Hierarchies ontological and ideological

Journal Article

Citation:

Liinebo, O.; Rayo, A. (2011) Hierarchies ontological and ideological – *Mind* (In press)

Hierarchies ontological and ideological

Øystein Linnebo and Agustín Rayo

January 2, 2011

1 Introduction

This remarkable passage was written by Kurt Gödel in 1933:

> only one solution [to the paradoxes] has been found, although more then 30 years have elapsed since the discovery of the paradoxes. This solution consists in the theory of types. (I mean the simple theory of types [...].)
>
> It may seem as if another solution were afforded by the system of axioms for the theory of aggregates, as presented by Zermelo, Fraenkel and von Neumann; but it turns out that this system is nothing else but a natural generalization of the theory of types, or rather, it is what becomes of the theory of types if certain superfluous restrictions are removed. ((Gödel, 1933), pp. 45-6)

The passage is remarkable for several reasons. One is of a technical nature. The theory of types is *stratified* in a way that standard set theory is not. Whereas the language of type-theory contains variables of many different sorts and imposes significant constraints on the ways in which these variables can be combined, the language of set-theory contains only one sort of variable, any instance of which is allowed to fill the argument-places of any predicate in the language. So one might have thought that much more than the removal of 'superfluous restrictions' would be required to get from the theory of simple types to a system like Zermelo-Fraenkel set theory.

Other reasons are more philosophical. Type theory is a generalization of higher-order logic. And higher-order logic is widely regarded as profoundly different from set theory in motivation, interpretation, and philosophical status. For instance, a great deal of attention has been given to the view that second-order languages—languages consisting of the first two levels of the theory of simple types—are 'ontologically innocent' in a way that the language of set-theory is not.[1] According to this line of thought, all it takes for a second-order sentence such as '$\exists X(X(\text{Susan}))$' to be true is for Susan to exist. In contrast, *sets* must exist in order for the corresponding set-theoretic sentence, '$\exists x(\text{Susan} \in x)$', to be true. So any formal similarities between set-theory and type-theory would belie a crucial difference in their *subject-matter*. The language of set-theory is about a particular kind of object—sets. The language of type-theory is not: it increases our expressive resources without imposing demands on our ontology. In short, where set theory describes an 'ontological hierarchy' of independently existing sets, higher-order logic—and, more generally, type theory—makes available an 'ideological hierarchy' of stronger and stronger expressive resources.[2]

The widely shared assumption that these two hierarchies are profoundly different plays a key role in many popular applications of higher-order logic. Ontological applications have already been mentioned. Other examples include uses of higher-order logic in semantics (especially to defend the possibility of absolutely general quantification)[3] and philosophy of mathematics (where it is used to defend the categoricity of important mathematical structures).[4] Gödel's claim that set theory is "nothing else but a a natural generalization of the theory of types" would appear to be a direct challenge to the assumption that the two hierarchies are importantly different, and therefore to the many applications of higher-order logic that depend on this assumption.

The aim of this paper is to develop a clearer understanding of Gödel's remark and the surrounding philosophical terrain. After reconstructing and defending the remark, we will discuss the claim that there is no significant philosophical difference between the 'ontological hierarchy' of sets and the 'ideological hierarchy' of type theory, and argue that there is much to be said on behalf of this claim.[5] We will also discuss some technical issues concerning

[1] See for instance (Boolos, 1984), (Rayo and Yablo, 2001), and (Wright, 2007).

[2] A related point is that whereas the set-theoretic membership relation is typically regarded as a non-logical predicate, the corresponding type-theoretic relation is typically regarded as a logical predicate.

[3] See (Boolos, 1985) and (Rayo and Uzquiano, 1999), and (Williamson, 2003).

[4] See for instance (Shapiro, 2000).

[5] This will make good on suggestions made in (Linnebo, 2003), pp. 87-88 and (Rayo, 2006), p. 248.

infinitary type theories, and generalize the programme of developing the semantics for higher-order languages in other higher-order languages.[6]

2 Gödel's 'superfluous restrictions'

As we have seen, Gödel saw Zermelo-Fraenkel set theory as resulting from the removal of certain 'superfluous restrictions' from the theory of simple types. What, exactly, are these restrictions? Gödel lists three:

(i) "In Russell's theory the process of going over to the next higher type—for instance, from classes of individuals to classes of classes of individuals—can be repeated only a finite number of times; i.e. to each class occurring in the system of *Principia Mathematica*, there is assigned a finite number n, indicating in how many steps the class under consideration can be reached, starting from the level of individuals. This number n can be arbitrarily large, but it must be finite. Now there is no reason whatever to stop the process of formation of types at this stage. You can, e.g., form the class of *all* classes of finite type, which, of course, is not of finite type, but may be called of type ω. It is clear how this process can be continued indefinitely." ((Gödel, 1933), pp. 46-47)

(ii) "Only the so-called pure types have been admitted [...], i.e. no class can be formed which contains classes of different type among its elements." (*ibid.*, p. 46)

(iii) "Propositions of the form $a \in b$ are regarded as meaningless (i.e. neither true nor false) if a and b are not of the appropriate types—if, for instance, a is of higher type than b. This complication can be removed simply by stating that $a \in b$ is to be false if a and b are not of the appropriate types." (*ibid.*, p. 46)

This calls for explanation. Let us begin with the notion of type theory. Ordinary first-order logic contains quantifiers that range over individuals. But logicians and philosophers have also been interested in systems which extend the language of first-order logic by adding new kinds of variables and quantifiers. The most familiar such system is second-order logic,

[6] See the references in foonote 3.

which adds *second-order variables* (i.e. variables taking the position of first-order predicates) and *second-order quantifiers* (i.e. quantifiers binding such variables).[7] But instead of adding just one set of variables and quantifiers, one could add a whole hierarchy.

The best known language of this sort is the language of the *simple theory of types*, where each variable is indexed with a natural number known as its *type* ('x^1' and 'y^{17}', for instance, are variables of type 1 and 17, respectively). On the classical interpretation of the language, variables of type 0 range over individuals; variables of type 1 range over classes of individuals; variables of type 2 over classes of classes of individuals; and so on. (This is how Gödel expresses himself in the passage that we quoted at the beginning of this article.) But today the most popular interpretations are based on a hierarchy of concepts, or on a hierarchy of pluralities, super-pluralities and beyond. There are some important differences between the axiomatic theories that are appropriate for these two interpretations; for instance, there are empty concepts but no empty pluralities. It will be useful to work with the conceptual interpretation in what follows, but we offer a discussion of the plural interpretation in Appendix A. We will let the atomic formulas of type theory be of the form '$s(t)$', rather than use the notation '$t \in s$', which we will reserve for set theory. For reasons of simplicity, we will not consider types for functions or polyadic relations. Such types would complicate matters greatly but would not substantially change the philosophical claims that follow (with one possible exception explicitly noted).

Although the language of the simple theory of types has only variables of finite types, more radical extensions of classical first-order languages are possible. In particular, we could allow variables whose type-indices are not just natural numbers but arbitrary ordinals. (Of course, this requires working against the background of ordinary set theory.) We say that a language is of *order* α when all of its variables have type-indices below α. This is a natural generalization of the familiar notion of a nth-order language, for finite n. For instance, a second-order language has variables of type 0, which range over individuals, and variables with type 1, which range over first-order concepts (or pluralities, or classes of individuals).

[7] Although second-order languages are extensions of first-order languages, the historical development was actually the reverse: the former languages were formulated and investigated before the latter. See (Shapiro, 2000), Ch. 7.

What is often known as a 'second-order variable' is thus regarded as a variable of type 1, and likewise for variables of greater finite order. The language of the simple theory of types is a language of order ω, as it has variables of all types below ω.

Type-theoretic languages may also contain constants. Specifically, a language of order α may contain constants of type less than or equal to α. This too is a natural generalization of familiar notions. For instance, a second-order language may contain constants of type 2, which are applicable to 'second-order variables' of type 1.

On the intended interpretation, the values of a variable of limit type will consist of the 'union' of the values of variables of all preceding levels, and the values of a variable of successor type will consist of the 'union' of the values of variables of the proceeding level plus all 'collections' of values of variables of that level.

Next we explain the notion of 'cumulativity'. A typed language is said to be *cumulative* when for any two terms s and t the string $s(t)$ is regarded as a well-formed formula just in case the type of s is strictly greater than the type of t. A typed language is said to be *non-cumulative* when the stricter requirement is imposed that the type of s must be precisely one greater than the type of t. (The languages of ordinary higher-order logic and simple type theory are non-cumulative languages in this sense.) A cumulative typed language thus counts more predications as well-formed than a non-cumulative typed language. But even a cumulative typed language deems many predications ill-formed: those where the type of the subject term is equal to or greater than that of the predicate term.

Removal of Gödel's three 'superfluous restrictions' from simple type theory amounts to the implementation of the following three changes:

(i) Allow infinite types.

(ii) Allow the type structure to be cumulative.

(iii) Allow a type-unrestricted notion of predication.

In the next three sections we will consider some arguments for making these changes. Although we believe our arguments to be consonant with Gödel's position, they go well beyond his brief discussion of the issue. Our goal is not an exegetical one. We wish to develop an interesting argument which is broadly in the spirit of Gödel's proposal.

3 Step one: Admitting infinite types

Why should we admit infinite types?

Gödel's initial answer seems to be simply: 'Because we can.' It makes perfect mathematical sense to admit infinite types. So what is to prevent us from doing so?[8]

We are sympathetic towards this simple answer. But we would like to explore an additional argument for thinking that one ought to admit infinite types. The argument relies on a controversial assumption:

> Absolute Generality
>
> One's first-order quantifiers can meaningfully be taken to range over absolutely all objects.

The reasons Absolute Generality is controversial are tied up in interesting ways with the question of whether there is a significant philosophical difference between type-theory and set-theory. Philosophers tend to fall into two main groups. Members of the first group accept Absolute Generality, and use it to argue that there is a significant philosophical divide between the 'ideological hierarchy' of higher-order languages and the 'ontological hierarchy' of set-theory. (They argue, for example, that whereas '$\exists X \forall y (Xy)$' is true when the range of the type-0 variables consists of absolutely every individual, the corresponding set-theoretic claim, '$\exists x \forall y (y \in x)$', must be false, on pain of paradox.) By contrast, members of the second group are skeptical of Absolute Generality but tend to believe that there is no significant philosophical divide between the two hierarchies—the only real difference between type-theory and set theory has to do with differences in syntactic restrictions on the corresponding languages.

[8]Other logicians too were motivated by this line of thought. See for instance, (Hilbert, 1926), p. 184 (p. 387 of translation), (Carnap, 1934), p. 186, and (Tarski, 1935), Section 7. Inspired by Hilbert, Gödel suggests introducing infinite types already in his famous incompleteness paper; see (Gödel, 1931), fn. 48a. Moreover, according to (Gödel, 1944), pp. 464, the constructible hierarchy L can be regarded as an extension of the orders of Russell's ramified theory of type into the transfinite.

If you are a member of the second group, you are unlikely to be moved by the main argument in this section. But you will also be less in need of an argument. For you will think there is no *philosophical* obstacle to extending type-theory into the transfinite: transfinite-type theory is a variant of set-theory with a particularly restrictive notation. So the only potential obstacle to transfinite type-theory arises from questions about its technical viability.

If you are a member of the first group, you will think that extending type-theory into the transfinite is philosophically substantial view. For you will think that each additional stage in the type-theoretic hierarchy presupposes an essentially new kind of expressive resource. But you will also believe in Absolute Generality, so you will take the argument in this section to have force.[9]

Now for the argument. It will be useful to start by presenting a rough version of the main idea—a more precise version of the argument will come later. Let a *generalized semantic theory* for a given language be a theory of all possible interpretations the language might take. If one were to assume that every possible interpretation of the language in question is captured by some (set-thereotic) *model* of the language, then one would see no difference between a generalized semantic theory for the language and a model theory for the language. But if you accept Absolute Generality, you will think that a language can admit of interpretations that don't correspond to any model in the set-theoretic sense. The standard interpretation of a language with quantifiers ranging over everything there is, for example, requires a domain consisting of everything there is. But there is no universal set. So there is no set-theoretic model capturing the standard interpretation of this language.

