in \mathsf{t} \mid \mathsf{g}(\mathsf{x}) \in \mathsf{t} \rangle \rightarrow \mathsf{glob}_\Theta(\mathsf{b}) \subseteq \mathsf{t} \rangle$
$\quad \langle \forall \mathsf{b} \mid \mathsf{glob}_\Theta(\mathsf{b}) = \{\mathsf{b}\} \cup \{\mathsf{g}(\mathsf{u}) : \mathsf{u} \in \mathsf{glob}_\Theta(\mathsf{b})\} \rangle$
$\quad \langle \forall \mathsf{b} \mid \{\mathsf{f} \subseteq \mathsf{glob}_\Theta(\mathsf{b}) \mid \langle \forall \mathsf{x} \in \mathsf{glob}_\Theta(\mathsf{b}) \mid \mathsf{x} \in \mathsf{f} \leftrightarrow \mathsf{g}(\mathsf{x}) \in \mathsf{f} \rangle\} = \{\emptyset, \mathsf{glob}_\Theta(\mathsf{b})\} \rangle$
END reachGlob

Fig. 10.2 Interfaces of two Ref THEORYS

$\{b\}$ which is closed under application of g to its own elements, as precisely stated by the claims which this THEORY will supply.

An example of the use of reachability through a global function is the construction of the set of all natural numbers intended *à la* von Neumann, which can be carried out in two steps:[13]

APPLY $\langle \mathsf{glob}_\Theta : \mathsf{count} \rangle \mathsf{reachGlob}(\mathsf{g}(\mathsf{X}) \mapsto \mathsf{X} \cup \{\mathsf{X}\}) \Rightarrow$ THM nats_a: $[Upwardcounting]$
$\quad \langle \forall\, \mathsf{y}, \mathsf{x}, \mathsf{z} \mid \mathsf{y} \in \mathsf{count}(\mathsf{x}) \,\&\, \mathsf{z} \in \mathsf{count}(\mathsf{y}) \rightarrow \mathsf{z} \in \mathsf{count}(\mathsf{x}) \rangle \,\&$
$\quad \langle \forall\, \mathsf{b}, \mathsf{x}, \mathsf{y} \mid \mathsf{b} \in \mathsf{count}(\mathsf{b}) \,\&\, (\mathsf{x} \in \mathsf{count}(\mathsf{b}) \,\&\, \mathsf{y} = \mathsf{x} \cup \{\mathsf{x}\} \rightarrow \mathsf{y} \in \mathsf{count}(\mathsf{b})) \rangle \,\&$
$\quad \langle \forall\, \mathsf{b}, \mathsf{t} \mid \mathsf{b} \in \mathsf{t} \,\&\, \langle \forall \mathsf{x} \in \mathsf{t} \mid \mathsf{x} \cup \{\mathsf{x}\} \in \mathsf{t} \rangle \rightarrow \mathsf{count}(\mathsf{b}) \subseteq \mathsf{t} \rangle \,\&$
$\quad \langle \forall\, \mathsf{b} \mid \mathsf{count}(\mathsf{b}) = \{\mathsf{b}\} \cup \{\mathsf{u} \cup \{\mathsf{u}\} : \mathsf{u} \in \mathsf{count}(\mathsf{b})\} \rangle.$

DEF nats: $[vonNeumann's naturalnumbers]$ $\mathbb{N} =_{\mathrm{Def}} \mathsf{count}(\emptyset)$.

It would be pointless to discuss here the inferential armory of Ref, because we are still in the phase of designing how to formalize the basic techniques underlying nonstandard analysis, and the expected outcome of such a formalization is best described by a plan concerning the core THEORY interfaces and by choices as to how implement some key definitions.

An important enhancement to the Zermelo-Fraenkel set theory came historically with von Neumann's introduction of an axiom,

$$\forall x\, \exists a\, \forall y \in x\ (\, a \in x \,\&\, y \notin a\,),$$

which *forbids* membership to form infinite chains $\ell_0 \ni \ell_1 \ni \ell_2 \ni \cdots$; this is tersely stated by singling out, for any given set x, a set a disjoint from x that belongs to x unless $x = \emptyset$. In Ref this principle is embodied by a construct, $\boldsymbol{arb}\,(X)$, such that

[13] What follows is not meant to imply that the definition of $\mathbb{N}$ shown is the ideal one.

$$\forall\, x \Big(\boldsymbol{arb}\,(x) \cap x = \emptyset \;\; \& \;\; \boldsymbol{arb}\,(x) \in x \cup \{x\} \Big)$$

(implying $\boldsymbol{arb}\,(\emptyset) = \emptyset$). The meaning of $\boldsymbol{arb}$ is competently handled by a most basic inference method of Ref.

To appreciate the usefulness of $\boldsymbol{arb}$, consider the THEORY whose interface appears on the left of Fig. 10.3. Upon receipt of a set n_0 that meets a given property P, this THEORY will return a set $\mathsf{transfInd}_\Theta$ still enjoying P but none of whose elements satisfies P. In its hidden internal working, transfInduction first applies the THEORY reachGlob seen in Fig. 10.2 to $g(X) = \boldsymbol{arb}\,(\{u \in X \mid P(u)\})$ and then applies the resulting glob_Θ to n_0 to get a set $N_0 = \{n_0\,,\; g\,n_0\,,\; g(g\,n_0), \ldots, \emptyset\}$ such that $\boldsymbol{arb}\,(\{w \in N_0 \mid P(w)\})$ is the sought $\mathsf{transfInd}_\Theta$.

In Ref the well-foundedness of membership also lies behind a definition mechanism based on $\in$-*recursion*, shown at work with the specification of img in Fig. 10.4 and which we will repeatedly use in the ongoing. A discussion about the syntax of $\in$-recursive definitions can be found in [35, pp. 216–217]); concrete illustrations of it will suffice here. A basic example is

$$\mathsf{rk}\,(X) \;=_{\text{Def}}\; \bigcup \Big\{ \mathsf{rk}\,(y) \cup \{\mathsf{rk}\,(y)\} \;:\; y \in X \Big\}\,,$$

defining the RANK of a set X. The mechanism at stake is akin to recursion as used in computer programming; like it, it resorts to a base case to avoid circularity: in fact, $\mathsf{rk}\,(X) = \emptyset$ when $X = \emptyset$, since obviously $\{\mathsf{rk}\,(y) \cup \{\mathsf{rk}\,(y)\} \;:\; y \in \emptyset\} = \emptyset$. But $\mathsf{rk}\,(X)$ might also be an infinite set (actually, a transfinite ordinal), a situation which will occur, e.g., when X is infinite or has some infinite elements.

THEORY $\mathsf{transfInduction}(\mathsf{n}_0, \mathsf{P}(\mathsf{X}))$	THEORY $\mathsf{finInduction}(\mathsf{n}_0, \mathsf{P}(\mathsf{X}))$
$\quad \mathsf{P}(\mathsf{n}_0)$	$\quad \mathsf{P}(\mathsf{n}_0)\,\&\,\mathsf{Finite}\,(\mathsf{n}_0)$
$\Rightarrow (\mathsf{transfInd}_\Theta)$	$\Rightarrow (\mathsf{fInd}_\Theta)$
$\quad \mathsf{P}(\mathsf{transfInd}_\Theta)\,\&\,\langle \forall \mathsf{k} \in \mathsf{transfInd}_\Theta \mid \neg \mathsf{P}(\mathsf{k}) \rangle$	$\quad \mathsf{P}(\mathsf{fInd}_\Theta)\,\&\,\langle \forall \mathsf{k} \subsetneq \mathsf{fInd}_\Theta \mid \neg \mathsf{P}(\mathsf{k}) \rangle$
END transfInduction	END finInduction

Fig. 10.3 Transfinite induction contrasted with finite induction. The former exploits the well-foundedness of $\in$ while the latter exploits the well-foundedness of $\subsetneq$ over finite sets. Other classical forms of induction, e.g., arithmetic induction or induction over ordinals, can be conveniently hooked to membership or inclusion

$$\begin{array}{l} \mathsf{img}(\mathsf{I}, \mathsf{B}) \;=_{\text{Def}}\; \textbf{if } \mathsf{I} = \emptyset \textbf{ then } \mathsf{B} \textbf{ else } \boldsymbol{arb}\,(\{\mathsf{g}(\mathsf{img}(\mathsf{j}, \mathsf{B})) : \mathsf{j} \in \mathsf{I}\})\ \textbf{fi} \\ \mathsf{glob}_\Theta(\mathsf{B}) \;=_{\text{Def}}\; \{\mathsf{img}(\mathsf{i}, \mathsf{B}) :\; \mathsf{i} \in \boldsymbol{\sigma}_\infty\} \end{array}$$

