4
Plurals and Set Theory

What is the relation between some things and their set? This is a hard question which has confounded many brilliant minds. We recall, for example, that Russell wrestled with the question:

> Is a class which has many terms to be regarded as itself one or many? Taking the class as equivalent simply to the numerical conjunction "*A* and *B* and *C* and etc.," it seems plain that it is many; yet it is quite necessary that we should be able to count classes as one each, and we do habitually speak of a class. Thus classes would seem to be one in one sense and many in another.
>
> (Russell 1903, Section 74)

We begin with a formal comparison between plural logic and set theory, which clarifies an important technical aspect of the question. After that, we address some philosophical issues concerning the relation between some things and their set. Our discussion yields an argument for primitive plurals, which we believe has more force than any of the arguments discussed in the previous chapter. More specifically, we argue that the expressive resources of plurals are needed to account for sets.

4.1 A simple two-sorted set theory

Assume we start with a singular first-order language whose quantifiers range over certain objects. Let us refer to these objects as *individuals*. We are interested in ways to talk simultaneously about many individuals.

The most familiar option, at least to anyone with some training in mathematics, is to use set theory. A set is a *single object* that has zero or more *elements*. Talking about a single set thus provides a way to talk about all of its elements simultaneously. For example, we can convey information about two individuals, say Russell and Whitehead, by talking about their

The Many and the One: A Philosophical Study of Plural Logic. Salvatore Florio and Øystein Linnebo, Oxford University Press.
DOI: 10.1093/oso/9780198791522.003.0004

set {Russell, Whitehead}. The information that they are philosophers can be conveyed by saying that every element of the set is a philosopher. Similarly, we can convey information about the natural numbers by talking about their set. The information that they are infinitely many can be conveyed by saying that their set is infinite. Suppose, more generally, that we want to talk about some objects. According to the present strategy, we can achieve this by talking about an associated set.

It is not obvious, however, that such a set exists. After all, the lesson of the set-theoretic paradoxes is that not every formula defines a set. The most famous example is Russell's paradox of the set of all sets that are not elements of themselves. Consider the formula that serves as a condition for membership in this would-be set: $x \notin x$. Suppose this formula defines a set R. Now ask: is R an element of itself? The answer is affirmative if and only if R satisfies the membership condition. In other words: $R \in R$ if and only if $R \notin R$. But this is a contradiction!

Thankfully, the problem posed by the set-theoretic paradoxes can be put off, at least for a while. The paradoxes do not arise when we consider only sets of individuals drawn from a fixed first-order domain. And for present purposes, this is all we need. So let us consider a very simple set theory, which satisfies our present needs but does not give rise to paradoxes.

We distinguish between individuals and sets of individuals. To do so, it is convenient to use a *two-sorted* language. Such languages are easily explained because they are implicit in various mathematical practices. For example, in geometry we often use one set of variables to range over points (say, $p_1, p_2, \ldots$) and another set of variables to range over lines (say, $l_1, l_2, \ldots$). We adopt a similar approach in our simple set theory, letting lower-case variables range over individuals ($x, y, \ldots$) and upper-case variables ($X, Y, \ldots$) range over sets of individuals. We refer to these as *individual variables* and *set variables*, respectively. If desired, we can of course add constants of either sort. There are sortal restrictions on the formation rules. For instance, the language has a membership predicate '$\in$' whose first argument can only be an individual term and whose second argument can only be a set term. Thus, '$a \in X$' means that the individual a is an element of the set X. In addition to the ordinary identity predicate, which can be flanked by any two individual terms, our extended language contains a set identity predicate, which can be flanked only by set terms. For convenience, we use the standard identity sign for both identity predicates. Given the restrictions just mentioned, it is impermissible to make identity claims involving both an individual and a set term (such as '$a = X$').

This two-sorted language, which we call $\mathcal{L}_{\text{SST}}$, will be the language of our simple set theory, SST. Let $\mathcal{L}^{+}_{\text{SST}}$ be the extended language obtained by adding predicates that take set terms as arguments. This is an optional extra to which we will return.