More generally, by assuming Absolute Generality one can prove that it is impossible to state a generalized semantic theory for an nth-order language in another nth-order language. It is, however, possible to state a generalized semantic theory for an nth-order language in an $(n+1)$th order language.[10] So—unless one is prepared to countenance a language for which generalized semantic theorizing is impossible—one has a motivation for

[9]The argument also presupposes an unrestricted comprehension scheme for each type; see Section 6.4 and Appendix B for details. Although this assumption is widely regarded as uncontroversial, it has been challenged in (Linnebo,). The ensuing dialectical situation is much like the one just discussed. Since the challenge is based on assumptions that reduce the divide between the two hierarchies, anyone sympathetic to the challenge will have independent reasons for believing that ascent in the ideological hierarchy is possible.

[10]For proofs and details, see (Rayo, 2006).

using languages of order $n + 1$ in one's theorizing whenever one is prepared to use languages of order n. Since we are certainly prepared to use *first*-order languages in our theorizing, it follows that we have a motivation for using languages of every finite type.

But if one is indeed prepared to theorize using languages of order n for each natural number n, there is no obvious reason for questioning the legitimacy of ω-order languages (i.e. a languages with variables of every finite type). For such languages would be made up entirely from vocabulary that had been previously deemed legitimate. And once one has an ω-level language there is a reason to climb further. For one can prove that it is impossible to state a generalized semantic theory for an ω-order language in another ω-order language. It is, however, possible to state a generalized semantic theory for an ω-order language in a language of even higher order. So—unless one is prepared to countenance a language for which generalized semantic theorizing is impossible—one has a motivation to ascend even further.

The argument can be generalized—provided we are willing to work against the background of ordinary ZF set theory or some other theory of ordinals. For we can then use the following two (plausible but non-trivial) principles to motivate the legitimacy of languages of type α for α an arbitrary ordinal:

- The Principle of Semantic Optimism

  Given an arbitrary language, it should be possible to articulate a generalized semantic theory for that language.

- The Principle of Union

  For λ a limit ordinal, suppose that one is prepared to countenance languages of order β for every $\beta < \lambda$. Then one should also countenance languages of order λ (i.e. languages containing variables of type β, for every $\beta < \lambda$), on the grounds that they would be made up entirely of vocabulary that had been previously deemed legitimate.

The argument runs as follows. Assuming Absolute Generality, it is impossible to state a generalized semantic theory for a β-order language in another β-order language. It is, however, possible to do so in a $(\beta + 1)$-order language (or a $(\beta + 2)$-order language if β is a limit ordinal). So the Principle of Semantic Optimism motivates ascent from β-order languages to languages of higher order, for arbitrary β. Moreover, the Principle of Union motivates ascent to λ-order languages, for λ an arbitrary limit ordinal, whenever ascent to β-order languages, for $\beta < \lambda$, has been motivated. But one should certainly admit the use of first-order languages in one's theorizing. So, by transfinite induction, one should admit use of α-level languages in one's theorizing, for arbitrary α. QED

Of course, one might attempt to challenge the premises on which the argument is based. In particular, we have not provided any systematic defense of the Principle of Semantic Optimism. In the absence of a more systematic defense of the premises, we cannot pretend to have provided an apodictic route from Absolute Generality to the the legitimacy of languages of type α for α an arbitrary ordinal. However, we maintain that the premises are plausible and that our argument thus succeeds in showing that friends of Absolute Generality are naturally led to the conclusion that there are at least as many types as ordinals.

It is also worth noting that the argument makes essential use of an 'external' theory of ordinals, which are used as type superscripts on variables. To emphasize the point, it is useful to consider a skeptic about set theory, who is unwilling to talk about ordinals. Such a skeptic might accept versions of the Principle of Semantic Optimism and the Principle of Union by saying 'for any language, there is a language of higher type', and 'for any determinate collection of languages, one could in principle construct a 'union' language which pools together all their resources'. But when the principles are stated in this way, there is no guarantee that the skeptic is indeed prepared to go 'all the way up', and end up with a hierarchy of languages isomorphic to the hierarchy of ordinals. For any limit ordinal λ, it is consistent with a reading of what the skeptic has said that her principles yield only languages of level β for $\beta < \lambda$.

A final point worth emphasizing is the huge idealization that languages of infinite order involve. Such languages are of course very different from the sorts of languages that humans are actually capable of using.[11] Our project is to use the resources of set theory to investigate

[11] A less extreme idealization would be to equip the semanticist with ordinal notations, which can be used as type indices. (Thanks here to Stewart Shapiro.)

possible expressive resources, regardless of whether these expressive resources can be mastered by human beings.

We present proofs of the main technical results underlying the argument in Appendix B. One result is negative and says that, for any ordinal β, it is impossible to give a generalized semantic theory for a β-order language in another β-order language. Another result is positive and says that, for any successor ordinal β, it is possible to give a generalized semantic theory for β-order languages in $(\beta + 1)$-order languages.

It is interesting to note that Tarski makes a claim that is closely related to our positive result (although concerned with truth *simpliciter* rather than truth in a model, which is our current concern):

> In fact, the setting up of a correct definition of truth for languages of infinite order would in principle be possible provided we had at our disposal in the metalanguage expressions of higher order than all variables of the language investigated. ((Tarski, 1935), p. 272)

A related claim was made by Gödel in 1933:

> For any formal system you can construct a proposition—in fact a proposition in the arithmetic of integers—which is certainly true if the system is free from contradiction but cannot be proved in the given system. Now if the system under consideration (call it S) is based on the theory of types, it turns out that exactly the next higher type not contained in S is necessary to prove this arithmetic proposition, i.e. this proposition becomes a provable theorem if you add to the system S the next higher type and the axioms concerning it." ((Gödel, 1933), p. 48)

4 Step two: 'Going cumulative'

Why 'go cumulative'? That is, why allow an entity of one type to apply to entities of all lower types rather than just entities of the immediately preceding type?

Again, Gödel's initial answer seems to be just 'Because we can'. Since it makes perfect mathematical sense to do so, what is to prevent us?

While we are sympathetic to this simple answer, we now explore a more systematic argument. This argument is based on the observation that it is very difficult not to 'go cumulative' once infinite types are allowed, as urged in Step 1. Assume for instance that we want to accept variables of type ω. To what expressions are these variables to apply? The standard non-cumulative answer is: to terms of the immediately preceding type. But there is no type which immediately precedes ω. The only natural response is to allow variables of type ω to apply to terms of any finite type, in other words, to allow type ω to be cumulative. Analogous considerations support the view that variables of any limit type must be allowed to be cumulative. But once we allow all variables of limit types to be cumulative, why not allow cumulativity for variables of successor type as well? This amounts to 'going completely cumulative'.

Can this argument be resisted? One option is to resist the assumption that there are no infinite descending chains of types. For if there were such a chain $\alpha, \alpha-1, \alpha-2, \ldots$, this would allow the type hierarchy to be strictly non-cumulative, with terms of one type applicable to all and only terms of the immediately preceding type. In fact, type theories with infinite descending chains of types are consistent provided that the corresponding theories without such chains of types are consistent.[12] But even if consistent, systems with such infinitely descending chains of types are problematic, since they can be shown to have no standard models.[13] Moreover, systems with infinitely descending chains of types seem poorly motivated. While a well-ordered theory of types can be motivated by an "iterative conception of types", there seems to be no analogous motivation for a theory of non-well-founded types. These

[12] *Proof sketch.* Assume there was a derivation of an inconsistency in the former system. Since this derivation is finite, it uses only finitely many different types. The derivation can therefore be imitated in the latter system. See (Wang, 1952) for discussion of type theories where the types are all integers rather than just the natural numbers.

[13] For a model to be 'standard', in the relevant sense, is for every collection of entities of a given type to constitute the sole instances of some entity of the next higher type. That no such model exists can be verified by using an observation from (Yablo, 2006). Say that an entity y^γ is *well-founded* iff there is no infinite descending chain $y^\gamma, y^{\gamma-1}, \ldots$ such that $y^{\gamma-k}(y^{\gamma-k-1})$ for every natural number k. Consider the collection $W^{\alpha-k}$ of well-founded entities of level $\alpha - k - 1$. Then we must have $W^{\alpha-k}(W^{\alpha-k-1})$ for each $k \in \omega$. But this yields an infinitely descending chain, which means that the entities $W^{\alpha-k}$ cannot be well-founded after all.

Why doesn't this problem render the relevant type theory inconsistent? It turns out that inconsistency is avoided only because the notion of well-foundedness cannot be expressed in the language, and because the argument is infinitely long and hence cannot be captured by a finite proof. However, when the theory is formulated in a system of form $\mathcal{L}_{\omega_1\omega_1}$, both obstacles are removed and an inconsistency can indeed be derived.

considerations make it reasonable to disallow infinitely descending sequences of types.

In fact, the addition of infinite types all of which are non-cumulative would anyway have had little theoretical value. Much of the point of infinite types is to have variables capable of applying directly to all (or at least most) of the preceding types. For instance, we may be motivated to go to an infinite level in order to develop the semantics of languages of all finite levels. But this requires expressions of an infinite type which is allowed to apply directly to expressions of all finite types.

A second assumption is that, if variables of limit types are allowed to be cumulative, then so must variables of successor types. This assumption too can be resisted. It is certainly *possible* to adopt a (partially cumulative) hierarchy which is cumulative at limit types but not at successor types. But there is no obvious philosophical motivation for doing so. If cumulativity is acceptable at limit types, there can be nothing fundamentally incoherent or impermissible about it. So then why not accept cumulativity also at successor types?[14] But this resistance would be philosophically unmotivated. If cumulativity is allowed at limit types, there can be no principled objection to it. It would thus be philosophically *ad hoc* not to allow successor types to be cumulative as well.

Notice, moreover, that the refusal to allow for cumulativity at successor types has little technical significance. For such cumulativity can be *simulated* by choosing to work only with variables and constants of limit levels. For instance, the effect of the claim $x^4(c^2) \wedge c^2(x^0)$ can be simulated by "relettering" it as $x^{\omega+\omega}(c^\omega) \wedge c^\omega(x^0)$.[15] This observation extends to arbitrary arguments in the fully cumulative language. Provided that the relettering respects the original ordering of the terms and is uniform throughout an entire argument, the relettered argument will be well-formed in the partially cumulative language and capable of doing the expressive and deductive work of the argument in the fully cumulative language.

A final assumption made in our argument for 'going cumulative' is that the types are

[14] In fact, when the type-theoretic hierarchy is understood in terms of plurals, super-plurals and beyond, there appears to be an independent motivation for cumulativity at successor types. Consider, for example, the sentence 'These people, those people, and that man compete' where a predication is made of a 'super-plurality' consisting of two ordinary pluralities and one individual. For discussion of 'super-pluralities' in natural language, see (Linnebo and Nicolas, 2008).

[15] More generally, formulas can be "relettered" in this way in any language of an order whose Cantor normal form ends with $\omega^2 \cdot k$ for natural number k.

linearly ordered. Together with the second assumption that there are no infinitely descending sequences of types, this allows us to identify the types with some initial segment of the ordinals. But why assume that the types are linearly ordered? For instance, why not allow terms for polyadic relations? Given any types $\tau_1, \ldots, \tau_n$, we can for instance admit terms of type $\langle \tau_1, \ldots, \tau_n \rangle$, which take n arguments of types $\tau_1, \ldots, \tau_n$ respectively. This corresponds to a natural partial (but non-linear) order $\leq$ on the types defined by $\tau_i < \langle \tau_1, \ldots, \tau_n \rangle$ for every i, and letting $\leq$ be the reflexive closure of $<$.[16]

It is worth noting that such relations are not available on the plural interpretation of type theory. For even if one admits 'higher' pluralities, the plural interpretation does not provide any primitive polyadic relations. (It is irrelevant to our present concern that polyadic relations can be *coded* by means of suitable 'higher' pluralities.)