Fig. 10.4 A viable specification of the iterated images of g and of the output symbol glob_Θ inside the THEORY reachGlob of Fig. 10.2. Here $\boldsymbol{\sigma}_\infty$ is a Ref's built-in constant subject to the assumption that $\boldsymbol{\sigma}_\infty \neq \emptyset\ \&\ \langle \forall \mathsf{x} \in \boldsymbol{\sigma}_\infty \mid \{\mathsf{x}\} \in \boldsymbol{\sigma}_\infty \rangle$

Let us digress briefly. The α's satisfying the equality $\alpha = \mathsf{rk}(\alpha)$ turn out to be precisely the sets known, after von Neumann, as *ordinal numbers*;[14] and it is not hard to prove, about the indexed class of sets which satisfy the conditions

$$\begin{aligned} V_\emptyset &= \emptyset\,, \\ V_{\gamma\cup\{\gamma\}} &= \mathscr{P}\big(V_\gamma\big) \quad \text{for every ordinal number } \gamma\,, \\ V_\lambda &= \textstyle\bigcup_{\beta\in\lambda} V_\beta \quad \text{for every nonnull ordinal } \lambda \text{ not of the form } \gamma\cup\{\gamma\} \end{aligned}$$

—historically called the *cumulative hierarchy*—, that V_α consists, for each ordinal α, of all sets whose ranks lie below α. Now consider the property

$$\mathscr{V}(L) \;\leftrightarrow_{\text{Def}}\; L = \bigcup\{\mathscr{P}(\ell) \;:\; \ell \in L \;\mid\; \mathscr{V}(\ell)\}\;.$$

In this new instance of $\in$-recursion, the reader can recognize a streamlined definition of the stages of the cumulative hierarchy: as one readily sees, $\mathscr{V}(\emptyset)$ holds; more generally, one can show that $\mathscr{V}(L)$ is logically equivalent to the existence of an ordinal α such that $L = V_\alpha$. We do not prove this fact but do call attention to it because a similar change of perspective will motivate our formalization of superstructures in the following section.

10.7 Top-Down Recognition of Superstructure Stages

Concerning the unusual way, just hinted at, of approaching the cumulative hierarchy, one might contend that it is presumably harder—or, if anything, less transparent—to infer directly from the definition of $\mathscr{V}(L)$ a statement such as

$$\Big(\mathscr{V}(L') \;\&\; \mathscr{V}(L'')\Big) \;\rightarrow\; (L' \subsetneq L'' \leftrightarrow L' \in L'')$$

than to prove, for any pair α, β of ordinals, the biimplications

$$\big(V_\beta \subsetneq V_\alpha \leftrightarrow \beta \in \alpha\big) \;\;\&\;\; \big(V_\beta \in V_\alpha \leftrightarrow \beta \in \alpha\big)\,.$$

A tentative reply is that transfinite induction of the kind schematized in Fig. 10.3 (left) is often a shortcut compared to a proof pattern relying on the theory of ordinals. On

[14] A common definition of ordinals, owing to a simplification due to Raphael Robinson, is:

$$\mathsf{Ord}(U) \;\leftrightarrow_{\text{Def}}\; \forall x\,(x \in U \rightarrow x \subseteq U)\;\&\;\forall x\,\forall y\big(\{x,y\} \subseteq U \rightarrow (x \in y \vee y \in x \vee x = y)\big)\,.$$

a smaller scale, as will now be seen, we can treat superstructures without numbering their stages: with virtually no recourse to natural numbers.[15]

We can exploit recursion to describe sets L which are *stages* of a superstructure. The first of the three definitions shown below is $\in$-recursive and specifies a function seeking a set s of *individuals* (recall Sect. 10.2) such that $\mathsf{s}_m = L$ for some $m \in \mathbb{N}$; if such an s exists, it can be found by repeated extraction

$$L \ni \log L \ni \log\log L \ni \cdots \ni \mathsf{s}$$

of the 'logarithm' of L, where $\ell = \log L$ momentarily means that $L = \ell \cup \mathscr{P}(\ell)$ (needless to say, this equation has either one or no solution—in the former case, $\emptyset \in L$ and hence L cannot be regarded as a set of individuals):

$$\begin{aligned}\mathsf{basis}(L) \;=_{\text{Def}}\; & \textbf{if } \emptyset \notin L \;\&\; L \cap \bigcup L = \emptyset \textbf{ then } L \\ & \textbf{elseif } \big(\exists\, \ell \mid L = \ell \cup \mathscr{P}(\ell) \;\&\; \mathscr{P}(\ell) \cap \bigcup(\ell \setminus \mathscr{P}(\ell)) = \emptyset\big) \\ & \textbf{then } \boldsymbol{arb}\,(\{\mathsf{basis}(\ell) : \ell \in L \mid L = \ell \cup \mathscr{P}(\ell)\}) \\ & \textbf{else } \{\emptyset\} \textbf{ fi}\,;\end{aligned}$$

$$\mathsf{Stage}(L, S) \;\leftrightarrow_{\text{Def}}\; L = \emptyset \vee \big(\mathsf{basis}(L) = S \;\&\; S \neq \{\emptyset\}\big)\,;$$

$$\begin{aligned}\mathsf{Ur}\,(S) \;\leftrightarrow_{\text{Def}}\; & \emptyset \notin S \;\&\; S \cap \bigcup S = \emptyset \;\& \\ & \big(\forall\, \ell \mid \mathsf{Stage}(\ell, S) \;\rightarrow\; \mathscr{P}(\ell) \cap \bigcup S = \emptyset\big)\,.\end{aligned}$$

The chain $L = L_0$, $L_{n+1} = \log L_n$ of logarithms surely has finite length but may end with a set L_m such that either $\emptyset \in L_m$ or $L_m \cap \bigcup L_m \neq \emptyset$ holds, in which cases L_m cannot serve as a set of individuals. When this happens, $\mathsf{basis}(L)$ will flag the failure by returning $\{\emptyset\}$; but failure can be detected earlier during the descent, should $\mathscr{P}(L_n) \cap \bigcup(L_n \setminus \mathscr{P}(L_n))$ be nonnull at some point. The predicate $\mathsf{Stage}(L, S)$ indicates L as a potential stage of the superstructure—if any—generated by its 'ultimate logarithm' $S = \mathsf{basis}(L)$ when the latter is obtained without failure; but even when so, S does not qualify as a set of individuals unless one can indefinitely ascend, starting with S, through stages none of which reveals the inner structure of its elements. The property $\mathsf{Ur}\,(S)$ captures the sense of our last remark.

Under the assumption $\mathsf{Ur}\,(\mathsf{s}_0)$, we have in fact checked with the assistance of Ref that $\widehat{\mathsf{s}}_0$ behaves as desired (see Fig. 10.5), even though genuine individuals ('*urelemente*' of the nature set forth in [14, p. 11]) do not exist in the von Neumann cumulative universe of all sets.

The interface, shown in Fig. 10.5, of the THEORY $\mathsf{superstructure}$ may look intimidating, the cause being that it exploits the *property* $\langle \exists \ell \mid \mathsf{Stage}(\ell, \mathsf{s}_0)\ \&\ X \in \ell \rangle$

[15] Natural numbers will play an irreplaceable role in the informal arguments providing the rationale for the formal constructions that follow; within the formal treatment, their collection $\mathbb{N}$ will act as a set whose infinitude is easiest to prove (and infinite sets will be crucial in Sect. 10.8).