We now formulate SST based on the axioms and rules of two-sorted classical logic. First, we adopt the axiom of extensionality for sets:

(S-Ext) $$\forall x(x \in X \leftrightarrow x \in Y) \rightarrow X = Y$$

Then, we adopt an axiom scheme of set comprehension:

(S-Comp) $$\exists X \forall x(x \in X \leftrightarrow \varphi(x))$$

where X does not occur free in $\varphi(x)$. The theory SST+ is obtained by adapting the rules and axioms of SST to the richer language $\mathcal{L}^{+}_{\text{SST}}$.

Notice how Russell's paradox is blocked by the use of separate sorts for individuals and their sets. In our two-sorted language, the membership condition for the offending set, namely $x \notin x$ is not even well formed.

4.2 Plural logic and the simple set theory compared

Let us compare how plural logic and the simple set theory talk about the many. Consider a domain of individuals to which both systems are applicable. (We will later address the important question of what, exactly, the conditions are under which each system is applicable.) Suppose we wish to talk about many individuals simultaneously. As we will now show, these two ways to talk about the many share a common structure.

The two languages share a common stock of variables x_i that take as their values one individual at a time. And each language has an additional stock of variables that are used to convey information about (loosely speaking) collections of individuals: plural variables xx_i, which take as their values many individuals simultaneously, or set variables X_i, which take as their values a single set of individuals. In addition, each language has a predicate for membership in a collection: $x_i \prec xx_j$ for "x_i is one of xx_j" or $x_i \in X_j$ for "x_i is an element of X_j".

This suggests that it should be straightforward to translate back and forth between the two languages. One can simply replace $\prec$ with $\in$ and xx_j with

X_j, and *vice versa*. In fact, things are *nearly* that simple. There are just two wrinkles to be ironed out:

- $\mathcal{L}_{\text{SST}}$ has an identity predicate that can be flanked by set terms, whereas $\mathcal{L}_{\text{PFO}}$ has no identity predicate that can be flanked by plural terms.
- SST postulates an empty set, whereas PFO has an axiom stating that every plurality is non-empty.

Fortunately, both problems are easily overcome. In Appendix 4.A, we show how to define a translation from each language to the other such that each sentence and its translation convey the same information, at least as far as the individuals are concerned, only that one sentence does so by utilizing plural resources, while the other uses set-theoretic resources.

As we explain in Appendices 4.A and 4.B, the translations satisfy the following important conditions:

(i) each translation is recursive, that is, there is an effective algorithm for carrying out the translation;
(ii) each translation commutes with the logical connectives (for example, the translation of a negation $\neg\varphi$ is the negation of the translation of φ);
(iii) every theorem of each of the two theories is translated as a theorem of the other theory (for example, every theorem of PFO is translated as a theorem of SST).

More generally, let τ be a translation from the language of one theory T_1 to that of another theory T_2 such that these three conditions are satisfied. Then, τ is said to be an *interpretation* of T_1 in T_2. Thus, what we show in the appendices is that each of our two theories PFO and SST can be interpreted in the other, and likewise with PFO+ and SST+.

It is important to be absolutely clear about what the mutual interpretability of two theories does and does not establish. Interpretability is a purely formal notion, which also allows us to recursively turn a model of one theory into a model of another. Thus, two mutually interpretable theories are equivalent for the purposes of formal logic. However, there is no guarantee that the equivalence will extend beyond those purposes.

To see this, suppose the two languages are meaningful. Then, there is no guarantee that the translation preserves the kinds of extra-logical properties that philosophers commonly discuss. For example, the translation need not preserve features of sentences such as:

- truth value;
- meaning (perhaps understood as the set of possible worlds at which a sentence is true);
- epistemic status (for example, a priori or a posteriori);
- ontological commitments.

It is often controversial whether a translation preserves these features. The translations presented in this chapter are no exception. Consider a nominalist who accepts a certain plural sentence but rejects its set-theoretic translation. This provides a perspective from which the translation does not preserve truth value and hence meaning.

4.3 Plural logic vs. set theory: classifying the options

What is the significance of the shared structure (or mutual interpretability) that we just observed? Is this *merely* a technical result? Or does the technical result have some broader philosophical significance?