On the conceptual interpretation, on the other hand, there is no obstacle to admitting primitive polyadic relations. Notice, however, that not just any kind of polyadic relation will help someone who wishes to allow infinitary types while holding on to non-cumulativity and the well-foundedness of their ordering $<$. It can be proved that for this goal to be attainable, we must allow expressions of infinite adicity.[17] One can for instance avoid cumulativity at level ω by admitting relations of countably infinite adicity with a separate argument place for expresions of each finite type. Although we see no principled objection to type theories which admit such relations, they would represent a major complication of our type theory. It is particularly worrisome that such systems will require not just infinitely long conjunctions and disjunctions, but also infinitely long strings of quantifiers. Such logics are known to be very complicated. Moreover, relying on relations of infinite adicity is contrary to a major concern in the foundations of mathematics, namely to develop simple and natural theories in which much of mathematics can be codified. Although the type theory that we have set out involves a huge idealization from our actual mathematical languages, we maintain that it is very clean and tidy compared with theories that include relations of infinite adicity and

[16] That is, we let $\tau \leq \tau'$ iff $\tau < \tau'$ or $\tau = \tau'$.

[17] *Proof sketch.* Assume there is an infinitary type. Then by the well-foundedness of $<$, there must be an $<$-minimal infinitary type τ. Let e be an expression of type τ. Assume e is of finite adicity, say $\tau = \langle \tau_1, \ldots, \tau_n \rangle$. Then by the minimality assumption, $\tau_1, \ldots, \tau_n$ must be finitary types. But then τ cannot be infinitary. Hence the adicity of τ cannot be finite after all.

infinitely long strings of quantifiers. So we submit that there are strong pragmatic reasons to be interested in systems where the types are linearly ordered.

Given the three assumptions just defended, a mathematical proof of full cumulativity is not far away.[18]

5 Step three: Admitting type-unrestricted predication

Why should we admit a type-unrestricted notion of predication? That is, why should we admit a predication as well-formed only when the type of the predicate is strictly greater than the type of the argument?

According to Gödel, this restriction on predication "can be removed simply by stating that [the predications previously regarded as ill-formed are] to be false if [the expressions] are not of the appropriate types." He argues that doing so does not reintroduce any danger of paradox and thus is mathematically unproblematic.

We are sympathetic towards this simple argument. But it is worth noting that a more systematic treatment of the issue is available: one can show that type-unrestricted predications can be treated as syntactic abbreviations for more complex formulas involving only type-restricted predication.

We begin by defining a type-unrestricted notion of identity. Let $x^\alpha \equiv_\gamma y^\beta$ abbreviate the claim that x^α and y^β are indiscernible by entities of some type γ greater than α and β [that is, $\forall z^\gamma(z^\gamma(x^\alpha) \leftrightarrow z^\gamma(y^\beta))$]. Let $x^\alpha \equiv y^\beta$ abbreviate $x^\alpha \equiv_\gamma y^\beta$, where $\gamma = \max(\alpha, \beta) + 1$.[19] We can then prove the following lemma, which shows that $\equiv$ behaves precisely like identity.[20]

Proposition 1 $\equiv$ is an equivalence relation. Moreover, if $\phi(x^\alpha)$ and $\phi(y^\beta)$ are well formed,

[18] The first and third assumptions ensure that the types can be taken to be an initial segment of the ordinals. We next make two further assumptions, which were implicit in the main text but should be uncontroversial. First we assume that type$(e) = \sup\{\text{type}(f) + 1 : \ulcorner e(f) \urcorner \text{ is well-formed}\}$. Then we assume that, if e applies to expressions whose types are cofinal in a limit ordinal λ, then e applies to expressions of all types $< \lambda$. The four assumptions just mentioned, plus the second one mentioned in the main text, entail full cumulativity. The proof is left as an exercise for the reader.

[19] Note that the first of these two definitions delivers a different two-place predicate '$\equiv_\gamma$' for each γ, rather than a single three-place predicate '$\equiv$' in which one of the arguments is filled by a type-variable. The second definition enables a contextual elimination of any occurrence of '$\equiv$' in favor of some appropriate predicate of the form '$\equiv_\gamma$'.

[20] Cf. (Degen and Johannsen, 2000), Proposition 1. The only type-theoretic assumption used in the proof is the comprehension scheme described in Section 6.4 below.

then we can prove:

$$x^\alpha \equiv y^\beta \to [\phi(x^\alpha) \leftrightarrow \phi(y^\beta)]$$

In the rest of the main text we will mostly assume that the role of identity is played by this equivalence relation $\equiv$. However, for technical reasons we will also allow languages to contain a primitive identity predicate which can be flanked by terms of the same type.[21]

Once this type-unrestricted version of identity is in place, defining a type-unrestricted notion of predication is straightforward. One simply uses type-unrestricted identity to raise the type of the predicate enough to ensure that the predication is legitimate. Consider two terms s^α and t^β of any types α and β available in the relevant language. Recall that the predication $s^\alpha(t^\beta)$ is defined only when $\beta < \alpha$. Whether or not this condition is met, we let $t^\beta \,\varepsilon\, s^\alpha$ abbreviate $\exists z^\gamma[s^\alpha \equiv z^\gamma \wedge z^\gamma(t^\beta)]$, where $\gamma = \max(\alpha, \beta+1)$. It is easy to see that $t^\beta \,\varepsilon\, s^\alpha \leftrightarrow s^\alpha(t^\beta)$ if $\beta < \alpha$. If one wants an untyped notion of predication, it is therefore sufficient to use '$t^\beta \,\varepsilon\, s^\alpha$' in place of '$s^\alpha(t^\beta)$'. (In what follows, we won't bother maintaining the two notations and read '$s^\alpha(t^\beta)$' as type-unrestricted.)[22]

This completes our defense of Gödel's claim that the three 'superfluous restrictions' should be lifted. As announced, we believe our defense to be broadly in the spirit of Gödel's proposal.[23] We also believe our defense vindicates Gödel's view that the three restrictions are superfluous. Lifting the first restriction amounts to an extension of the original theory rather than a revision of it. The second and third restriction would constitute genuine modifications of the theory if taken as primitive. But as emphasized above, the effect of lifting these restrictions can, if desired, be simulated by means of coding. And when simulated, the second and third modifications amount to no modification at all.

[21] More specifically, the technical results in Appendix B would have to be weakened if we did not have a primitive identity predicate.

[22] This definition requires a slight modification of Gödel's claim that predications previously regarded as ill-formed should be regarded as false. Assume $x^0 \equiv x^\omega$ and $y^1(x^0)$. Then $y^1(x^\omega)$ comes out true on our definition. More generally, our definition allows a variable of lower type to be true of a variable of higher type if the type of the latter is needlessly high. However, Gödel's claim remains true if all types are minimal.

[23] Notice in particular that our argument for the first step, like Gödel's, is based on observations about the expressive limitations of languages of any fixed level: see the passage quoted at the end of Section 3, as well as surrounding text from (Gödel, 1933).

6 The transformation of type theory to iterative set theory

We now turn to Gödel's claim that removing the the restrictions effects a transformation of type theory to iterative set theory. Recall that in the passage with which we started Gödel claims that the axiomatization of set theory "presented by Zermelo, Fraenkel and von Neumann [...] is nothing else but a natural generalization of the theory of types, or rather, it is what becomes of the theory of types if certain superfluous restrictions are removed."

6.1 Translating between type theory and a restricted set theory

We begin by providing precise characterizations of the relevant higher-order languages and some analogous (but unusual) set-theoretic languages, and by showing how to translate between these languages.

The pure cumulative language $\mathcal{L}^\alpha$ of order α is defined by a simultaneous induction on formulas and terms. The language is based on the usual connectives, an identity predicate that may be flanked by any two terms of equal type, a countably infinite supply of variables for each type $\beta < \alpha$, and quantifiers capable of binding these variables. In addition, for every formula $\phi(x^\gamma)$ with x^γ free and $\gamma + 1 < \alpha$, there is an *abstraction term* $(\lambda x^\gamma\, \phi(x^\gamma))^{\gamma+1}$ of type $\gamma + 1$. Intuitively, this term stands for the $(\gamma+1)$-level concept of being an x^γ such that $\phi(x^\gamma)$.

Recall that the values of a variable of limit type will consist of the 'union' of the values of variables of all preceding levels, and the values of a variable of successor type will consist of the 'union' of the values of variables of the proceeding level plus all 'collections' of values of variables of that level.

Let $\mathcal{L}^\alpha_\in$ be the usual language of set theory augmented with new primitive terms 'V_β' for each $\beta < \alpha$ and the usual formation rules except that all quantification must be restricted to some V_β. Note that in this language the ranks are provided 'from the outside', rather than generated and described within the theory as is customary in ordinary ZF. This parallels the way in which type theory receives its types 'from the outside'.

We define a translation $*$ from $\mathcal{L}^\alpha$ to $\mathcal{L}^\alpha_\in$ as follows:

(1) $$s^\beta = t^\beta \stackrel{*}{\mapsto} s^\beta = t^\beta$$

(2) $$s^\gamma(t^\beta) \stackrel{*}{\mapsto} t^\beta \in s^\gamma$$

(3) $$\forall x^\gamma \phi \stackrel{*}{\mapsto} (\forall x^\gamma \in V_\gamma)\phi^*$$

(4) $$(\lambda x^\gamma\, \phi(x^\gamma))^{\gamma+1} \stackrel{*}{\mapsto} \{x^\gamma \in V_\gamma : \phi^*(x^\gamma)\}$$

Note that the occurrences of type-indices on the right-hand side is not required by the syntax of $\mathcal{L}^\alpha_\in$. A variable of the form x^α can nevertheless be construed as a first-order variable by regarding the superscript not as a type index but rather as having the same function as primes (as in x' and x''), namely that of producing distinct first-order variables.

What about the reverse direction? Here the main challenge is to find a way to map the untyped variables of $\mathcal{L}^\alpha_\in$ to the typed variables of $\mathcal{L}^\alpha$. Fortunately, every bound variable of $\mathcal{L}^\alpha_\in$ is implicitly typed, since the quantifier doing the binding must be restricted to some V_β. This allows us to describe a reverse translation $\dagger$, defined on all sentences of $\mathcal{L}^\alpha_\in$. First we add type indices to variable-occurrences: if v is bound by a quantifier restricted to V_β, its occurrences should be rewritten v^β. Then we translate as follows:

(5) $$s^\gamma = t^\beta \stackrel{\dagger}{\mapsto} s^\gamma \equiv t^\beta$$

(6) $$t^\beta \in s^\gamma \stackrel{\dagger}{\mapsto} s^\gamma(t^\beta)$$

(7) $$(\forall x^\gamma \in V_\gamma)\phi \stackrel{\dagger}{\mapsto} (\forall x^\gamma)\phi^\dagger$$

where $s^\gamma \equiv t^\beta$ and $s^\gamma(t^\beta)$ are 'unpacked' as described in Section 5. If cumulativity is regarded as primitive, the resulting formulas will involve bound variables of types as high as $\max(\gamma, \beta)+1$ and $\max(\gamma, \beta + 1) + 1$ respectively. This means that the translation $\dagger$ will always map formulas of $\mathcal{L}^\alpha_\in$ to formulas of $\mathcal{L}^{\alpha+2}$ (or $\mathcal{L}^\alpha$, if α is a limit ordinal).

We now describe a way in which the translations preserve truth value. Assume a type theorist and a set theorist confront a domain of individuals. The type theorist is interested in the ideological hierarchy of concepts based on these individuals. By contrast, the set theorist is interested in the ontological hierarchy of sets based on these individuals regarded as urelements. Assume that the individuals in question form a set. (The possibility that they not form a set will be considered in the next subsection.) Then it is easy to verify that the translations preserve truth value. For in both cases a limit level consists of the 'union' of the values of variables of all proceeding levels. And in both cases a successor level consists of the 'union' of the values of variables of the proceeding level plus all 'collections' of values of variables of that level. This ensures that the concepts of type γ are isomorphic to the sets in V_γ, for γ an arbitrary ordinal. It follows that $*$ must map every truth of $\mathcal{L}^\alpha$ to a true sentence of $\mathcal{L}^\alpha_\in$, and that $\dagger$ must map every truth of $\mathcal{L}^\alpha_\in$ to a true sentence of $\mathcal{L}^{\alpha+2}$.