THEORY universe($\mathfrak{U}(X)$)
$\mathfrak{U}(\emptyset)$ & $\left\langle \forall x, y \mid \mathfrak{U}(x) \,\&\, \mathfrak{U}(y) \to \mathfrak{U}(\{x,y\}) \right\rangle$
$\left\langle \forall x, y, z \mid \mathfrak{U}(x) \,\&\, \mathfrak{U}(y) \,\&\, \{y,z\} \subseteq x \to \mathfrak{U}(z) \right\rangle$
$\Rightarrow$
$\left\langle \forall x, y \mid \mathfrak{U}(x) \,\&\, \mathfrak{U}(y) \to \mathfrak{U}(\langle x,y\rangle) \right\rangle$
$\left\langle \forall f, x \mid \mathsf{Is\_map}(f) \,\&\, \mathfrak{U}(f) \,\&\, \left\langle \forall x \in \mathbf{dom}(f) \mid \mathfrak{U}(x) \right\rangle \,\&\, \left(f = \emptyset \vee \left\langle \exists q \in f \mid \mathfrak{U}(q) \right\rangle \right) \to \mathfrak{U}(f|x) \right\rangle$
END universe

THEORY superstructure(s_0)
Ur(s_0)
$\Rightarrow$ (sstr_Θ)
$\mathsf{Stage}(\emptyset, s_0) \,\&\, \mathsf{Stage}(s_0, s_0) \,\&\, s_0 \neq \{\emptyset\}$ & $\left\langle \forall \ell \mid \mathsf{Stage}(\ell, s_0) \,\&\, \ell \neq \emptyset \to s_0 \subseteq \ell \right\rangle$ & $s_0 \cap \bigcup (s_0 \setminus \mathscr{P}(s_0)) = \emptyset \,\&\, s_0 \setminus \mathscr{P}(s_0) = s_0$
$\left\langle \forall \ell \mid \mathsf{Stage}(\ell, s_0) \,\&\, s_0 = \emptyset \to \ell \subseteq \mathscr{P}(\ell) \right\rangle$ & $\left\langle \forall \ell \mid \mathsf{Stage}(\ell, s_0) \,\&\, s_0 = \emptyset \to \mathsf{Stage}(\mathscr{P}(\ell), s_0) \right\rangle$
$\left\langle \forall \ell \mid \mathsf{Stage}(\ell, s_0) \to (\ell = \emptyset \vee s_0 = \ell \setminus \mathscr{P}(\ell)) \,\&\, \ell \cap \bigcup (\ell \setminus \mathscr{P}(\ell)) = \emptyset \right\rangle$
$\left\langle \forall \ell \mid \mathsf{Stage}(\ell, s_0) \,\&\, (s_0 = \emptyset \vee \ell \neq \emptyset) \to \mathsf{Stage}(\ell \cup \mathscr{P}(\ell), s_0) \right\rangle$ & $\mathsf{Stage}(s_0 \cup \mathscr{P}(s_0), s_0) \,\&\, \emptyset \in s_0 \cup \mathscr{P}(s_0)$
$\left\langle \forall x, \ell \mid x \in \ell \setminus s_0 \,\&\, \mathsf{Stage}(\ell, s_0) \to \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, \mathscr{P}(x) \subseteq \mathscr{P}(h) \right\rangle \right\rangle$
$\left\langle \forall x, \ell, y, m, z \mid x \in \ell \,\&\, y \in m \,\&\, \{y,z\} \subseteq x \,\&\, \mathsf{Stage}(\ell, s_0) \,\&\, \mathsf{Stage}(m, s_0) \to \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, z \in h \right\rangle \right\rangle$
$\left\langle \forall \ell, m \mid \ell \not\subseteq m \,\&\, \mathsf{Stage}(\ell, s_0) \,\&\, \mathsf{Stage}(m, s_0) \to m \subseteq \ell \right\rangle$
$\left\langle \forall x, \ell, y, m \mid x \in \ell \,\&\, y \in m \,\&\, \mathsf{Stage}(\ell, s_0) \,\&\, \mathsf{Stage}(m, s_0) \to \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, \{x,y\} \in h \right\rangle \right\rangle$
$\left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, \emptyset \in k \right\rangle$ &
$\left\langle \forall x, y \mid \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, x \in h \right\rangle \,\&\, \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, y \in k \right\rangle \to \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, \{x,y\} \in k \right\rangle \right\rangle$ &
$\left\langle \forall x, y, z \mid \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, x \in h \right\rangle \,\&\, \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, y \in k \right\rangle \,\&\, \{y,z\} \subseteq x \to \left\langle \exists m \mid \mathsf{Stage}(m, s_0) \,\&\, z \in m \right\rangle \right\rangle$
$\left\langle \forall x, y \mid \left\langle \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, x \in h \right\rangle \,\&\, \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, y \in k \right\rangle \to \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, \langle x,y\rangle \in k \right\rangle \right\rangle$ &
$\left\langle \forall f, x \mid \mathsf{Is\_map}(f) \,\&\, \left\langle \exists k \mid \mathsf{Stage}(k, s_0) \,\&\, f \in k \right\rangle \,\&\, \left\langle \forall u \in \mathbf{dom}(f), \exists h \mid \mathsf{Stage}(h, s_0) \,\&\, u \in h \right\rangle \,\&\,$
$\left(f = \emptyset \vee \left\langle \exists q \in f, \exists m \mid \mathsf{Stage}(m, s_0) \,\&\, q \in m \right\rangle \right) \to \quad \left\langle \exists n \mid \mathsf{Stage}(n, s_0) \,\&\, f|x \in n \right\rangle \right\rangle$
$\left\langle \forall x \mid \left\langle \exists \ell \mid \mathsf{Stage}(\ell, s_0) \,\&\, x \in \ell \right\rangle \leftrightarrow x \in \mathsf{sstr}_\Theta \right\rangle$
END superstructure

Fig. 10.5 Interfaces of the THEORYS of universes and superstructures

as a temporary surrogate of the sought $\widehat{s}_0$. Only its final claim shows that the X's enjoying that property form a set, namely the output parameter sstr_Θ to be then actualized as $\widehat{s}_0$ outside the THEORY; even so we can exploit the said property as a universe to get, through the THEORY universe, derived closure properties. Observe, in fact, that the second-to-last and penultimate claim of superstructure match the assumptions of universe and its internally derived conclusions.

The moral is that our recursive characterization of the stages of a superstructure disclosed handy patterns to our formal reasoning about them; however, at one point we had to resort to a construction from below, closer in spirit to [14, Sect. 1.3]: this happened when it came to ascertaining that the union-class of all the stages is, in fact, a set. For that purpose, we applied the THEORY reachGlob (see Fig. 10.2 above) to the actual input parameter **if** $X = \emptyset \,\&\, s_0 \neq \emptyset$ **then** s_0 **else** $X \cup \mathscr{P}(X)$ **fi**, thus getting a function glob whence the sought superstructure was obtained simply by taking $\mathsf{sstr}_\Theta = \bigcup \mathsf{glob}(\emptyset)$. The following triad of equations conveys the idea, in functionally equivalent terms:

$$\mathsf{nextStage}(L) = \textbf{if } L = \emptyset \,\&\, \mathsf{s}_0 \neq \emptyset \textbf{ then } \mathsf{s}_0 \textbf{ else } L \cup \mathscr{P}(L) \textbf{ fi} \,,$$

$$\mathsf{stage}(I) = \boldsymbol{arb}\,(\{\mathsf{nextStage}(\mathsf{stage}(j)) : \; j \in I\}) \;,$$

$$\mathsf{sstr}_\Theta = \bigcup \{\mathsf{stage}(i) : \; i \in \boldsymbol{\sigma}_\infty\} \;.$$

These equations, in fact, adjust the construction of Fig. 10.4 to the case at hand; as said under that figure, $\boldsymbol{\sigma}_\infty$ is a Ref's built-in witnessing that infinite sets exist.

10.8 Forging Companion Sets of Individuals

When undertaking the construction of a standard universe, in practice one starts with a pre-defined, infinite basis—say the set $\mathbb{R}$ of all real numbers—whose elements may have an inner structure that prevents their direct use as individuals. If so, how can we conceal their structure? We need a technique for converting a set s' whatsoever into a set s'' so that $\mathsf{Ur}\,(s'')$ holds and there is a one-one correspondence between s' and s''.

One plainly sees that $\mathsf{Ur}\,(s'')$ cannot hold if any set of finite rank belongs to s''; on the other hand, imposing that $\emptyset \notin s''$ and that all elements of elements of s'' share the same infinite rank r suffices to ensure that $\mathsf{Ur}\,(s'')$ holds—one shows inductively, in fact, that each stage originating from s'' is the union of a set of finite rank with a set whose elements have ranks exceeding r. This observation makes it rather easy to conceive an injection ur whose domain is the given s' and whose set of values, $s'' = \{\mathsf{ur}\,x : \; x \in s'\}$, can serve as basis in place of s' in the construction of the standard superstructure. Should any rationale arise for doing so, we can even tune the range of s'' by means of an auxiliary 'gauge' set c', as suggested by the interface of the THEORY $\mathsf{urification}$ in Fig. 10.6.