When the structure of one theory can be recovered within that of another, this raises the question of whether one of the theories can be eliminated in favor of the other. In the present context, there are three options. First, one may eliminate pluralities in favor of sets. Second, one may proceed in the opposite direction and eliminate sets in favor of pluralities. Finally, one may refrain from any elimination and retain both pluralities and sets. All three options have their defenders.

First, some philosophers hold that the plural locutions found in English and many other natural languages should be eliminated in favor of talk about sets. We mentioned in Chapter 2 that Quine is an advocate of this view; see also Resnik 1988. For Quine at least, this is at root a claim about regimentation into our scientific language. It is indisputable that many natural languages contain plural locutions. But our best scientific theory of the world has no need for such locutions. This theory is to be formulated in a singular language—that is, a language lacking plural resources—whose quantifiers also range over sets. When regimenting natural language into this scientific language, the plural locutions of the former should be analyzed by means of the set talk of the latter. In short, for scientific purposes, we should eschew plural resources and instead rely on set-theoretic resources. These resources also suffice to interpret “the vulgar” (as Quine once put it), that is, to regiment the plural resources indisputably found in English and other natural languages.

Second, other philosophers insist that sets should be eliminated in favor of pluralities. That is, we can and should interpret ordinary set talk without relying on set-theoretic resources ourselves. A classic paper by Black (1971) can be read as advocating this view.[1] More recently, Oliver and Smiley have expressed considerable sympathy for the view, claiming to have at least shifted the burden of proof onto its opponents (2016, 316–17).

Lastly, one may hold that neither system should be eliminated in favor of the other, because both plural logic and set theory are legitimate and earn their keep in our best scientific theory. Following Cantor and Gödel, this is the view that we will defend. Suppose we are right that both systems should be retained. Then a host of questions arise concerning their relation. We will be particularly concerned with two such questions.

(a) Every non-empty set obviously corresponds to a plurality, namely the elements of the set. What about the other direction? Does every plurality correspond to a set? If not, under what conditions do some things form a set?[2]
(b) Suppose that some objects form a set. Can these objects be used to shed light on, or give an account of, the set that they form?

Before addressing these two questions, let us explain why we reject both the elimination of pluralities in favor of sets and that of sets in favor of pluralities.

4.4 Against the elimination of pluralities in favor of sets

One reason against the elimination of pluralities in favor of sets is the paradox of plurality, discussed in Section 3.4. The paradox arises in untyped approaches to sets, where sets are regarded as objects alongside others. Ordinary set theory is such an approach, unlike our system SST. The argument

[1] Rafal Urbaniak (2013) has argued that Leśniewski can be read in the same way. This reading is disputed by Oliver and Smiley who take Leśniewski to be “an orthodox singularist” about plurals (2016, 15).

[2] See Hewitt 2015 for a useful overview of this issue.

begins, we recall, by observing that the following sentence seems trivially true:

(4.1) There are some objects such that any object is one of them if and only if that object is not an element of itself.

Suppose that plural resources are to be eliminated in favor of set-theoretic ones. Then it is natural to regiment (4.1) as follows:

(4.2) There is a set of which any object is an element if and only if that object is not an element of itself.

In symbols:

(4.3) $\exists x\big(\text{set}(x) \wedge \forall y(y \in x \leftrightarrow y \notin y)\big)$

This, of course, is an instance of the familiar Russell sentence, which is inconsistent.

While this is a powerful argument, we saw that several responses are possible. Quine might try to dismiss the plural talk about sets in (4.1) as just confused talk about sets in two different guises and as having no place in the ideal language of science. This is a logically coherent view for him to take. However, this blunt dismissal of "the vulgar" is ultimately hard to sustain. We find it difficult to deny that English speakers do understand plural talk about sets. A charitable interpretation of "the vulgar" should not deny this fact.

A more promising option is to deny the possibility of absolutely general quantification. If absolute generality is unattainable, then the door is open to claiming that (4.2) is true but that the witness to the existence claim lies beyond the range of the embedded universal quantifier ('$\forall y$' in the formalization), with the result that paradox is averted.[3]

Yet another option is to deny that (4.1) is true. Of course, (4.1) is just an instance of plural comprehension. But perhaps plural comprehension isn't always permissible! Any plurality must presumably be properly circumscribed. So when we are reasoning about a domain that cannot be circumscribed—such as the domain of absolutely everything—not every condition can be used to define a plurality.