Our argument for the claim that the translations preserve truth-value takes place against a background consisting of both type theory and set theory. The argument is thus only available to people who are willing to presuppose ordinary set theory. We help ourselves to two assumptions about the relationship between the hierarchies. Firstly, if the two hierarchies are isomorphic up to a certain level, then there is a unique way to extend this isomorphism one level further. Secondly, if there are compatible isomorphisms for all levels $\gamma < \lambda$, then there is a unique way to combine them into an isomorphism for level λ. Based on these two non-trivial assumptions, our argument clarifies the relationship between type theory and set theory by establishing that the translations preserve truth-value.

A proof-theoretic comparison of the two hierarchies—which relies on much weaker assumptions—will be provided in Section 6.4.

6.2 A set of all individuals?

What if the individuals that the type theorist and the set theorist confront do not form a set? For instance, these individuals may comprise *the entire universe of sets*. In such cases, the type theorist appears to be at an advantage, as she can consider the ideological hierarchy built on top of these individuals; that is, the ideological hierarchy where all the sets figure as individuals. The set theorist appears unable to match this. For if we really start with *all the sets*, then there are no *further* sets to be formed that can play the role of the type-theorist's 'higher-order entities'.

Gödel himself is not concerned that a given domain of individuals may fail to form a set:

> all the classes occurring in this system [of axioms] can be considered as a new domain of individuals and used as a starting point for creating still higher types. There is no end to this process. ((Gödel, 1933), p. 47)

Just as in type theory it is always permissible to go to the next higher type, in set theory it is always permissible to regard a domain of sets as a set of some still larger domain—or so the passage seems to suggest. A similar idea is familiar from (Zermelo, 1930). If this is right, then it is always possible to regard a given domain of individuals as a set in some extended process of set formation. If the universe of sets is open-ended in this way, then the set theorist will not, after all, be at any disadvantage vis-à-vis the type theorist.

A related but weaker point can be made even without Gödel's and Zermelo's view about the open-endedness of the universe of sets. Type theorists propose an ideological hierarchy in which the totality of sets may figure as individuals. But each level of this ideological hierarchy is isomorphic to the corresponding level of 'further sets' which Gödel and Zermelo claim can be formed. So at least from a mathematical point of view, there is no deep difference between ascending in this ideological hierarchy and taking further ontological steps of the sort that Gödel and Zermelo are advocating. We return, in the final section, to the question whether there might be an important *philosophical* difference between these two isomorphic ways of extending a given hierarchy.

6.3 Translating from ordinary set theory into type theory

In Section 6.1 we showed how to provide type-theoretic translations of set-theoretic sentences all of whose quantifiers are explicitly restricted to some V_β. But the language of ordinary set theory allows quantifiers which are not explicitly restricted in this way. So we do not yet have a translation-method for the language of ordinary set theory.

The standard response to this sort of problem is to map each set-theoretic sentence ϕ onto the sentence $\phi^{(\alpha)}$, which is the result of restricting all the quantifiers in ϕ to V_α, for some suitable ordinal α. This yields a syntactic translation from $\mathcal{L}_\in$ to $\mathcal{L}_\in^{\alpha+1}$. To what extent does this translation give us what we want?

Let us start with the question of preservation of truth-value. Recall that Zermelo set theory, Z, is ZF minus the axiom scheme of Replacement. Z is satisfied by V_λ for every limit ordinal $\lambda \geq \omega + \omega$, so the translation $\phi \mapsto \phi^{(\lambda)}$ maps every theorem of Z to a truth of $\mathcal{L}_\in^{\lambda+1}$. By choosing larger ordinals we can get stronger results. Since ZF is satisfied by V_κ for every strongly inaccessible κ, the translation $\phi \mapsto \phi^{(\kappa)}$ will map every theorem of ZF to a truth of $\mathcal{L}_\in^{\kappa+1}$. (Of course, this appeal to strong inaccessibles involves going beyond ZFC.)[24]

Now consider the requirement that the translation-method preserve not just the truth-value of the original sentence, but also its *intended interpretation*. Whether or not one thinks that the translation $\phi \mapsto \phi^{(\alpha)}$ can satisfy this stricter requirement will depend on one's attitude towards sets. If one accepts Absolute Generality, and believes that it is possible to quantify over all sets, then one will think that the stricter requirement fails to be satisfied. For although we are capable of quantifying over all sets, no claim resulting from this translation involves such quantification.

However, if one adopts the view of Gödel and Zermelo that we discussed in the previous subsection, then one may think that $\phi \mapsto \phi^{(\alpha)}$ is able to preserve intended interpretation. For one may think that seemingly unrestricted set-theoretic quantifiers are always tacitly restricted to V_Ω, the domain of sets that have been recognized so far—a domain that could be used 'as a starting point' to characterize an even larger domain of sets.

When the translation $\phi \mapsto \phi^{(\alpha)}$ is composed with the translation $\phi \mapsto \dagger\phi$, from Section 6.1, one gets a function that translates each sentence in the language of ordinary set theory into a sentence of transfinite type theory. This means that, regardless on one's attitude towards Absolute Generality, one should think that there is a truth-value-preserving translation of ordinary set-theory into transfinite type theory, modulo the existence of suitably large cardinals. Depending on one's attitude towards Absolute Generality, one might also think that the translation preserves intended interpretation.

[24] Assuming a suitable reflection principle can be justified, one can even show that there is a cardinal ξ such that V_ξ satisfies the same $\mathcal{L}_\in$-sentences as the universe. (See (Shapiro, 1987), pp. 323–4.) Then the translation $\phi \mapsto \phi^{(\xi)}$ maps every truth of $\mathcal{L}_\in$ to a truth of $\mathcal{L}_\in^{\xi+1}$.

6.4 A proof-theoretic comparison of type theory and set theory

Let *pure cumulative logic of order* α be the $\mathcal{L}^\alpha$-theory T^α which arises by making the following additions to a complete axiomatization of first-order logic without identity. First, we extend the usual introduction and elimination rules for the quantifiers to quantification of all types.

Second, we adopt the following axiom scheme of type-raising: $\forall x^\alpha\, \exists x^\beta (x^\alpha \equiv x^\beta)$ for any $\beta \geq \alpha \geq 0$.[25]

Third, we adopt a *limit rule* which expresses the idea that every limit type $\lambda < \alpha$ is 'the union' of all preceding types. If we have proved $\forall x^\gamma\, \phi(x^\gamma)$ for every $\gamma < \lambda$, we may infer $\forall x^\lambda\, \phi(x^\lambda)$.

Fourth, we adopt a *comprehension scheme* for every type γ such that $\gamma + 1 < \alpha$:

$$\text{(Comp)} \qquad (\lambda x^\gamma\, \phi(x^\gamma))^{\gamma+1}(t^\delta) \leftrightarrow \phi(t^\delta)$$

provided $\phi(t^\delta)$ is well-formed.

Fifth, we adopt an axiom scheme of extensionality. Consider two non-empty entities x^β and y^γ, where $\beta \leq \gamma < \alpha$. Assume x^β and y^γ are coextensional in the sense that any entity falling under the former is equivalent to an entity falling under the latter, and that any entity falling under the latter is equivalent to an entity falling under the former. Then the axiom of extensionality says that $x^\beta \equiv y^\gamma$.[26]

Finally, we adopt an axiom scheme to the effect that no individual applies to any other entity: $\neg x^0(y^\alpha)$.[27]

How much of ordinary ZF set theory can be obtained from type theory? Let's work with the composite translation $\phi \mapsto \dagger\phi^{(\alpha)}$ described above. The question is then which axioms of ZF are mapped by this translation to theorems of $T^{\alpha+3}$ (as opposed to truths of $\mathcal{L}^{\alpha+3}$, which

[25] In the absence of a primitive identity predicate there is an alternative but equivalent approach based on a simple modification of the introduction and elimination rules for the quantifiers. See (Degen and Johannsen, 2000), p. 149 for details. Given the version of Extensionality set out below, we can prove every instance of our axiom scheme where $\alpha \geq 1$. The trick is to consider the term $(\lambda u^\beta\, x^\alpha(u^\beta))^{\beta+1}$ and to make a judicious use of Proposition 3 of (Degen and Johannsen, 2000).

[26] The non-emptiness requirement is needed to handle the case where β or γ is 0. For individuals can be distinct from each other and from the empty concept although none applies to anything else.

[27] Our axiomatization follows that of (Degen and Johannsen, 2000) with two minor exceptions needed to allow distinct individuals: the extensionality rule has been restricted to *non-empty* entities (which are thus non-individuals), and the final axiom scheme replaces Degen and Johannsen's 'null rule'.

is what we studied above). (In more technical parlance, the question is how much of ZF can be *interpreted* in $T^{\alpha+3}$ under the mentioned translation.) This question receives a partial answer by the following proposition:

Proposition 2 (Degen and Johannsen) Let $\lambda \geq \omega + \omega$ be a countable limit ordinal and $\alpha \geq \lambda + 3$. Then T^α proves $\dagger\phi^{(\lambda)}$ for each theorem ϕ of Zermelo set theory Z. (In more technical parlance, Z is interpreted in T^α under the mentioned translation.)

We refer the reader to (Degen and Johannsen, 2000) for a proof.[28]

7 Philosophical assessment

Our discussion so far can be thought of as an extended argument for the following two claims: (1) the sorts of considerations that motivate an interest in the first few levels of the theory of types also motivate an ascent up to higher and higher types (Section 3 and Appendix B); (2) the resulting hierarchy becomes (as Gödel observes) very set-theoretic in character (Sections 4–6). This shows that there is no deep *mathematical* difference between the ideological hierarchy of type theory and the ontological hierarchy of set theory. But there might of course be an important *philosophical difference*. The purpose of this final section is to discuss whether this is so.

One possibility is to take our formal results to show that one of the hierarchies should be eschewed in favor of the other. For instance, a Quinean might see the formal results as an ultimate vindication of the claim that type theory is 'set-theory in sheep's clothing' (Quine, 1986). More specifically, she might use the formal resemblance between the type theoretic hierarchy and the set theoretic hierarchy to cast doubt on the idea that the language of type-theory makes sense when taken at face value: it only makes sense as a notational variant

[28]The translations of several of the axioms of Z (such as Extensionality restricted to non-empty sets, Empty-set, Pairing, Union, Power-set, Separation, and Infinity) can be proved directly in our type-theory. This has the advantage over Degen and Johannsen's approach of yielding explicit proofs and of avoiding the requirement that the limit ordinal λ be countable. We don't know whether (the translations of the instances of) Replacement can be established in this way.

Gödel intended to provide a more internal motivation of Replacement, where the levels α are generated internally rather than provided from outside, as above. But it is unclear exactly how the argument is supposed to work. For some scholarly discussion, see (Feferman, 1995), esp. note f.

of the language of set-theory. Alternatively, one might see the formal results as rendering the set-theoretic hierarchy obsolete, by showing that type-theory gives us all the expressive resources of set-theory without any of its ontological commitments. (It is worth keeping in mind, however, that we used an external theory of ordinals to set up the type-theoretic hierarchy, and it is not obvious that this could be avoided.)

Our own view is that it would be a mistake to eschew one of the hierarchies over the other. We would like to suggest instead that the two hierarchies—ideological and ontological—constitute difference perspectives on the same subject-matter. The remainder of this section is an effort to clarify what this means.