This THEORY receives sets s', c' such that $s' \cup c'$—and hence $\mathsf{rk}\,(s' \cup c')$—is infinite; it manufactures and produces in output a function ur_Θ sending injectively each $x \in s'$ to a set $\mathsf{ur}_\Theta(x)$ all of whose elements have rank $\mathsf{rk}\,(s' \cup c')^+ = \mathsf{rk}\,(s' \cup c') \cup$

Fig. 10.6 Gauged transformation of a set s' whatsoever into a set of individuals

DEF $=_{\text{Def}} \mathsf{X} \cup \{\mathsf{X}\}$

THM $\neg\mathsf{Finite}(\mathsf{R}) \,\&\, \emptyset \notin \mathsf{S} \,\&\, \langle \forall \mathsf{u} \in \bigcup \mathsf{S} \mid \mathsf{rk}(\mathsf{u}) = \mathsf{R} \rangle \;\rightarrow\; \mathsf{Ur}(\mathsf{S})$

THEORY $\mathsf{urification}(\mathsf{s}', \mathsf{c}')$
$\quad \neg\mathsf{Finite}(\mathsf{s}') \vee \neg\mathsf{Finite}(\mathsf{c}')$
$\Rightarrow (\mathsf{ur}_\Theta)$
$\quad \langle \forall \mathsf{x} \in \mathsf{s}', \mathsf{y} \in \mathsf{s}' \mid \mathsf{ur}_\Theta(\mathsf{x}) = \mathsf{ur}_\Theta(\mathsf{y}) \;\rightarrow\; \mathsf{x} = \mathsf{y} \rangle$
$\quad \langle \forall \mathsf{v} \in \bigcup \{\mathsf{ur}_\Theta(\mathsf{x}) : \mathsf{x} \in \mathsf{s}'\} \mid \mathsf{rk}(\mathsf{v}) = \mathsf{rk}(\{\mathsf{s}' \cup \mathsf{c}'\}) \rangle$
$\quad \langle \forall \mathsf{u} \in \{\mathsf{ur}_\Theta(\mathsf{x}) : \mathsf{x} \in \mathsf{s}'\} \mid \mathsf{rk}(\mathsf{u}) = \mathsf{rk}(\{\mathsf{s}' \cup \mathsf{c}'\})^+ \rangle$
$\quad \mathsf{Ur}(\{\mathsf{ur}_\Theta(\mathsf{x}) : \; \mathsf{x} \in \mathsf{s}'\})$
END $\mathsf{urification}$

$\{\mathsf{rk}(s' \cup c')\}$, where $R^+ =_{\text{Def}} R \cup \{R\}$. The definition of ur_Θ—internally hidden, insofar as immaterial outside the THEORY urification—could well be

$$\mathsf{ur}_\Theta(\mathsf{X}) =_{\text{Def}} \left\{ \mathsf{s}' \setminus \{\mathsf{X}\} \cup \{\mathsf{s}' \cup \mathsf{c}'\} \right\} .$$

What really counts to us is that $\mathsf{Ur}\left(\{\mathsf{ur}_\Theta(x) : x \in s'\}\right)$ holds, as we aimed at.

To see a more sophisticated exploitation of the THEORY at hand, suppose next that we are given a set s along with an infinite set i' that we want to use as index set for enlarging s, seen as a standard set of individuals, into a set w of nonstandard individuals. To ease the discussion, we momentarily dismiss the concurrence issue debated in Sect. 10.4; we will content ourselves with an ultrafilter none of whose elements is a finite set, over (a counterpart i'' of) i'.

First move. Convert i' into a set i'' so that all indices j in i'' have the same infinite rank r, exceeding the rank of s, and there is a one-one correspondence $\mathsf{u}(X)$ between i' and i'':

$$\begin{array}{l} \text{APPLY } (\mathsf{ur}_\Theta : \mathsf{u}) \ \mathsf{urification}(\, \mathsf{s}' \mapsto i',\ \mathsf{c}' \mapsto s\,) \Rrightarrow \cdots \\ \text{DEF} \quad i'' =_{\text{Def}} \{\mathsf{u}(\mathsf{x}) : \mathsf{x} \in i'\} \end{array}$$

Second move. Observe that when $\mathscr{W}$ is a set of functions from i'' to s then each element of $\bigcup \mathscr{W}$ is an ordered pair $\langle j, x\rangle = \{\{j\}, \{j, x\}\}$, whose rank is infinite. Trivially $\emptyset \notin \mathscr{W}$ and hence $\mathsf{Ur}(\mathscr{W})$ holds.

Third move. Introduce an *ultrafilter* æ such that

$$\bigcup \text{æ} = i'' \ \text{ and } \ \text{æ} \supseteq \{i'' \setminus \{j\} : j \in i''\} ,$$

and at this point specify $\mathscr{W}$ as follows:

$$\begin{array}{rl} \rho(g) & =_{\text{Def}} \boldsymbol{arb}\left(\{h \in s^{i''} \mid \{j \in i'' \mid h\, j = g\, j\} \in \text{æ}\}\right) , \\ \mathscr{W} & =_{\text{Def}} \left\{\rho(g) : g \in s^{i''}\right\} . \end{array}$$

Now regard this $\mathscr{W}$ and its subset

$$\mathscr{S} =_{\text{Def}} \{h \in \mathscr{W} \mid (\exists\, y \mid \mathsf{dom}(\,(i'' \times \{y\}) \cap h) \in \text{æ})\} ,$$

respectively, as the 'wide' and the 'small' set of all nonstandard individuals and of the standard ones: it should be clear that $\mathscr{S}$ can act as a counterpart of the original s, in view of the natural correspondence between the two.

What precedes has offered clues about how to implement the THEORY whose interface is shown in the lower part of Fig. 10.7.

$$
\begin{array}{l}
\textsc{Def} \qquad \mathsf{X} \bigtriangleup \mathsf{Y} \;=_{\text{Def}}\; \mathsf{X} \cup \mathsf{Y} \setminus \mathsf{X} \cap \mathsf{Y} \\
\textsc{Def}\ \ \mathsf{Ultrafilter}(\text{Æ}) \;\leftrightarrow_{\text{Def}}\; \Big\langle \forall \mathsf{x} \mid (\mathsf{x} \cap \bigcup \text{Æ}) \in \text{Æ} \vee (\bigcup \text{Æ} \setminus \mathsf{x}) \in \text{Æ} \Big\rangle \,\&\, \\
\qquad\qquad\qquad\qquad\qquad\quad \Big\langle \forall \mathsf{x} \in \text{Æ}, \mathsf{y} \in \text{Æ} \mid \mathsf{x} \cap \mathsf{y} \in \text{Æ} \setminus \{\emptyset\} \Big\rangle \\
\textsc{Thm}\ \ \Big\langle \forall \mathsf{x} \in \mathsf{B}, \mathsf{y} \in \mathsf{B} \mid \mathsf{x} \cap \mathsf{y} \in \mathsf{B} \setminus \{\emptyset\} \,\&\, \mathsf{x} \subseteq \ddot{\mathsf{I}} \Big\rangle \,\&\, \mathsf{B} \neq \emptyset \;\rightarrow \\
\qquad\qquad\qquad\qquad\qquad\quad \Big\langle \exists \text{æ} \mid \mathsf{Ultrafilter}(\text{æ}) \,\&\, \mathsf{B} \subseteq \text{æ} \,\&\, \ddot{\mathsf{I}} = \bigcup \text{æ} \Big\rangle \\
\textsc{Theory}\ \mathsf{individuation}\big(\mathsf{s}, \mathsf{i}''\big) \\
\quad \mathsf{s} \neq \emptyset \\
\quad \neg \mathsf{Finite}(\mathsf{i}'') \\
\quad \Big\langle \forall \mathsf{j} \in \mathsf{i}'' \mid \mathsf{rk}(\mathsf{s}) \in \mathsf{rk}(\mathsf{j}) \,\&\, \mathsf{rk}(\mathsf{j}) = \mathsf{rk}(\boldsymbol{arb}(\mathsf{i}'')) \Big\rangle \\
\Rightarrow \big(\text{æ}_\Theta, \mathsf{w}_\Theta\big) \\
\quad \mathsf{Ultrafilter}(\text{æ}_\Theta) \,\&\, \bigcup \text{æ}_\Theta = \mathsf{i}'' \,\&\, \Big\{ \mathsf{i}'' \setminus \{\mathsf{j}\} : \mathsf{j} \in \mathsf{i}'' \Big\} \subseteq \text{æ}_\Theta \\
\quad \Big\langle \forall \mathsf{g} \in \mathsf{w}_\Theta, \mathsf{h} \in \mathsf{w}_\Theta \setminus \{\mathsf{g}\} \mid \mathsf{Svm}(\mathsf{g}) \,\&\, \mathbf{dom}(\mathsf{g}) = \mathsf{i}'' \,\&\, \mathbf{dom}(\mathsf{g} \bigtriangleup \mathsf{h}) \in \text{æ}_\Theta \Big\rangle \\
\quad \Big\langle \forall \mathsf{y} \mid \Big\langle \exists \mathsf{h} \in \mathsf{w}_\Theta \mid \mathbf{dom}((\mathsf{i}'' \times \{\mathsf{y}\}) \cap \mathsf{h}) \in \text{æ}_\Theta \Big\rangle \leftrightarrow \mathsf{y} \in \mathsf{s} \Big\rangle \\
\quad \Big\langle \forall \mathsf{g} \mid \mathsf{Svm}(\mathsf{g}) \,\&\, \mathbf{dom}(\mathsf{g}) = \mathsf{i}'' \,\&\, \mathbf{dom}((\mathsf{i}'' \times \mathsf{s}) \cap \mathsf{g}) \in \text{æ}_\Theta \;\rightarrow \\
\qquad\qquad\qquad\qquad\qquad\quad \Big\langle \exists \mathsf{h} \in \mathsf{w}_\Theta \mid \mathbf{dom}(\mathsf{g} \bigtriangleup \mathsf{h}) \notin \text{æ}_\Theta \Big\rangle \Big\rangle \\
\quad \neg \mathsf{Finite}(\mathsf{rk}(\boldsymbol{arb}(\mathsf{i}''))) \,\&\, \emptyset \notin \mathsf{w}_\Theta \,\&\, \Big\langle \forall\, p \in \bigcup \mathsf{w}_\Theta \mid \mathsf{rk}(p) = \mathsf{rk}(\boldsymbol{arb}(\mathsf{i}''))^{++} \Big\rangle \\
\quad \mathsf{Ur}(\mathsf{w}_\Theta) \\
\textsc{End}\ \ \mathsf{individuation}
\end{array}
$$