[3] See discussion in Section 3.4.

We won't attempt to resolve the matter here. The last two responses to the paradox of plurality raise big questions that we discuss in the final two chapters. Instead, we wish to lay out another—and, we believe, more compelling—reason why pluralities should not be eliminated in favor of sets. The reason is simply that pluralities are needed to give an account of sets.[4] So if pluralities were eliminated in favor of sets, we could not use plural reasoning to give such an account. In sum, to retain an attractive account of sets in terms of pluralities, we cannot eliminate plurals.

What is the promised account of sets in terms of pluralities? It is useful to recall how Cantor, the father of modern set theory, sought to explain the concept of set.

> By a 'manifold' or 'set' I understand in general every many which can be thought of as one, i.e. every totality of determinate elements which can be bound together into a whole through a law [. . .].
>
> (Cantor 1883, 43; our translation)[5]

That is, a set is a "many thought of as one". Of course, it is far from clear how this is to be understood. (An explication will be proposed shortly.) But there can be no doubt that Cantor sought to understand a set in terms of the many objects that are its elements and that are somehow "thought of as one".

> By a 'set' we understand every collection into a whole M of determinate, well-distinguished objects m of our intuition or our thought (which will be called the 'elements' of M). We write this as: $M = \{m\}$.
>
> (Cantor 1895, 481; our translation)[6]

It is tempting to read Cantor's variable 'm' as a plural variable (see also Oliver and Smiley 2016, 4–5). So, in line with our notation, let us replace

[4] One might attempt to deny the need for such an account by adopting a more structuralist conception of set, where a set is characterized in terms of its structural relations to all other sets rather than in terms of some particularly intimate relation to its elements. See Parsons 2008, Chapter 4, for useful discussion. However, we insist that there is also a more ontological conception of set, especially in the case of hereditarily finite sets, which regards a set as "constituted" by its elements. In fact, such a conception is suggested by a liberal view of definitions, to be described shortly.

[5] The original reads: 'Unter einer Mannichfaltigkeit oder Menge verstehe ich nämlich allgemein jedes Viele, welches sich als Eines denken lässt, d.h. jeden Inbegriff bestimmter Elemente, welcher durch ein Gesetz zu einem Ganzen verbunden werden kann [. . .].'

[6] The original reads: 'Unter einer Menge verstehen wir jede Zusammenfassung M von bestimmten wohlunterschiedenen Objekten m unsrer Anschauung oder unseres Denkens (welche die Elemente von M genannt werden) zu einem Ganzen. In Zeichen drücken wir dies so aus: $M = \{m\}$.'

this variable with '*mm*'. A set M is then said to be a collection into one of some well-distinguished objects *mm*, namely the elements of M. And we write $M = \{mm\}$.

What about the empty set? Here there is a threat of a mismatch. While standard set theory accepts an empty set, traditional plural logic does not accept an empty plurality. But we are confident that this threat can be addressed. One option is to break with traditional plural logic and accept an empty plurality, perhaps on the grounds that, although this isn't how plurals work in English and many other natural languages, there are coherent languages where plurals do behave in this way (see Burgess and Rosen 1997, 154–5). Another option is to break with standard set theory and abandon the empty set. However, we would prefer not to deviate from successful scientific practice, in this case set theory, unless there are compelling reasons to do so. Finally, an elegant option proposed (in a different context) by Oliver and Smiley (2016, 88) is to allow "co-partial functions", that is, functions that can have a value even where the argument is undefined. Suppose the 'set of' operation $xx \mapsto \{xx\}$ is such a function. Then, applied to an undefined argument, this function can have the empty set as its value.