It is useful to start with some preliminaries. Logicians and philosophers are divided in their reactions to Russell's Paradox. Some are liberal about set formation and assume that there is a set of Fs whenever there is a definite fact of the matter about what Fs there are. *Liberalists* of this sort see the Paradox as showing that there is no definite fact of the matter about what sets there are, and therefore that the hierarchy of sets is 'open-ended'.[29] The alternative reaction to Russell's paradox is to be restrictive about set formation and be prepared to deny that there is a set of Fs even when there is a definite fact of the matter about what Fs there are. *Non-liberalists* of this sort can thus maintain that there is a definite fact of the matter about what sets there are, and that Russell's Paradox relates to one of many cases where some definite range of Fs fail to form a set.[30] This also enables non-liberalists to deny that there is any interesting sense in which the hierarchy of sets is 'open-ended'.

These two reactions to the Paradox tend to be linked to different ways of thinking about sets. Non-liberalists tend to see the existence of sets as a substantial matter. (Metaphorically: when God created Bruno, He did not *thereby* create Bruno's singleton; an *additional* action on God's part would be required to bring the set into existence.) Accordingly, for it to be the case that Bruno is a member of his singleton it is not *sufficient* that Bruno exist. It is required, *in addition*, that Bruno's singleton exist and that it bear a certain relation to Bruno. Similarly, for Bruno to be a member of the set of elephants it is not *sufficient* that Bruno be an elephant. It is required, *in addition*, that there be a set of elephants and that it bear

[29] Examples of liberalists include (Gödel, 1933), (Zermelo, 1930), (Parsons, 1977).

[30] Examples of non-liberalists include (Boolos, 1984), (Lewis, 1991) and (McGee, 1997).

a certain relation to Bruno. On this view, it is natural to think that there is a definite fact of the matter about what sets exist. (Whichever sets God found it in His wisdom to create, those are the sets that exist.) And—assuming there is a definite fact of the matter about which sets exist—it is a truth of (higher-order) logic that there are more collections of sets than sets. So one immediately gets the result that not every collection of sets forms a set.

In contrast, liberalists tend to deny that there is a substantial gap between the existence of Bruno and the existence of his singleton, or between there being a definite fact of the matter about what elephants there are and the existence of a set of all elephants. Different liberals develop this idea in different ways. Some deny that there is any gap whatsoever:[31] for Bruno to be a member of the set of elephants *just is* for Bruno to be an elephant, and for Bruno to be a member of his singleton *just is* for Bruno to be identical to Bruno. So when God creates Bruno he *thereby* makes it the case that his singleton exists. On this view, 'Bruno is an elephant' and 'Bruno is a member of the set of elephants' describe the *same* feature of reality. Accordingly, the use of a singular term like 'the set of elephants' increases one's expressive resources *without increasing one's ontological commitments*. (This is *not* to say that 'Bruno is a member of the set of elephants' carries no commitment to the set of elephants. It certainly does. The point is that such commitment was already presupposed by relevant uses of the term 'elephant'. For it is nothing over and above commitment to there being a definite fact of the matter about what elephants there are: for the set of elephants to exist *just is* for there to be a definite fact of the matter about what elephants there are.)

Other liberalists accept that there is an ontological gap between the existence of Bruno and the existence of his singleton but deny that the gap is a substantive one.[32] These liberalists accept that 'Bruno's singleton is an elephant-singleton' carries an ontological commitment not carried by 'Bruno is an elephant', namely a commitment to sets. But they argue that sets are 'lightweight' objects whose existence does not make any substantive demand on the world. Perhaps the existence of Bruno's singleton amounts to nothing more than the coherence or legitimacy of some set theory which asserts the existence of this set. Or perhaps it amounts to nothing more than an ability to refer to the set, make true predications of

[31] This is the view of AR.
[32] This is the view of ØL.

it, and distinguish it from other sets. Regardless of the details, liberalists of this sort insist that, if a legitimate mathematical practice endorses the truth of a sentence whose logical form reveals a commitment to sets, then no additional contribution from the world is needed to guarantee the existence of sets.

Let's put aside internal disagreements among liberalists and focus on their common conviction that there is no substantial ontological gap between there being a definite fact of the matter about what Fs there are and the existence of a set of all Fs. We have seen that this conviction, when applied to Russell's Paradox, yields the conclusion that there can be no definite fact of the matter about what sets there are. (For if there were, there would be a Russell set, which would yield a contradiction.) This conclusion should come as no surprise to liberalists. For the liberalist view makes it impossible to draw a clean separation between the question of how one might extend one's expressive resources and the question of how many sets exist. By increasing one's expressive resources in the right sort of way, one is led to recognize additional ontology. So insofar as one believes that the process of extending one's expressive resources is essentially open-ended, one should also think that the hierarchy of sets is essentially open-ended—and therefore that there is no definite fact of the matter about what sets there are.

The connection between type theory and set theory that has been developed in this paper can be used to make some of the liberalist ideas more explicit. Let ϕ be a sentence of the set-theoretic language $\mathcal{L}^{\alpha}_{\in}$ (introduced in section 6.1), and let $\phi^{\dagger}$ be the type-theoretic sentence that results from applying the †-translation. Then the liberalist may claim that there is no substantial difference between the features of the world described by ϕ and $\phi^{\dagger}$. She may claim, for example, that there is no substantial difference between the feature of reality described by the set-theoretic sentence '$\exists x \in V_1(\text{Bruno} \in x)$' and the feature of reality described by the type-theoretic sentence '$\exists x^1(x^1(\text{Bruno}))$'. (This, of course, applies only to sentences in $\mathcal{L}^{\alpha}_{\in}$. But that is not something liberals are likely to see as a cost, since they think there is no definite fact of the matter about what sets there are.)

For the non-liberalist, the connection between type theory and set theory can be expected to have a very different upshot. On any interesting interpretation of the type-theoretic hierarchy, sentences such as '$\exists x \in V_1(\text{Bruno} \in x)$' and '$\exists x^1(x^1(\text{Bruno}))$' will be seen as making very different ontological demands on the world: only the former will be seen as requiring the existence of sets. So the non-liberalist will see it as a mistake to think that the type-theoretic and set-theoretic hierarchies constitute different perspectives on the same subject-matter. She might instead take one of the attitudes alluded to earlier, and see the formal connection between the two hierarchies as evidence that one of the hierarchies should be eschewed in favor of the other. As noted above, one possibility is for the non-liberalist to see the ideological hierarchy as superior, on the grounds that it delivers expressive resources without ontological extravagance. Another possibility is for her to conclude that the ideological hierarchy is too good to be true, and conclude that it cannot be made sense of when taken at face value.

There is, however, a third—and more intriguing—possibility. The non-liberalist might come to see the connection between type theory and set theory as a reason for moving in the liberalist direction. For one might have thought that a big selling point of non-liberalism was its tidy ontology: there is no need to countenance an open-ended hierarchy of sets, and no reason to doubt the truth of Absolute Generality. But once one notices that Absolute Generality can be used to motivate ascent into higher and higher levels of the ideological hierarchy, one might come to see the supposed tidiness of non-liberalism as an illusion.

We noted in Section 3 that Absolute Generality—together with the Principle of Semantic Optimism and the Principle of Union—can be used to motivate ascent to languages of type α for α an arbitrary ordinal. This entails that the type-theoretic hierarchy has at least as many levels as there are ordinals. But it does not, by itself, deliver the conclusion that the hierarchy is open-ended. By strengthening the Principle of Union, however, it is possible to generate an argument for open-endedness:

- The Principle of Union (Strengthened Version)

  Let C by any definite collection of type-theoretic languages. Then one should countenance the legitimacy of the 'union' language L_C, which results from pooling together the resources of every language in C.

The argument is straightforward. Suppose for *reductio* that the type-theoretic hierarchy is not open-ended. There is, in other words, a definite collection of type-theoretic languages. It follows from the Principle of Union that one should countenance the legitimacy of the 'union' language L_T which results from pooling together the resources of every type-theoretic language. Assuming ABSOLUTE GENERALITY, it is impossible to state a generalized semantic theory for L_T in L_T. But, by the Principle of Semantic Optimism, there should be a language in which such a theory could be articulated. And the most natural way of doing so is by countenancing the legitimacy of further types.

If this is right, then open-endedness is here to stay. One can try to move it from one's ontology to one's ideology, but nothing of any real substance will have changed.[33]

[33]For their many helpful comments, we would like to thank Eduardo Barrio, Roderick Batchelor, Javier Castro Albano, Roy Cook, Wolfgang Degen, Salvatore Florio, David Nicolas, Ignacio Ojea, Lavinia Picollo, Pedro Santos, Gabriel Uzquiano, and Philip Welch, as well as seminar participants at the Institute for Philosophical Research of the National Autonomous University of Mexico, and audiences at the Universities of Bristol, Buenos Aires, Lund, and São Paulo, and at the 2009 meeting of the Swiss Society for Logic and the Philosophy of Science. Special thanks are due to Stewart Shapiro.

Appendices

A The plural hierarchy

In the main text we work on the assumption that the type-theoretic hierarchy is interpreted conceptually: that type-1 variables stand for first-level concepts (i.e. concepts applying to objects), type-2 variables stand for second-level concepts (i.e. concepts applying to first-order concepts; or concepts applying to first-order concepts *and individuals*, if one thinks of cumulativity as primitive rather than simulated), and so forth. But the hierarchy could also be interpreted *plurally*: one could read type-1 variables as plural terms, type-2 variables as 'super-plural' terms, and so forth, all the way up.[34]

Some changes are needed to accommodate a plural interpretation of the type-theoretic hierarchy. First the *syntax* of the language needs to be adjusted. On the conceptual interpretation, variables of higher type stand for concepts, and are therefore of the same syntactic category as predicates. Since predicates have argument-places, this means that variables of higher type should also be thought of as having argument-places. On the plural interpretation, in contrast, variables of higher type are *terms*. Type-1 variables, for example, should be thought of as sharing a syntactic category with the plural term 'the horses', rather than the predicate '... is a horse'. Accordingly, the formula '$x^1(z^0)$' should be thought of as syntactically analogous to 'z^0 is among the horses' (or '$\text{Among}(z^0, \text{the horses})$'), not as syntactically analogous to '$\text{Horse}(z^0)$'. Pluralists should therefore think of '$x^{\beta+1}(z^\beta)$' as a syntactic abbreviation for '$\text{Among}^{\beta+2}(z^\beta, x^{\beta+1})$', where '$\text{Among}^{\beta+2}$' is a new logical predicate.

The plural interpretation requires an adjustment of *the deductive system* as well. There are two main differences that need to be accounted for. The first concerns the question of whether variables of higher-type can take 'empty' values. For instance, is there an 'empty' level-1 entity, that is, an entity x^1 such that $\forall z^0\ (\neg x^1(z^0))$? On the conceptual interpretation, the answer is 'yes': there is a first-level concept with no instances. But on the plural interpretation, the answer is 'no': it is not the case that there are some things such that nothing

[34] A plural interpretation of the finite types is discussed in (Rayo, 2006).

is one of them. The pluralist should therefore adopt axioms which reflect facts of this sort. She might, for instance, set forth an axiom $\forall x^{\alpha+1}\,\exists z^{\alpha}(x^{\alpha+1}(z^{\alpha}))$ for every suitable α, and restrict comprehension as suggested below.

The second difference concerns the question of whether to admit 'singleton pluralities'. Consider for instance the plural term 'the things identical to Socrates', or in lambda-notation: $(\lambda u^0(u^0 = \text{Socrates}))^1$. It is natural to think that this plural term refers to Socrates and nothing else. There is no need to posit as a referent some 'singleton plurality' which consists of nothing but Socrates but which is somehow endowed with properties distinct from those of the philosopher. Rather, the referent can simply be taken to be Socrates himself. (Note that this uses cumulativity; otherwise a term of type 1 could not denote an object.)