Fig. 10.7 Transformation of a set s into a set w_Θ of nonstandard individuals

10.9 Set-Encoding of Bounded-Quantifier Formulae

Before we can exploit Ref to state and prove propositions such as the transfer principle (not to mention Łoś's theorem, see Sect. 10.3), we must devise a set-encoding of terms and formulae that enables easy specifications of how to

(A) evaluate a term or formula under a set-assignment for its variables,
(B) determine the truth value of a sentence,
(C) replace a free variable by a constant within a term or formula,

and the like. Then we will be able to reason formally with Ref about the languages of specific universes.

The set-theoretic representation of terms and formulae can be conceived of rather liberally. By seeing each universe $\mathscr{U}$ as embedded in the class of all sets, which is Ref's domain of discourse, we will in particular

- treat the different languages $\mathscr{L}_{\mathscr{U}}$ by a single encoding instead of separately,
- specify the function val that evaluates a 'term' t under a set-assignment v for the variables occurring in it so that $\mathsf{val}(t, v)$ yields a result even when t does not encode a term.

THEORY termEncoding()
⇒ (cst_Θ , Pair_Θ , Appl_Θ , lft_Θ , rgt_Θ)
$\langle \forall c, k \mid \mathsf{cst}_\Theta(c) = \mathsf{cst}_\Theta(k) \rightarrow c = k \rangle$
$\langle \forall c, i, p, q, x, y, z \mid i \in \mathbb{N} \,\&\, \mathsf{Pair}_\Theta(p) \,\&\, \mathsf{Appl}_\Theta(q) \rightarrow \{\mathsf{cst}_\Theta(c), i, p, q\} \neq \{x, y, z\} \rangle$
$\langle \forall x, y \mid \emptyset \notin \{x, y\} \rightarrow \langle \exists p \mid \langle \forall s \mid \mathsf{Pair}_\Theta(s) \,\&\, \mathsf{lft}_\Theta(s) = x \,\&\, \mathsf{rgt}_\Theta(s) = y \leftrightarrow s = p \rangle \rangle \rangle$
$\langle \forall x, y \mid \emptyset \notin \{x, y\} \rightarrow \langle \exists q \mid \langle \forall s \mid \mathsf{Appl}_\Theta(s) \,\&\, \mathsf{lft}_\Theta(s) = x \,\&\, \mathsf{rgt}_\Theta(s) = y \leftrightarrow s = q \rangle \rangle \rangle$
$\langle \forall s \mid \mathsf{Pair}_\Theta(s) \vee \mathsf{Appl}_\Theta(s) \rightarrow \emptyset \notin \{\mathsf{lft}_\Theta(s), \mathsf{rgt}_\Theta(s)\} \rangle$
$\langle \forall t \mid \{\mathsf{lft}_\Theta(t), \mathsf{rgt}_\Theta(t)\} \subseteq t \cup \{\emptyset\} \rangle$
END termEncoding

Fig. 10.8 THEORY about the set-encoding of terms

About one feature of the representation of the syntax, we see no reason for being flexible: each variable $\boldsymbol{x}_n$ will be encoded by its subscript n, a positive integer.

The THEORY interface displayed in Fig. 10.8 formulates the constraints to which we submit our encoding of terms, effected via two properties, Pair and Appl, and three functions: lft , rgt, and cst. The two properties are meant to indicate which sets encode terms of the respective forms $\langle \ell, r \rangle$ and $(\ell \restriction r)$; $\mathsf{lft}(p)$ and $\mathsf{rgt}(p)$ will provide, when applied to a set p that encodes a term of the form $\langle \ell, r \rangle$, the two sets encoding the immediate subterms, ℓ and r respectively; $\mathsf{lft}(q)$ and $\mathsf{rgt}(q)$ will behave likewise when q encodes a term of the form $(\ell \restriction r)$. As for cst, it will send each set c to a constant $\boldsymbol{c}$ designating it *univocally*: not only $\mathsf{cst}(c) \neq \mathsf{cst}(k)$ must hold whenever $c \neq k$, but we require also that $\neg\mathsf{Pair}(\mathsf{cst}(c))$, $\neg\mathsf{Appl}(\mathsf{cst}(c))$, and $\mathsf{cst}(c) \notin \mathbb{N}$, to avoid 'collision' between $\mathsf{cst}(c)$ and any set encoding a non-constant term. Unambiguous readability also demands that $p \notin \mathbb{N}$, $q \notin \mathbb{N}$, and $p \neq q$ hold when $\mathsf{Appl}(p)$ and $\mathsf{Pair}(q)$ hold. This is the rationale behind the first two claims issued by the THEORY termEncoding. To understand the third, fourth, and fifth claim thereof, think of $\emptyset$ as non-encoding set *par excellence*: for every pair x, y of sets which differ from $\emptyset$, we want unique sets p, q to exist such that $\mathsf{lft}(p) = \mathsf{lft}(q) = x$, $\mathsf{rgt}(p) = \mathsf{rgt}(q) = y$, and $\mathsf{Pair}(p)$, $\mathsf{Appl}(q)$ hold; conversely, we want $\mathsf{lft}(s)$ and $\mathsf{rgt}(s)$ to differ from $\emptyset$ when either $\mathsf{Pair}(s)$ or $\mathsf{Appl}(s)$ holds. The last claim of termEncoding plays a technical role: since the only built-in kind of recursion in Ref is $\in$-recursion, by imposing that the immediate subterms of any compound term t (as encoded by a set) *belong* to t, this claim will ease the recursive definition of functions over all terms.

Figure 10.9 suggests one way of implementing the wanted functions and properties inside termEncoding, based on the remark that when $\emptyset \notin \{x, y\}$ the projections x, y can be retrieved from both variants $\langle x, y \rangle \cup \{x, y\}$, $\langle x, y \rangle \cup \{x, y\} \cup \{\emptyset\}$ (the former of which equals $\{x, y\}^+ \cup \{\{x\}\}$) of Kuratowski's pair $\langle x, y \rangle$.