What is it for many objects to be "thought of as one" or collected "into a whole"? Let us attempt to shed some light on this idea. Many philosophers and mathematicians believe that the elements of a set are somehow "prior to" the set itself and that the set is somehow "constituted" by its elements.[7] Assume *xx* form a set $\{xx\}$. Then the objects *xx* can be used to give an account of $\{xx\}$. That is, properties and relations involving the set are explained in terms of properties and relations involving the plurality of its elements. Why is a an element of $\{xx\}$? An answer immediately suggests itself: because a is one of *xx*. Why is $\{xx\}$ identical with $\{yy\}$? Again, the answer seems obvious: because *xx* are the very same objects as *yy*.[8]

Of course, in their current form, these remarks are highly programmatic. The promised account needs to be spelled out. We do this in Section 12.3 by defending a liberal view of definitions. Here is the rough idea behind the view: it suffices for a mathematical object to exist that an adequate definition of it can be provided. The adequacy in question is understood as follows. Suppose we begin with a "properly circumscribed" domain of

[7] See, e.g., Parsons 1977a and Fine 1991.

[8] This account contrasts with some earlier contributions to the metaphysics of sets, e.g. Lewis 1991 and Oliver and Smiley 2016 (Chapter 14). We believe our account coheres better with the remarks by Cantor (discussed above) and by Gödel (discussed in Section 4.6).

objects standing in certain relations.[9] We would like to define one or more additional objects. Suppose our definition determines the truth of any atomic statement concerned with the desired "new" objects by means of some statement concerned solely with the "old" objects with which we began. Then, according to the liberal view, the definition is permissible.

To illustrate the point, let us apply the view to the case of sets. Suppose we begin with some properly circumscribed domain of objects. For every plurality of objects xx from this domain, we postulate their set $\{xx\}$, with the understanding that atomic statements concerned with any new sets should be assessed in the following way.

(i) $\{xx\} = \{yy\}$ if and only if $xx \approx yy$.
(ii) $a \in \{xx\}$ if and only if $a \prec xx$.

Notice how this account determines the truth of any atomic statement concerned with the "new" sets solely in terms of the "old" objects with which we began, as required by the liberal view.[10]

We also observe that this account distinguishes a set from its singleton, as is customary in comtemporary set theory. By (i), we have $\{xx\} = \{\{xx\}\}$ just in case xx is coextensive with $\{xx\}$. We contend that this coextensionality claim is false. Suppose xx are two or more in number. Then cardinality considerations alone ensure its falsity. Alternatively, suppose xx consist of a single object a. Then the coextensionality claim is equivalent to $a = \{a\}$, which is false because a is an element of its own singleton but not, we may suppose, of itself.

To sum up, we argue that pluralities should be retained alongside sets, so that the former can be used to shed light on the latter. This account of sets draws essentially on our liberal view of definitions.

4.5 Against the elimination of sets in favor of pluralities

The view discussed in the preceding section retains sets but gives an account of them in terms of pluralities. One may wonder whether a more radical approach is possible. Why not simply eliminate sets in favor of pluralities?

[9] In Part IV of the book, the notion of being properly circumscribed will play an important role and will be analyzed under the label of *extensional definiteness*.
[10] Mereological sums provide another example of this liberal view of definitions; see Section 5.8.

Black's (1971) classic discussion suggests a view of this sort.[11] He observes that ordinary language often talks about sets: expressions such as 'my set of chessmen' or 'that set of books' feel fairly natural to English speakers. By reflecting on ordinary uses of the word 'set', he argues, we can come to see the intimate connection between talk about a set and about its elements. More specifically, we can come to realize that basic uses of the word 'set' are simply substitutes for plural expressions such as plural descriptions or lists of terms. In his example, the sentence 'a certain set of men is running for office' is what he calls an "indefinite surrogate" for the statement that, say, Tom, Dick, and Harry are running for office (Black 1971, 631).

Black recognizes that there is a gap between ordinary uses of the word 'set' and its uses in mathematics. For instance, ordinary speakers untrained in abstract mathematics often have misgivings about the empty set. If sets are collections of things, how can there be a collection of nothing whatsoever? Despite such misgivings, Black contends that we can rely on our ordinary understanding of plurals to make sense of "idealized" uses of the word 'set' as it occurs in mathematics.