This second difference requires a modification of the plural comprehension scheme to ensure that the ordinary 'lambda-conversion' is available only when the comprehension formula is at least doubly instantiated, and that lambda-terms whose comprehension formula is singly instantiated simply denote this single instance:[35]

$$\text{(Comp}'\text{)} \qquad \begin{cases} (\exists_{\geq 2} x^{\gamma}\,\phi(x^{\gamma}) \rightarrow (\lambda x^{\gamma}\,\phi(x^{\gamma}))^{\gamma+1}(t^{\delta}) \leftrightarrow \phi(t^{\delta})) \\ \exists y^{\gamma}[\forall u^{\gamma}(y^{\gamma} \equiv u^{\gamma} \leftrightarrow \phi(u^{\gamma})) \rightarrow (\lambda x^{\gamma}\,\phi(x^{\gamma}))^{\gamma+1} \equiv y^{\gamma}] \end{cases}$$

where the former clause is subject to the proviso that $\phi(t^{\delta})$ is well-formed.

Our comparison in Section 6 of the type-theoretic hierarchy with the set-theoretic hierarchy presupposed a conceptual interpretation of the hierarchy and a corresponding deductive system. But we have seen that the plural interpretation calls for a slightly different deductive system. This raises the question whether our comparison carries over to the plural interpretation of the type-theoretic hierarchy.

Fortunately, versions of the original results do carry over to the plural case. Let $V_{\alpha}^{\geq 2}$ be

[35] What about the claim that Socrates is among the things identical to Socrates? On the current proposal this claim is formalized as $(\lambda u^0(u^0 = \text{Socrates}))^1(\text{Socrates})$, which is false because the lambda-term denotes Socrates, and in the intended model no entity applies to itself (or, in official language, 'is among' itself). Anyone troubled by this quirk is invited to re-define the predicate 'Among' to stand for the reflexive closure of the relation currently associated with that predicate, that is, to lay down: $\text{Among}^{\beta+1}(x^{\alpha}, y^{\beta}) \leftrightarrow x^{\alpha} = y^{\beta} \vee y^{\beta}(x^{\alpha})$, where $\alpha < \beta$.

the subset of V_α that is generated by iterating a *restricted* version of the powerset operation:

$$\mathcal{P}^{\geq 2}(x) = \{z : z \subseteq x \;\wedge\; \exists w \exists v(w \neq v \wedge w \in z \wedge v \in z)\}$$

(We will assume that there are at least two urelements, as this construction will otherwise be trivial.) By induction on α we can then show that the translations $*$ and $\dagger$ map truths about the plural hierarchy of level α to truths about $V_\alpha^{\geq 2}$ and *vice versa.* Moreover, it is easy to show that the effect of the full powerset operation can be simulated by using only the restricted powerset operation. In particular, for any limit ordinal λ, an isomorphic copy of V_λ can be found as a subset of $V_\lambda^{\geq 2}$. It follows that V_λ is isomorphic to a sub-hierarchy of the plural hierarchy of level λ. It should also be possible to provide an interpretation of Zermelo set theory in the plural analogue of the type theory $T^{\lambda+3}$, thus establishing a plural analogue of Theorem 2.

It is also worth noting that—on the assumption that there are at least two urelements—one can prove plural analogues of Theorems 4 and 5 from Appendix B.

B Generalized Semantics

A fundamental tool in our constructions will be the ability to code ordered pairs.

Theorem 1 (n-tuples) Given any entities $x^{\gamma_1}, \ldots, x^{\gamma_n}$, we can code for the ordered n-tuple of these entities by means of a single entity x^γ, where γ is the maximum of the γ_i. We will designate this entity x^γ as $\langle x^{\gamma_1}, \ldots, x^{\gamma_n} \rangle^\gamma$.

Proof and details will be provided in Section B.2.

Recall from Section 2 that a language of order α contains variables of all orders $< \alpha$ and constants of all orders $\leq \alpha$.

Theorem 2 (Negative result) We cannot develop a generalized semantics for a language of order α in another language of order α.

Proof. The crucial observation is that, for $\alpha = \beta + 1$ a successor ordinal, there are more possible assignments of semantic value to a constant c^α than there are entities of level β. This cardinality claim is a higher-order version of Cantor's Theorem. Assume for contradiction that we have a collection of ordered pairs $\langle x^\beta, y^\alpha \rangle^\alpha$ which codes an onto correlation of entities of order β with entities of order α. Let Δ^α be the entity consisting of all and only those x^β such that x^β does not fall under the entity which this correlation associates with x^β. Since the correlation has been assumed to be onto, there must be a δ^β which is correlated with Δ^α. But then we get $\Delta^\alpha(\delta^\beta) \leftrightarrow \neg\Delta^\alpha(\delta^\beta)$, which yields a contradiction.

To develop a generalized semantics for a language of order $\alpha = \beta + 1$, one needs to talk about arbitrary assignments of semantic value to constants c^α of type α. But, by the observation above, this requires variables of order $\geq \alpha$. This proves the theorem for the special case in which α is a successor ordinal.

The case where α is a limit ordinal is also straightforward. In order to develop a generalized semantics for a language of order α in a language of order α, one would have to represent arbitrary assignments of semantic values to c^α by means of some variable x^γ where $\gamma < \alpha$. But by the cardinality result we know this to be impossible. $\dashv$

Theorem 3 (Positive result) For any successor level α, we can develop a generalized semantics for a language of order α in a language of order $\alpha + 1$. (*A fortiori*, for any limit level λ, we can develop a generalized semantics for a language of order λ in a language of order $\lambda + 2$.)

A proof will be provided in the next section.

B.1 Generalized semantics for $\mathcal{L}^\alpha$

When investigating the semantics of a language $\mathcal{L}^\alpha$, we will always work relative to some base theory which is strong enough to code the syntax of $\mathcal{L}^\alpha$. As usual, let $\ulcorner e \urcorner$ be the Gödel number of an expression e.

We need to talk about assignments to syntactic entities. There are two kinds of assignment. All constants will be assigned a semantic value by a *model*. And all variables will be assigned a value by a *variable assignment* (or *assignment* for short). We will use primitive predicates $\textsc{Asgn}^{\beta+1}(A^\beta)$ and $\textsc{Mod}^{\alpha+1}(M^\alpha)$ for the notions of assignment and model respectively. In Section B.3 we formulate axioms concerning these predicates which incorporate the obvious requirements on entities that are to play the role of assignments and models: namely that a model specify a domain and associate constants with entities of appropriate level which are based on the domain, and that an assignment associate variables with entities of appropriate level.[36] We show in Section B.3 that, when the order of the language $\mathcal{L}^\alpha$ is a successor ordinal $\alpha = \beta + 1$, the assignment can be chosen to be a type-β entity A^β. The fact that the type of the assignment can be chosen to be β rather than α will be crucial to the construction below. By contrast, where α is a limit ordinal, the assignment cannot be of a level lower than α. The model M^α can chosen to be of level α, but no lower, regardless of whether α is a successor or a limit.

In Section B.3 we also define an interpretation operation $[\![\cdot]\!]^\alpha_{M^\alpha}$ which for any constant c of the language $\mathcal{L}^\alpha$ outputs the entity $[\![c]\!]^\alpha_{M^\alpha}$ which the model M^α associates with c, and another interpretation operation $[\![\cdot]\!]^\beta_{A^\beta}$ which for any variable v of the language outputs the entity $[\![v]\!]^\beta_{A^\beta}$ which the assignment A^β assigns to v. Where there is no danger of confusion, we use the undecorated notation $[\![t]\!]$.

Definition 1 Say that an assignment B^β is a *v-variant* of an assignment A^β iff B^β is an assignment such that for any variable $u \neq v$, we have $[\![u]\!]^\beta_{A^\beta} = [\![u]\!]^\beta_{B^\beta}$.

Definition 2 Consider a successor ordinal $\alpha = \beta + 1$. We define the notion of satisfaction in a model M^α relative to an assignment A^β as follows:

1. If ϕ is a formula of the form $t^\gamma(t_1^{\gamma_1}, \ldots, t_n^{\gamma_n})$ for an n-place term t of arguments of the types indicated and where $1 \leq n \leq 3$, then:

   $$\mathrm{Sat}\langle \ulcorner\phi\urcorner, A^\beta\rangle \text{ iff } [\![t]\!]\langle [\![t_1^{\gamma_1}]\!], \ldots, [\![t_n^{\gamma_n}]\!]\rangle$$

2. If ϕ is a formula of the form $t_1^\gamma = t_2^\gamma$ for two terms t_1^γ and t_2^γ of the same type, then:

   $$\mathrm{Sat}\langle \ulcorner\phi\urcorner, A^\beta\rangle \text{ iff } [\![t_1^\gamma]\!] = [\![t_2^\gamma]\!]$$

[36] For bookkeeping reasons we do not require that the entities associated with the variables be based on the domain in question. This is instead taken care of by clause 5 of Definition 1 below.

3. If ϕ is a formula of the form $\neg\psi$, then:

   $\text{Sat}\langle\ulcorner\phi\urcorner, A^\beta\rangle$ iff it is not the case that $\text{Sat}\langle\ulcorner\psi\urcorner, A^\beta\rangle$

4. If ϕ is a formula of the form $\psi_1 \wedge \psi_2$, then:

   $\text{Sat}\langle\ulcorner\phi\urcorner, A^\beta\rangle$ iff $\text{Sat}\langle\ulcorner\psi_1\urcorner, A^\beta\rangle$ and $\text{Sat}\langle\ulcorner\psi_2\urcorner, A^\beta\rangle$

5. If ϕ is a formula of the form $\exists v\, \psi$ for some variable v, then:

   $\text{Sat}\langle\ulcorner\phi\urcorner, A^\beta\rangle$ iff there is a v-variant B^β of A^β, which assigns to v an entity based on the domain associated with M^α, such that $\text{Sat}\langle\ulcorner\psi\urcorner, B^\beta\rangle$

Two brief remarks are in order. Each of clauses 1 and 2 is shorthand for several clauses, which cover the cases of each term being either a constant or a variable, and, where the term t in clause 1 is a constant, the cases of its being of adicity 1, 2, or 3. We choose not to allow for constants of arbitrary finite adicity in order to keep the number of clauses finite and thus be able to conjoin them in definitions given below.[37] Secondly, the notion of an entity x^γ being 'based on' a domain $D^1_{M^\alpha}$ is the obvious one: every object in the transitive closure of x^γ is in $D^1_{M^\alpha}$, where the notion of transitive closure is defined in Section B.3.

As observed by (Tarski, 1935), we can convert the implicit definition to an explicit one.[38] Consider the result of formalizing the above definition, conjoining all of its clauses and replacing every occurrence of $\text{Sat}\langle\ulcorner\phi\urcorner, A^\beta\rangle$ by $Y^\alpha\langle\ulcorner\phi\urcorner, A^\beta\rangle$. The resulting formula has four free variables, namely $\ulcorner\phi\urcorner$,A^β, Y^α, and M^α. (Note that M^α occurs free in instances of clauses 1 and 2 where $[\![t]\!]$ is short for $[\![t]\!]_{M^\alpha}$.) Let $\Phi(\ulcorner\phi\urcorner, A^\beta, Y^\alpha, M^\alpha,)$ abbreviate this formula. Let $\textsc{True}(\ulcorner\phi\urcorner, A^\beta, M^\alpha)$ abbreviate the following formula:

$$\forall Y^\alpha[\Phi(\ulcorner\phi\urcorner, Y^\alpha, M^\alpha, A^\beta) \rightarrow Y^\alpha\langle\ulcorner\phi\urcorner, A^\beta\rangle]$$

Intuitively, $\textsc{True}(\ulcorner\phi\urcorner, A^\beta, M^\alpha)$ should be read as: ϕ is true in M^α under assignment A^β.

Theorem 4 The formula $\textsc{True}(\ulcorner\phi\urcorner, A^\beta, M^\alpha)$ satisfies the recursion clauses set out in Definition 1.

[37] However, if one was willing to allow countably infinite conjunctions, one could lift this restriction and allow for arbitrary finite adicity.

[38] As Tarski observes in fn. 1 on p. 175, the method for doing so goes back to (Frege, 1879) and (Dedekind, 1888).