Assuming that terms are encoded according to a quintuple such as the one produced by termEncoding, it is easy to implement their evaluation thus developing the THEORY evalTerm whose interface is shown in Fig. 10.10.

$$
\begin{array}{|l|}
\hline
\mathsf{cst}(C) =_{\text{Def}} \{\{C\}\} \\
\mathsf{lft}(T) =_{\text{Def}} \boldsymbol{arb}\left(\{x:\ x\in T, y\in T \mid T\setminus\{\emptyset\} = \{x,y\}^{+}\cup\{\{x\}\}\}\right) \\
\mathsf{rgt}(T) =_{\text{Def}} \boldsymbol{arb}\left(\{y:\ x\in T, y\in T \mid T\setminus\{\emptyset\} = \{x,y\}^{+}\cup\{\{x\}\}\}\right) \\
\mathsf{Pair}(P) \leftrightarrow_{\text{Def}} \emptyset\notin P \,\&\, \emptyset\notin\{\mathsf{lft}(P), \mathsf{rgt}(P)\} \\
\mathsf{Appl}(Q) \leftrightarrow_{\text{Def}} \emptyset\in Q \,\&\, \emptyset\notin\{\mathsf{lft}(Q), \mathsf{rgt}(Q)\} \\
\hline
\mathsf{Lit}(L) \leftrightarrow_{\text{Def}} \left\langle \exists s,t \mid L\setminus\{\emptyset\} = \langle s,t\rangle \ \&\ t\notin\{\emptyset,s\} \ \&\ s\neq\emptyset \right\rangle \\
\mathsf{Qnt}(Q) \leftrightarrow_{\text{Def}} \left\langle \exists x,t \mid Q\setminus\{\emptyset\} = \langle x,t\rangle \ \&\ t\notin\{\emptyset,x\} \ \&\ x\in\mathbb{N} \right\rangle \\
\hline
\end{array}
$$

Fig. 10.9 A viable implementation of the quintuple needed to encode terms, followed by encodings of literals of the forms $(s \in t)$, $(s \notin t)$ and of bounded quantifiers of the forms $(\exists x_n \in t)$, $(\forall x_n \in t)$. Equality can be eliminated in terms of membership

$$
\begin{array}{l}
\textsc{Theory}\ \mathsf{evalTerm}\big(\mathsf{th}(\mathsf{N},\mathsf{V}),\mathsf{cst}(\mathsf{S}),\mathsf{Pair}(\mathsf{P}),\mathsf{Appl}(\mathsf{P}),\mathsf{lft}(\mathsf{P}),\mathsf{rgt}(\mathsf{P})\big) \\
\quad \big\langle \forall \mathsf{c},\mathsf{k} \mid \mathsf{cst}(\mathsf{c})=\mathsf{cst}(\mathsf{k})\rightarrow \mathsf{c}=\mathsf{k} \big\rangle \\
\quad \big\langle \forall \mathsf{c},\mathsf{i},\mathsf{p},\mathsf{q},\mathsf{x},\mathsf{y},\mathsf{z} \mid \mathsf{i}\in\mathbb{N} \,\&\, \mathsf{Pair}(\mathsf{p}) \,\&\, \mathsf{Appl}(\mathsf{q}) \rightarrow \{\mathsf{cst}(\mathsf{c}),\mathsf{i},\mathsf{p},\mathsf{q}\} \neq \{\mathsf{x},\mathsf{y},\mathsf{z}\} \big\rangle \\
\quad \big\langle \forall \mathsf{x},\mathsf{y} \mid \emptyset\notin\{\mathsf{x},\mathsf{y}\} \rightarrow \big\langle \exists \mathsf{p},\forall \mathsf{s} \mid \mathsf{Pair}(\mathsf{s}) \,\&\, \mathsf{lft}(\mathsf{s})=\mathsf{x} \,\&\, \mathsf{rgt}(\mathsf{s})=\mathsf{y} \leftrightarrow \mathsf{s}=\mathsf{p} \big\rangle \big\rangle \\
\quad \big\langle \forall \mathsf{x},\mathsf{y} \mid \emptyset\notin\{\mathsf{x},\mathsf{y}\} \rightarrow \big\langle \exists \mathsf{q},\forall \mathsf{s} \mid \mathsf{Appl}(\mathsf{s}) \,\&\, \mathsf{lft}(\mathsf{s})=\mathsf{x} \,\&\, \mathsf{rgt}(\mathsf{s})=\mathsf{y} \leftrightarrow \mathsf{s}=\mathsf{q} \big\rangle \big\rangle \\
\quad \big\langle \forall \mathsf{s} \mid \mathsf{Pair}(\mathsf{s}) \vee \mathsf{Appl}(\mathsf{s}) \rightarrow \emptyset\notin\{\mathsf{lft}(\mathsf{s}),\mathsf{rgt}(\mathsf{s})\} \big\rangle \\
\quad \big\langle \forall \mathsf{t} \mid \{\mathsf{lft}(\mathsf{t}),\mathsf{rgt}(\mathsf{t})\} \subseteq \mathsf{t}\cup\{\emptyset\} \big\rangle \\
\Rightarrow (\mathsf{val}_{\Theta}) \\
\quad \big\langle \forall \mathsf{c},\mathsf{v} \mid \mathsf{val}_{\Theta}\big(\mathsf{cst}(\mathsf{c}),\mathsf{v}\big)=\mathsf{c} \big\rangle \\
\quad \big\langle \forall \mathsf{p},\mathsf{v} \mid \mathsf{Pair}(\mathsf{p}) \rightarrow \mathsf{val}_{\Theta}(\mathsf{p},\mathsf{v}) = \langle \mathsf{val}_{\Theta}\big(\mathsf{lft}(\mathsf{p}),\mathsf{v}\big), \mathsf{val}_{\Theta}\big(\mathsf{rgt}(\mathsf{p}),\mathsf{v}\big)\rangle \big\rangle \\
\quad \big\langle \forall \mathsf{q},\mathsf{v} \mid \mathsf{Appl}(\mathsf{q}) \rightarrow \mathsf{val}_{\Theta}(\mathsf{q},\mathsf{v}) = \mathsf{val}_{\Theta}\big(\mathsf{lft}(\mathsf{q}),\mathsf{v}\big)\restriction \mathsf{val}_{\Theta}\big(\mathsf{rgt}(\mathsf{q}),\mathsf{v}\big) \big\rangle \\
\quad \big\langle \forall \mathsf{n},\mathsf{v} \mid \mathsf{n}\in\mathbb{N}\setminus\{\emptyset\} \rightarrow \mathsf{val}_{\Theta}(\mathsf{n},\mathsf{v}) = \mathsf{th}(\mathsf{n},\mathsf{v}) \big\rangle \\
\textsc{End}\ \mathsf{evalTerm}
\end{array}
$$

Fig. 10.10 A THEORY about the evaluation of terms

This THEORY receives, along with a quintuple of the said kind, a function $\mathsf{th}(N, V)$ supplying the value of the N-th variable in a set-valued assignment V; it manufactures and produces in output the evaluating function val_{Θ}. In order to represent a set-valued assignment it suffices to use a finite-length list which must, in its turn, be modeled somehow: in a manner—we propose—complying with the THEORY interface shown in Fig. 10.11.

The property Lst produced by the THEORY list is meant to indicate which sets represent lists; the dyadic function th associates with any such set ℓ the number of components of the list and the sets occupying those components.[16] Specifically, supposing that $\mathsf{Lst}(\ell)$ holds, $\mathsf{th}(0, \ell)$ will exceed by one the overall number of components of ℓ, and $\mathsf{th}(n, \ell)$ will provide the n-th component of ℓ when $0 < n < \mathsf{th}(0, \ell)$. It should be clear from this explanation that the three claims issued by list state that:

[16] One way of implementing lists is discussed in [30, pp. 127–128].

$$
\begin{array}{l}
\textsc{Theory}\ \mathsf{list}(\,)\\
\Rightarrow (\,\mathsf{Lst}_\Theta\,,\,\mathsf{th}_\Theta\,)\\
\quad \Big\langle \forall \ell \mid \mathsf{Lst}_\Theta(\ell) \rightarrow \mathsf{th}_\Theta(\emptyset,\ell) \in \mathbb{N} \setminus \{\emptyset\} \Big\rangle\\
\quad \Big\langle \forall \ell, \mathsf{m} \mid \mathsf{Lst}_\Theta(\ell)\,\&\,\mathsf{Lst}_\Theta(\mathsf{m})\,\&\,\Big\langle \forall \mathsf{n} \in \mathsf{th}_\Theta(\emptyset,\ell) \mid \mathsf{th}_\Theta(\mathsf{n},\ell) = \mathsf{th}_\Theta(\mathsf{n},\mathsf{m})\Big\rangle \rightarrow \ell = \mathsf{m}\Big\rangle\\
\quad \Big\langle \forall \mathsf{m}, \mathsf{h}, \mathsf{x} \mid \mathsf{h} \in \mathbb{N} \rightarrow \Big\langle \exists \ell \mid \mathsf{Lst}_\Theta(\ell)\,\&\,\mathsf{th}_\Theta(\emptyset,\ell) = \mathsf{h}^+\,\&\\
\qquad\qquad\qquad\qquad \Big\langle \forall \mathsf{n} \in \mathsf{h} \setminus \{\emptyset\} \mid \mathsf{th}_\Theta(\mathsf{n},\ell) = \mathsf{th}_\Theta(\mathsf{n},\mathsf{m})\Big\rangle\,\&\\
\qquad\qquad\qquad\qquad \big(\mathsf{h} = \emptyset \vee \mathsf{th}_\Theta(\mathsf{h},\ell) = \mathsf{x}\big)\Big\rangle\Big\rangle\\
\textsc{End}\ \mathsf{list}
\end{array}
$$