Proof sketch. We need to verify an existential claim: that for any model M^α there is a Y^α such that for any $\mathcal{L}^\alpha$-formula ϕ and assignment A^β we have $\Phi(\ulcorner\phi\urcorner, Y^\alpha, M^\alpha, A^\beta)$. Once the existential claim has been established, the remainder of the proof is routine.

To verify the existential claim, we begin by defining the notion of a *good* sequence of length n ($n \in \omega$). For fixed M^α, we shall say that $\text{Good}_{M^\alpha}(x^\alpha, n)$ is true just in case the following three conditions are met:

1. x^α applies to the triple $\langle 0, \phi, A^\beta\rangle^\beta$ just in case ϕ is an atomic formula of $\mathcal{L}^\alpha$, A^β is an assignment, and either of the following conditions obtains:

    (a) ϕ is of the form $t^\gamma(t_1^{\gamma_1}, \ldots, t_m^{\gamma_m})$ ($m \leq 3$) and $[\![t]\!]\langle[\![t_1^{\gamma_1}]\!], \ldots, [\![t_m^{\gamma_m}]\!]\rangle$.

    (b) ϕ is of the form $t_1^\gamma = t_2^\gamma$ and $[\![t_1^\gamma]\!] = [\![t_2^\gamma]\!]$.

2. For $0 < k \leq n$, x^α applies to the triple $\langle k, \phi, A^\beta\rangle^\beta$ just in case ϕ is an $\mathcal{L}^\alpha$-formula of complexity k, A^β is an assignment, and one of the following conditions obtains:

    (a) ϕ is of the form $\neg\psi$, and x^α fails to apply to the triple $\langle k-1, \psi, A^\beta\rangle^\beta$.

    (b) ϕ is of the form $\psi_1 \wedge \psi_2$, and x^α applies to triples $\langle k_1, \psi_1, A^\beta\rangle^\beta$ and $\langle k_2, \psi_2, A^\beta\rangle^\beta$ for some $k_1, k_2 < k$, at least one of which immediately precedes k.

    (c) ϕ is of the form $\exists v\ \psi$, and, for some v-variant B^β of A^β, which assigns to v an entity based on the domain associated with M^α, x^α applies to the triple $\langle k-1, \psi, B^\beta\rangle^\beta$.

3. for $k > n$, x^α applies to no triples of the form $\langle k, \phi, A^\beta\rangle^\beta$.

One can use the comprehension scheme to show that there exists an x^α such that $\text{Good}_{M^\alpha}(x^\alpha, 0)$, and to show that if there exists an x^α such that $\text{Good}_{M^\alpha}(x^\alpha, k)$, then there must also exist an x^α such that $\text{Good}_{M^\alpha}(x^\alpha, k+1)$. By induction on the natural numbers, it follows that, for every n, there is an x^α such that $\text{Good}_{M^\alpha}(x^\alpha, n)$. One can then let Y^α consist of the pairs $\langle\phi, A^\beta\rangle^\beta$ such that, for some x^α and some $n \in \omega$, $\text{Good}_{M^\alpha}(x^\alpha, n)$ and $x^\alpha(\langle n, \phi, A^\beta\rangle^\beta)$. It is now routine to verify that Y^α witnesses the existential claim with which we began. $\dashv$

This concludes our characterization of the notion of satisfaction in a model for the case in which $\alpha = \beta + 1$. What about the limit case? When α is a limit ordinal, assignments will have to be of order α. This means that the satisfaction predicate has to be of order $\alpha+1$. So an explicit definition of satisfaction in a model under an assignment can only be given in a language of order $\alpha+2$.

More fine-grained results are possible if we consider various restricted languages. For instance, we can let a *basic* language of order α be one with no non-logical constants of order α. Then it is possible to show that one can to develop a generalized semantics for a basic language of order α in a non-basic language of the same order.[39] However, we will here be concerned with higher-order languages in general, not with the special class of basic ones. Our reason is that basic languages are unstable. Once variables of a certain level have been admitted, it is artificial not to allow non-logical predicates applying to these variables. In particular, on the plural interpretation there is already a top-level *logical* predicate applicable to these variables, namely 'Among'. So there is nothing wrong with top-level predicates *per se*. This suggests that it should also be legitimate to introduce *non-logical* predicates of this sort.

B.2 Ordered pairs

We shall assume that the theory of $\mathcal{L}^1$ (which is a sublanguage of the language $\mathcal{L}^\alpha$ with which we are concerned) contains the resources to code the ordered pair of any two individuals of type 0 by means of some third individual of type 0. This assumption can, as is well known, be met in a variety of different ways. (As is well known, a minimal requirement is that the domain be infinite.) To fix ideas, we shall assume that the base theory contains enough set theory to code ordered pairs of individuals in the familiar Wiener-Kuratowski fashion. So we set forth the following definition:

$$OP^1(x^0, y^0, z^0) \ \equiv_{df} \ z^0 = \{\{x^0, y^0\}, \{x^0\}\}$$

We will now show that this allows us to characterize a new predicate $OP^{\beta+1}(a^0, x^\beta, y^\beta)$ for each ordinal β between 0 and α. Intuitively, $OP^{\beta+1}(a^0, x^\beta, y^\beta)$ says that y^β codes the 'ordered

[39] See (Rayo, 2006), p. 244.

pair' composed of the individual a^0 and the higher-order entity x^β.

The new predicates are characterized axiomatically. For each β between 0 and α, we set forth an *existence* axiom for $OP^{\beta+1}$:

$$\forall a^0 \forall x^\beta \exists y^\beta \; OP^{\beta+1}(a^0, x^\beta, y^\beta)$$

and a family of *pairing* axioms for $OP^{\beta+1}$, consisting of a separate axiom for each $\delta < \beta$:[40]

$$OP^{\beta+1}(a^0, x^\beta, y^\beta) \leftrightarrow \forall v^\delta \big(y^\beta(v^\delta) \leftrightarrow \exists u^\delta (x^\beta(u^\delta) \wedge OP^\beta(a^0, u^\delta, v^\delta))\big)$$

Intuitively, this says that y^β is the collection of codes for ordered pairs of a^0 and members u^δ of x^β.

Why should we believe these axioms? Where β is finite, the answer is straightforward: We can give an explicit analysis of $OP^{\beta+1}(a^0, x^\beta, y^\beta)$ from which the relevant pairing axioms follow and such that the existence axiom is implied by the comprehension scheme for entities of type $\beta+1$. For instance $OP^2(a^0, x^1, y^1)$ can be analyzed as $\forall v^0(y^1(v^0) \leftrightarrow \exists u^0(x^1(u^0) \wedge v^0 = \langle a^0, u^0\rangle))$. If we were willing to adopt infinitary languages, then this analysis could be continued to infinite types as well, thus eliminating the need for the new primitive predicates $OP^{\beta+1}$ and the axioms governing these predicates. Although we won't adopt this strategy, it shows that the axioms adopted above are not *ad hoc* but can be motivated in much the same way as the comprehension axioms described in Section 6.4. Analogous considerations apply to the other primitive predicates adopted below and the axioms governing them.

It will be useful to introduce a notational abbreviation. For any formula ϕ, we let:

$$\phi(\langle a^0, x^\beta\rangle^\beta) \;\equiv_{df}\; \exists y^\beta(OP^{\beta+1}(a^0, x^\beta, y^\beta) \wedge \phi(x^\beta))$$

One can then prove the following by induction on β:

Lemma 1 $\langle a^0, x^\beta\rangle^\beta = \langle b^0, y^\beta\rangle^\beta \leftrightarrow (a^0 = b^0 \wedge x^\beta = y^\beta)$

[40] In fact, for the special case in which $\beta = \theta + 1$, the pairing axiom corresponding to $\delta = \theta$ is all we need. The same goes for the axioms concerning $\text{Union}^{\beta+1}$, $\text{Proj}^{\beta+1}$, and $\text{Trans}^{\gamma+1}$ below.

Once we have pairs of the form $\langle a^0, x^\beta\rangle^\beta$ in place, it is possible to introduce pairs of the form $\langle x^\beta, y^\beta\rangle^\beta$. The intuitive idea is that we let $\langle x^\beta, y^\beta\rangle^\beta = \langle 0, x^\beta\rangle^\beta \cup \langle 1, y^\beta\rangle^\beta$. Formally, for each β between 0 and α, we introduce the following notational abbreviation:

$$\phi(\langle x^\beta, y^\beta\rangle^\beta) \ \equiv_{df} \ \exists z^\beta(\text{Union}^{\beta+1}(\langle 0, x^\beta\rangle^\beta, \langle 1, y^\beta\rangle^\beta, z^\beta) \wedge \phi(z^\beta))$$

Here 0 and 1 are arbitrary individuals of type 0 such that $0 \neq 1$. Union$^{\beta+1}$ is a new atomic predicate, characterized axiomatically. For each β between 0 and α, we set forth an *existence* axiom for Union$^{\beta+1}$:

$$\forall x^\beta \forall y^\beta \exists z^\beta \ \text{Union}^{\beta+1}(x^\beta, y^\beta, z^\beta)$$

and a family of *union* axioms for Union$^{\beta+1}$, one for each $\delta < \beta$:

$$\text{Union}^{\beta+1}(x^\beta, y^\beta, z^\beta) \rightarrow \forall w^\delta(z^\beta(w^\delta) \leftrightarrow (x^\beta(w^\delta) \vee y^\beta(w^\delta)))$$

Next, it will be useful to introduce a family of *projection* functions Proj$^{\beta+1}(a^0, y^\beta, x^\beta)$ ($\beta < \alpha$). Intuitively, Proj$^{\beta+1}(a^0, y^\beta, x^\beta)$ says that $y^\beta = \langle a^0, x^\beta\rangle^\beta \cup z^\beta$, for some z^β whose transitive closure includes no pairs of the form $\langle a^0, x^0\rangle^0$. Formally, we give an axiomatic characterization of the projection functions. For each β between 0 and α, we set forth an *existence* axiom for Proj$^{\beta+1}$:

$$\forall a^0 \forall y^\beta \exists x^\beta \ (\text{Proj}^{\beta+1}(a^0, y^\beta, x^\beta))$$

and a family of *projection* axioms for Proj$^{\beta+1}$, one for each $\delta < \beta$:

$$\text{Proj}^{\beta+1}(a^0, y^\beta, x^\beta) \rightarrow \forall z^\delta(x^\beta(z^\delta) \leftrightarrow y^\beta(\langle a^0, z^\delta\rangle^\delta))$$

This allows us to introduce a useful notational abbreviation. For any formula ϕ,

$$\phi(\pi^\beta_{a^0}(y^\beta)) \ \equiv_{df} \ \exists x^\beta(\text{Proj}^{\beta+1}(a^0, y^\beta, x^\beta) \wedge \phi(x^\beta))$$

By induction on β, it is easy to show that $\pi^{\beta}_{a^0}(\langle a^0, x^{\beta}\rangle^{\beta}) = x^{\beta}$. Similarly, one can show that if z^{β} is the union of $\langle a^0, x^{\beta}\rangle^{\beta}$ and $\langle b^0, y^{\beta}\rangle^{\beta}$, then $\pi^{\beta}_{a^0}(z^{\beta}) = x^{\beta}$ and $\pi^{\beta}_{b^0}(z^{\beta}) = y^{\beta}$. This makes it straightforward to prove:

Lemma 2 $\langle x^{\beta}, y^{\beta}\rangle^{\beta} = \langle z^{\beta}, w^{\beta}\rangle^{\beta} \leftrightarrow (x^{\beta} = z^{\beta} \wedge y^{\beta} = w^{\beta})$

As usual, ordered n-tuples for finite n can be characterized on the basis of ordered pairs. For $n > 2$, we set forth the following abbreviation:

$$\langle x^{\beta}_1, \ldots, \ldots, x^{\beta}_n\rangle^{\beta} \ \equiv_{df} \ \langle x^{\beta}_1, \langle x^{\beta}_2, \ldots, x^{\beta}_n\rangle^{\beta}\rangle^{\beta}.$$

Theorem 1 is now immediate.