Fig. 10.11 A THEORY of lists

1. the length of every list is a finite ordinal;
2. the equality criterion for lists ℓ, m is that ℓ and m have the same length h and the same n-th component for $n = 1, \ldots, h$;
3. from every triple m, h, x consisting of a list m, a natural number h, and a set x, one can obtain a list ℓ of length h whose last component—if any—is x and whose n-th component is $\mathsf{th}(n, m)$ for $n = 1, \ldots, h-1$; viz.:

$$\ell = \begin{cases} \langle\,\rangle & \text{if } h = 0\,,\\ \langle \mathsf{th}(1,m), \ldots, \mathsf{th}(h-1,m), x\rangle & \text{otherwise.}\end{cases}$$

For a sparing encoding of formulae, we can think of equality as a derived construct; a logical equivalence by which it can be eliminated is in fact $(s = t) \leftrightarrow (\exists \boldsymbol{x}_n \in \langle s, s\rangle)(t \in \boldsymbol{x}_n)$, where $\boldsymbol{x}_n$ does not occur in s or in t. It is also advisable to treat conjunction and disjunction as polyadic connectives, so that the only formulae which need to be encoded directly are the ones of the forms $(s \in t)$, $(s \notin t)$, $(\exists \boldsymbol{x}_n \in t)\left(\bigwedge_{i=0}^{h} \varphi_i\right)$, and $(\forall \boldsymbol{x}_n \in t)\left(\bigvee_{j=0}^{k} \psi_j\right)$, where each φ_i and each ψ_j has in its turn one of these forms. An expedient way of representing a multiple conjunction or disjunction, that owes much to Martin Davis for its dissemination in the early 1960s, is as the sets of conjuncts or disjuncts, respectively;[17] we will rely on this representation for completing our endeavor.

10.10 Related Work

> Often [· · ·] the nonstandard definition of a concept is simpler than the standard definition (both intuitively simpler and simpler in a technical sense, such as quantifiers over lower types or fewer alternations of quantifiers). As a result, nonstandard analysis sometimes makes it easier to find proofs. [4, p. 37]

[17] This way of representing formulae in conjunctive normal form is widely used today. In recent years [32] resorted to it, to give a Ref-based correctness proof for the DPLL satisfiability algorithm.

In what follows, we rely on [6] as an up-to-date comparative survey on systems which offer automated proof abilities related to real analysis. Some of the formalizations supported by such systems characterize real numbers axiomatically, as a given set with specific operations and properties; others construct real numbers either from rational Cauchy sequences or as Dedekind cuts. Nonstandard analysis is available in ACL2(r) and in Isabelle/HOL (see [26, 27], respectively): both achievements are reminiscent of [2, 3].

The semi-automated theorem prover ACL2, which ACL2(r) potentiates, offers limited support to quantifier handling (cf. [27, pp. 323–324]); in order to circumvent that difficulty, ACL2(r) focuses on the extension ${}^*\mathbb{R}$ of the reals. With hyperreal numbers, in fact, the quantifier alternation $\forall\, \varepsilon > 0\ \exists\, \delta > 0 \ldots$ which affects the usual formulas about limits becomes unnecessary, hence the proofs benefit from a higher degree of automation. The formalism of ACL2(r) is based on an axiomatization of ${}^*\mathbb{R}$ as an autonomous domain.

The Isabelle/HOL-mechanization of real analysis, on the other hand, introduces the standard, along with the nonstandard, definition of each concept; thereby, 'users will have the freedom either to stick with classical (standard) techniques, use nonstandard ones, or a combination of both' [26, p. 161]. 'Our first task', the author notes, 'each time we introduce a new concept from analysis, is to prove that the two definitions are equivalent' [26, p. 150]. Thus, albeit implicitly, the *transfer principle* plays a central role. It is 'neither an axiom nor a theorem, but a meta-theorem, since it applies to theorem statements' and, as such, 'it is not directly proved in Isabelle/HOL' [6]; nevertheless, since this principle informs the general pattern followed by all the equivalence proofs, the ultrapower construction of the hyperreals presupposed that a proof of Zorn's lemma and a theory of filters and ultrafilters were developed for Isabelle/HOL (cf. [26, p. 145]).

As an eventual reward of the exploration discussed in this paper, we hope to get Ref-based, nonstandard proofs of theorems of real analysis and to check by means of Ref many of the results presented in Martin Davis's chapter on hyperreal numbers [14, Sects. 2.3–2.8]. However, a formal remake of real analysis along unconventional lines is only an incidental issue here. As discussed at the beginning, we rather feel confronted with a proof-engineering issue—akin to metamathematical extensibility—which our proof assistant could tackle well because a proof of the relevant meta-theorem can be set up with relative ease in a full-fledged set theory.

After all, the guidelines for a Ref-based development of analysis which J. T. Schwartz sketched in [35, Chap. 5] stick to the tradition; the use of nonstandard methods can lead to much simpler and more elegant proofs than the classical ones, but one can contend that it calls for an extra amount of work spent on preliminary constructions, which may be out of scale with a proof of Rolle's theorem (to cite a result of analysis proper).

For a large-scale endeavor, this additional work is justified by considerations such as the following:

> Not only does nonstandard analysis provide a rigorous treatment of infinitesimals in the area of mathematics where they were originally used, it also gives elegant approaches to some ideas that developed later.
>
> [4, p. 37]

10.11 Concluding Remarks

> The well known theorem of Gödel shows that every system of logic is in a certain sense incomplete, but at the same time it indicates means whereby from a given system L of logic a more complete system L' may be obtained. By repeating the process we get a sequence L, $L_1 = L'$, $L_2 = L'_1$, $L_3 = L'_2$, ... of logics each more complete than the preceding.
>
> (A. M. Turing, 1938)

The authors have at this point prepared the ground for verifying, with a proof-checker based on set theory, the propositions in the first chapter of [14].[18] A variant of the Zermelo-Fraenkel set theory, postulating global choice, regularity and infinity,[19] underlies the logical armory of the proof-checker, Ref, on which our experimental activity relies. The formally checked proofs regard, for the time being, only certain parts of our planned work: in particular, we proved the conclusions of the THEORYs about universes, superstructures, and 'urification' shown in Figs. 10.5 and 10.6, as well as the unique readability of the sets that encode terms inside the THEORY termEncoding (see Figs. 10.8 and 10.9); the proof of the ultrafilter theorem was available from the outset,[20] along with many minor but useful facts about finiteness, rank, ordinals, the set constructs $\mathscr{P}$, $\bigcup$, etc.

In the phase on which we have reported, anyway, our work has been mainly architectural: given the availability of a second-order construct, 'THEORY', supporting modularization and proof reuse in Ref, we deem it wise to invest in designing the THEORY interfaces before formalizing proofs meticulously.

---

[18] A website reporting on our experiment is at
http://www2.units.it/eomodeo/InitialSetupForNonStandardAnalysis.html,
http://aetnanova.units.it/scenarios/InitialSetupForNonStandardAnalysis/.

[19] In Ref the well-foundendess of membership and statements of the axiom of choice easily result from the availability of the construct ***arb*** discussed in Sect. 10.6, thanks to the interplay of ***arb*** with abstract set formers; infinity is embodied by Ref's built-in constant $\boldsymbol{\sigma}_\infty$.

[20] For a Ref-based proof of Zorn's lemma (whence the ultrafilter theorem follows easily), see [35, pp. 373–405]. This lemma was used in Ref's proof of the maximal ideal theorem for Boolean algebras as presented in [9].

We are confident that we can finish the envisaged proof-development tasks without getting entangled in unforeseen difficulties. Then, as said in the introduction, we must adopt schemes of notation and extended rules of inference that conveniently assist Ref's users in exploitations of the nonstandard methods.

Even after those enhancements, Ref's theory will be a conservative extension of the specific set theory available in Ref's initial endowment. A more challenging and intriguing view on the extensibility of proof-checkers should cope with the progressive extension of theories, in a frame of mind close to some of Alan Turing's early investigations (see [37]).

Acknowledgements Discussions with Francesco Di Cosmo helped in polishing this paper. The first author acknowledges partial support from the Polish National Science Centre research project DEC-2011/02/A/HS1/00395; and the second author from the project FRA-UniTS (2014) "*Learning specifications and robustness in signal analysis*.