B.3 Assignments and models

We will now characterize primitive predicates $\textsc{Mod}^{\alpha+1}(M^{\alpha})$ and $\textsc{Asgn}^{\beta+1}(A^{\beta})$ corresponding to the notions of model and assignment, as well as the interpretation operations $[\![t]\!]^{\alpha}_{M^{\alpha}}$ and $[\![t]\!]^{\beta}_{A^{\beta}}$.

Intuitively, a model for $\mathcal{L}^{\alpha}$ is an entity M^{α} of type α which codes a domain and an interpretation for each constant c^{γ}_i of $\mathcal{L}^{\alpha}$ ($\gamma \leq \alpha$). It will be defined in such a way as to allow for the following decodings:

1. The *domain* of the model will be coded by a level-1 entity $D^1_{M^{\alpha}}$, which will be defined by way of the projection function, as follows:

$$D^1_{M^{\alpha}} = \pi^{\alpha}_{`\forall'}(M^{\alpha}).$$

2. The *interpretation* of a constant c^{γ}_i of $\mathcal{L}^{\alpha}$ ($\gamma \leq \alpha$) will be coded by a level-α entity $[\![c^{\gamma}_i]\!]^{\alpha}_{M^{\alpha}}$, which will be defined by way of the projection function, as follows:

$$[\![c^{\gamma}_i]\!]^{\alpha}_{M^{\alpha}} = \pi^{\alpha}_{c^{\gamma}_i}(M^{\alpha}).$$

Four requirements need to be imposed. Firstly, the domain $D^1_{M^\alpha}$ of M^α must be non-empty. Secondly, M^α must represent the interpretation of a constant c^γ_i as being based entirely on individuals from $D^1_{M^\alpha}$ (more precisely: whenever z^0 is in the 'transitive closure' of $[\![c^\gamma_i]\!]^\alpha_{M^\alpha}$, z^0 must be in D^1_M). Thirdly, we need to insist that $[\![c^\gamma_i]\!]^\alpha_{M^\alpha}(z^\delta)$ fail to be the case whenever $\delta \geq \gamma$. Finally, there must be no 'empty' constants of type 0.

We do this formally by way of the primitive predicate $\textsc{Mod}^{\alpha+1}(M^\alpha)$, which is characterized axiomatically. We begin with an existence axiom: $\exists x^\alpha\ \textsc{Mod}^{\alpha+1}(x^\alpha)$. Next, we introduce an axiom to guarantee that the domain of a model is non-empty:

$$\textsc{Mod}^{\alpha+1}(x^\alpha) \rightarrow \exists z^0\ D^1_{x^\alpha}(z^0)$$

We now need to capture the fact that the transitive closure of $[\![c^\gamma_i]\!]^\alpha_{M^\alpha}$ consists entirely of individuals in $D^1_{M^\alpha}$. But to do so it is necessary to characterize transitive closure. We proceed by introducing a family of new primitive predicates $\textsc{Trans}^{\gamma+1}(x^\gamma, y^0)$ $(\gamma \leq \alpha)$. Intuitively, $\textsc{Trans}^{\gamma+1}(x^\gamma, y^0)$ expresses the thought that y^0 is in the transitive closure of x^γ. Formally, $\textsc{Trans}^{\gamma+1}$ is characterized axiomatically. For $\gamma = 0$, we set forth a single axiom:

$$\textsc{Trans}^1(x^0, y^0) \leftrightarrow x^0 = y^0.$$

For $0 < \gamma \leq \alpha$, we set forth an axiom for each $\delta < \gamma$:

$$\textsc{Trans}^{\gamma+1}(x^\gamma, y^0) \leftrightarrow \exists z^\delta (x^\gamma(z^\delta) \wedge \textsc{Trans}^\gamma(z^\delta, y^0)).$$

With the notion of transitive closure in place, we may introduce the following axiom to capture the fact that the transitive closure of $[\![c^\gamma_i]\!]^\alpha_{M^\alpha}$ consists entirely of individuals in $D^1_{M^\alpha}$:

$$\textsc{Mod}^{\alpha+1}(x^\alpha) \rightarrow \forall z^0(\textsc{Trans}^{\alpha+1}([\![c^\gamma_i]\!]^\alpha_{x^\alpha}, z^0) \rightarrow D^1_{x^\alpha}(z^0))$$

Next, we need to ensure that $[\![c^\gamma_i]\!]^\alpha_{M^\alpha}(z^\delta)$ fails to be the case whenever $\delta \geq \gamma$. This requires an axiom for each δ such that $\gamma \leq \delta < \alpha$:

$$\textsc{Mod}^{\alpha+1}(x^{\alpha}) \rightarrow \forall z^{\delta}(\neg [\![c_i^{\gamma}]\!]^{\alpha}_{x^{\alpha}}(z^{\delta}))$$

Finally, we need an axiom to ensure that there are no 'empty' constants of type 0:

$$\textsc{Mod}^{\alpha+1}(x^{\alpha}) \rightarrow \exists z^{0}([\![c_i^{0}]\!]^{\alpha}_{x^{\alpha}} = z^{0})$$

This completes our characterization of $\textsc{Mod}^{\alpha+1}(x^{\alpha})$.

The next step is to characterize the notion of a variable assignment. Intuitively, an assignment for $\mathcal{L}^{\alpha}$ is an entity A^{β} of type β (for $\alpha = \beta + 1$) which codes an interpretation for each variable x_i^{γ} of $\mathcal{L}^{\alpha}$ ($\gamma < \alpha$). As in the case of models, it will be defined in such a way as to allow for the following decoding:

$$[\![x_i^{\gamma}]\!]^{\beta}_{A^{\beta}} = \pi^{\beta}_{x_i^{\gamma}}(A^{\beta}).$$

Two requirements need to be imposed. Firstly, we need to insist that $[\![x_i^{\gamma}]\!]^{\beta}_{A^{\beta}}(z^{\delta})$ fail to be the case whenever $\delta \geq \gamma$. Secondly, every variable of type 0 must receive an assignment.

Formally, this is done by way of the primitive predicate $\textsc{Asgn}^{\beta+1}(A^{\beta})$, which is characterized axiomatically. We begin with an existence axiom:

$$\exists y^{\beta}\ \textsc{Asgn}^{\beta+1}(y^{\beta})$$

Next, we need to ensure that $[\![x_i^{\gamma}]\!]^{\beta}_{A^{\beta}}(z^{\delta})$ fails to be the case whenever $\delta \geq \gamma$. This requires an axiom for each δ greater or equal to γ but below β:

$$\textsc{Asgn}^{\beta+1}(y^{\beta}) \rightarrow \forall z^{\delta}(\neg [\![x_i^{\gamma}]\!]^{\beta}_{y^{\beta}}(z^{\delta}))$$

Finally, we need an axiom to ensure that every variable of type 0 receives an assignment:

$$\textsc{Asgn}^{\beta+1}(y^{\beta}) \rightarrow \exists z^{0}([\![x_i^{0}]\!]^{\beta}_{y^{\beta}} = z^{0})$$

References

Benacerraf, P. and Putnam, H., editors (1983). *Philosophy of Mathematics: Selected Readings*, Cambridge. Cambridge University Press. Second edition.

Boolos, G. (1984). To Be is to Be a Value of a Variable (or to Be Some Values of Some Variables). *Journal of Philosophy*, 81(8):430–449. Reprinted in (Boolos, 1998).

Boolos, G. (1985). Nominalist Platonism. *Philosophical Review*, 94(3):327–344. Reprinted in (Boolos, 1998).

Boolos, G. (1998). *Logic, Logic, and Logic.* Harvard University Press, Cambridge, MA.

Carnap, R. (1934). *Logische Syntax der Sprache.* Springer, Wien. Translated as *The Logical Syntax of Language* (New York: Harcourt, Brace and Co, 1937).

Dedekind, R. (1888). *Was Sind und Was Sollen die Zahlen?* Vieweg, Braunschweig. English translation in (Ewald, 1996).

Degen, W. and Johannsen, J. (2000). Cumulative higher-order logic as a foundation for set theory. *Mathematical Logic Quarterly*, 46(2):147–170.

Ewald, W. (1996). *From Kant to Hilbert: A Source Book in the Foundations of Mathematics*, volume 2. Oxford University Press, Oxford.

Feferman, S. (1995). Introductory note to **1933o*. In (Gödel, 1995).

Frege, G. (1879). Begriffsschrift: Eine der arithmetischen nachgebildete Formelsprache des reinen Denkens. Translated and reprinted in (van Heijenoort, 1967).

Gödel, K. (1931). Über formal unentscheidbare Sätze der *Principia Mathematica* und verwandter Systeme, I. *Monatshefte für Mathematik und Physik*, 38:173–198.

Gödel, K. (1933). The present situation in the foundations of mathematics. In (Gödel, 1995).

Gödel, K. (1944). Russell's Mathematical Logic. In (Benacerraf and Putnam, 1983).

Gödel, K. (1995). *Collected Works*, volume III. Oxford University Press, Oxford.

Hilbert, D. (1926). Über das Unendliche. *Mathematische Annalen*, 95:161–190. Translated as "On the Infinite" in (van Heijenoort, 1967).

Lewis, D. (1991). *Parts of Classes*. Blackwell, Oxford.

Linnebo, Ø. Pluralities and Sets. Forthcoming in *Journal of Philosophy*.

Linnebo, Ø. (2003). Plural Quantification Exposed. *Noûs*, 37(1):71–92.

Linnebo, Ø. and Nicolas, D. (2008). Superplurals in English. *Analysis*, 68(3):186–197.

McGee, V. (1997). How we learn mathematical language. *Philosophical Review*, 106:35–68.

Parsons, C. (1977). What Is the Iterative Conception of Set? In Butts, R. and Hintikka, J., editors, *Logic, Foundations of Mathematics, and Computability Theory*. Reidel, Dordrecht. Reprinted in (Benacerraf and Putnam, 1983) and (Parsons, 1983).

Parsons, C. (1983). *Mathematics in Philosophy*. Cornell University Press, Ithaca, NY.

Quine, W. V. (1986). *Philosophy of Logic, Second Edition*. Harvard, Cambridge, Massachusetts.

Rayo, A. (2006). Beyond Plurals. In Rayo, A. and Uzquiano, G., editors, *Unrestricted Quantification: New Essays*. Oxford.

Rayo, A. and Uzquiano, G. (1999). Toward a Theory of Second-Order Consequence. *Notre Dame Journal of Formal Logic*, 40(3):315–25.

Rayo, A. and Yablo, S. (2001). Nominalism through De-Nominalization. *Noûs*, 35(1):74–92.

Shapiro, S. (1987). Principles of reflection and second-order logic. *Journal of Philosophical Logic*, 16:309–333.

Shapiro, S. (2000). *Foundations without Foundationalism: A Case for Second-Order Logic*. Oxford University Press, Oxford.

Tarski, A. (1935). Der Wahrheitsbegriff in den formalisierten Sprachen. *Studio Philosophica*, 1:261–405. English translation in (Tarski, 1983), pp. 152-278.

Tarski, A. (1983). *Logic, Semantics, Metamathematics.* Oxford University Press, Oxford, second edition.

van Heijenoort, J., editor (1967). *From Frege to Gödel*, Cambridge, MA. Harvard University Press.

Wang, H. (1952). Negative types. *Mind*, 61:366–8.

Williamson, T. (2003). Everything. In Hawthorne, J. and Zimmerman, D., editors, *Philosophical Perspectives 17: Language and Philosophical Linguistics.* Blackwell, Boston and Oxford.

Wright, C. (2007). On quantifying into predicate position. In M. Leng, A. P. and Potter, M., editors, *Mathematical Knowledge*, pages 150–74. Oxford University Press, Oxford.

Yablo, S. (2006). Circularity and paradox. In T. Bolander, V. H. and Pedersen, S., editors, *Self-Reference*. CSLI Press.

Zermelo, E. (1930). Über Grenzzahlen und Mengenbereiche. *Fundamenta Mathematicae*, 16:29–47. Translated in (Ewald, 1996).