References

1. Anastasio, S. (Coordinating Editor) (2015). In memory of Jacob Schwartz. *Notices of the AMS*, 473–490.
2. Ballantyne, A. M. (1991). The Metatheorist: Automatic proofs of theorems in analysis using non-standard techniques, Part II. In R. S. Boyer (Ed.), *Automated reasoning: Essays in Honor of Woody Bledsoe* (pp. 61–75). Dordrecht, The Netherlands: Kluwer Academic.
3. Ballantyne, A. M., & Bledsoe, W. W. (1977). Automatic proofs of theorems in analysis using nonstandard techniques. *Journal of the ACM*, *24*(3), 353–374.
4. Blass, A. (1978). Book reviews of *Applied nonstandard analysis*, by Martin Davis, *Introduction to the theory of infinitesimals*, by K. D. Stroyan and W. A. J. Luxemburg, and *Foundations of infinitesimal calculus*, by H. Jerome Keisler. *Bull. Amer. Math. Soc.*, 84(1):34–41, 1978.
5. Bledsoe, W. W. (1977). Non-resolution theorem proving. *Artificial Intelligence*, *9*(1), 1–35.
6. Boldo, S., Lelay, C., & Melquiond, G. (2015). Formalization of real analysis: A survey of proof assistants and libraries. *Mathematical Structures in Computer Science*, 38 pp.
7. Burstall, R., & Goguen, J. (1977). Putting theories together to make specifications. In R. Reddy (Ed.), *Proceedings of the 5th International Joint Conference on Artificial Intelligence* (pp. 1045–1058). Cambridge, MA.
8. Cantone, D., Omodeo, E. G., & Policriti, A. (2001). *Set Theory for Computing. From Decision Procedures to Declarative Programming with Sets*. Monographs in Computer Science. Springer.
9. Ceterchi, R., Omodeo, E. G., & Tomescu, A. I. (2014). The representation of Boolean algebras in the spotlight of a proof checker. In L. Giordano, V. Gliozzi, & G. L. Pozzato, (Eds.), *CILC 2014: Italian Conference on Computational Logic*, volume 1195 http://ceur-ws.org/Vol-1195/, ISSN 1613-0073, pp. 287–301. CEUR Workshop Proceedings, July 2014.
10. Chinlund, T. J., Davis, M., Hinman, P. G., & McIlroy, M. D. (1964). Theorem-proving by matching. Technical report, Bell Telephone Laboratories, Incorporated, Murray Hill, New Jersey.
11. Cohen, P. J. (1966). *Set Theory and the Continuum Hypothesis*. Mathematics Lecture Note Series. Reading, Massachusetts: W. A. Benjamin, Inc.
12. Davis, M. (1960). A program for Presburger's algorithm. *Summaries of talks presented at the Summer Institute of Symbolic Logic in 1957 at Cornell University* (vol. 2, pp. 215–223). Princeton, NJ. Communications Research Division, Institute for Defense Analyses. Reprinted as "A computer program for Presburger's algorithm" in [36, pp. 41–48].

13. Davis, M. (1963). Eliminating the irrelevant from mechanical proofs. *Proceedings of Symposia in Applied Mathematics* (vol. 15, pp. 15–30). Providence, RI: AMS. Reprinted in [36, pp. 315–330]; Russian transl. in *Kiberneticheskiy sbornik. Novaya seriya*, 7, 1970, pp. 160–179.
14. Davis, M. (1977). *Applied nonstandard analysis*. Wiley. Reprinted with corrections Dover, 2005. Russian translation, Izdatel'stvo Mir, Moscow 1980. Japanese translation 1977.
15. Davis, M. (2001). The early history of automated deduction. In J. A. Robinson & A Voronkov, (Eds.), *Handbook of Automated Reasoning* (pp. 3–15). Elsevier and MIT Press.
16. Davis, M. (2013). *Jack Schwartz meets Karl Mar*x. In *[22, pp. 23–37]*.
17. Davis, M., & Fechter, R. (1991). A free variable version of the first-order predicate calculus. *Journal of Logic and Computation*, *1*(4), 431–451.
18. Davis, M., & Hersh, R. (1972). Nonstandard analysis. *Scientific American*, *226*, 78–86.
19. Davis, M., & Putnam, H. (1958). *Feasible computational methods in the propositional calculus*. Technical report, Rensselaer Polytechnic Institute, Research Division, Troy, New York.
20. Davis, M., & Putnam, H. (1960). A computing procedure for quantification theory. *Journal of the ACM*, *7*(3):201–215. Reprinted in [36, pp. 125–139].
21. Davis, M., & Schonberg, E. (2011). Jacob Theodore Schwartz 1930–2009. *Biographical Memoirs of the National Academy of Sciences*,19 pp.
22. Davis, M., & Schonberg, E. (Eds.). (2013). *From Linear Operators to Computational Biology: Essays in Memory of Jacob T. Schwartz*. Springer.
23. Davis, M. & Schwartz, J. T. (1977). Correct-program technology/Extensibility of verifiers—Two papers on Program Verification with Appendix of Edith Deak. Technical Report No. NSO-12, Courant Institute of Mathematical Sciences, New York University.
24. Davis, M. & Schwartz, J. T. (1979). Metamathematical extensibility for theorem verifiers and proof-checkers. *Computers and Mathematics with Applications*, *5*, 217–230. Also in [25, pp. 120–146].
25. Davis, M., Logemann, G., & Loveland, D. W. (1962). A machine program for theorem-proving. Communications of the Association for Computing. *Machinery*, *5*(7), 394–397.
26. Fleuriot, J. D. (2000). On the mechanization of real analysis in Isabelle/HOL. In M. Aagaard & J. Harrison, (Eds.), *Theorem Proving in Higher Order Logics, 13th International Conference, TPHOLs 2000, Portland, Oregon, USA, 14–18 August 2000, Proceedings*, volume 1869 of *Lecture Notes in Computer Science* (pp. 145–161). Springer.
27. Gamboa, R., & Kaufmann, M. (2001). Nonstandard analysis in ACL2. *Journal of Automated Reasoning*, *27*(4), 323–351.
28. Keisler, H. J. (1976). *Foundations of infinitesimal calcuus*. Boston, MA: Prindle, Weber & Schmidt, Inc.
29. Omodeo, E. G. (1982). The Linked Conjunct method for automatic deduction and related search techniques. *Computers and Mathematics with Applications*, *8*, 185–203.
30. Omodeo, E. G. (2012). The Ref proof-checker and its "common shared scenario". In M. Davis & E. Schonberg, (Eds.), *From Linear Operators to Computational Biology: Essays in Memory of Jacob T. Schwartz* (pp. 121–131). Springer.
31. Omodeo, E. G., & Schwartz, J. T. (2002). A 'Theory' mechanism for a proof-verifier based on first-order set theory. In A. Kakas & F. Sadri, (Eds.), *Computational logic: Logic programming and beyond—Essays in honour of Bob Kowalski, part II* (vol. 2408, pp. 214–230). Springer.
32. Omodeo, E. G., & Tomescu, A. I. (2008). Using ÆtnaNova to formally prove that the Davis-Putnam satisfiability test is correct. *Le Matematiche*, *63*(1), 85–105.
33. Policriti, A. (1988). Decision procedures for elementary sublanguages of set theory. IX. Unsolvability of the decision problem for a restricted class of the Δ_0-formulas in set theory. *Communications on Pure and Applied Mathematics41*(2), 221–251.
34. Robinson, J. A. (1967). Review: Martin Davis, Eliminating the irrelevant from mechanical proofs. *Journal of Symbolic Logic*, *32*(1), 118–119.
35. Schwartz, J. T., Cantone, D., & Omodeo, E. G. (2011). *Computational logic and set theory—Applying formalized logic to analysis*. Springer.
36. Siekmann, J., & Wrightson, G. (Eds.). (1983). *Automation of reasoning 1: Classical papers on computational logic 1957–1966*. Berlin, Heidelberg: Springer.

37. Turing, A. M. (1939). Systems of logic based on ordinals. *Proceedings of the London Mathematical Society*, *2*(45), 161–228.
38. Weyhrauch, R. W. (1977). A users manual for FOL. Technical Report MEMO AIM-235.1, Stanford University, Stanford, CA, USA.
39. Yarmush, D. L. (1976). The Linked Conjunct and other algorithms for mechanical theorem-proving. Technical Report IMM 412, Courant Institute of Mathematical Sciences, New York University.