There is an obvious difficulty for Black's contention. Talk of *sets of sets* is ubiquitous in mathematics and, as we will see shortly, such "nested" sets are essential to the now-dominant iterative conception of set. How can we account for these uses of the word 'set'? If talk about sets is shorthand for talk about pluralities, then sets of sets would seem to correspond to higher-level pluralities, that is, "pluralities of pluralities".[12]

It is controversial whether such higher-level pluralities make sense, but a putative example is given in following sentence.

(4.4) My children, your children, and her children competed against each other.

The subject of this sentence appears to be a "nested" plural, that is, a plural expression formed by combining three other plural expressions. Arguably, this nesting of the subject is semantically significant. The claim is not merely that all the children in question compete against each other but that they do so in teams, each team comprising the children of each parent. We return to the question of whether there are higher-level pluralities in Chapter 9.

[11] We should note that this is not the only way to read Black. It is not entirely clear whether he proposes an eliminative reduction or favors some form of non-eliminative reductionism. An eliminative proposal is developed by Hossack (2000), who appeals to plurals and plural properties to eliminate sets.

[12] For proposals along these lines, see Simons 2016 and Oliver and Smiley 2016, Section 15.1.

While the availability of higher-level pluralities is a necessary condition for the envisaged elimination of sets, it is not sufficient. As observed, the language of mathematics talks extensively about sets and appears to treat these as objects. If possible, it would be good to take this language at face value. The account of sets in terms of pluralities outlined in the previous section allows us to do just that. This provides a reason to retain sets even if higher-level pluralities are available. The reason is even stronger for those who accept other mathematical objects such as numbers. If numbers are accepted, why not also accept sets?

4.6 The iterative conception of set

Suppose we retain both pluralities and sets, giving up on any attempt to eliminate one in favor of the other. How, then, to account for nested sets? This means going beyond the simple set theory discussed in Section 4.1 to form a stronger set theory, where the threat of paradox re-emerges. The standard response to this threat is the so-called iterative conception of set. One of the first clear expressions of this conception is given in a famous passage by Gödel.[13]

> The concept of set, however, according to which a set is anything obtainable from the integers (or some other well-defined objects) by iterated application of the operation "set of", and not something obtained by dividing the totality of all existing things into two categories, has never led to any antinomy whatsoever; that is, the perfectly "naive" and uncritical working with this concept of set has so far proved completely self-consistent.
>
> (Gödel 1964, 180)

The passage calls for some explanation. First, Gödel distinguishes the iterative conception of set from a problematic conception based on the idea of "dividing the totality of all existing things into two categories." Consider a condition that any object may or may not satisfy. One might then attempt to use this condition to divide the totality of all objects into two sets: the set of objects that satisfy the condition and the set of those that don't. But this approach to sets is problematic: as we have seen, it gives rise to Russell's paradox.

[13] The passage contains some footnotes, which we elide.

By contrast, the iterative conception starts with the integers or "some other well-defined objects". We are then told to consider iterated applications of the operation "set of". An example will help. Suppose we start, at what we may call stage 0, with two objects, say a and b. The "set of" operation can be applied to any plurality of objects available at stage 0 to form their set. Thus, at stage 1, which results from the application of this operation to the objects available at stage 0, we have the following sets: $\varnothing$, $\{a\}$, $\{b\}$, and $\{a, b\}$. So, at stage 1, we have six objects, namely a and b together with four sets that were not available at stage 0. Now we can apply the "set of" operation again, this time to the objects available at stage 1. This yields sets such as $\{\varnothing, a\}$, $\{\{a\}, \{b\}\}$, and many others. Note that, by this procedure, the objects available at any given stage form a set at the next stage.

There is a more systematic way to describe what takes us from one stage α to the next stage $\alpha + 1$. For any set S, let its powerset, $\wp(S)$, be the set of all subsets of S, that is:

$$\wp(S) = \{x : x \subseteq S\}$$

Suppose the objects available at stage α are the elements of V_α. Then at stage $\alpha + 1$ we form all the subsets of V_α. So, at stage $\alpha + 1$, we have the elements of V_α as well as those of $\wp(V_\alpha)$. In symbols: $V_{\alpha+1} = V_\alpha \cup \wp(V_\alpha)$. Again, we have by this procedure that all the sets available at stage α, taken together, form a set at stage $\alpha + 1$.

In fact, we want to consider really long iterations of the "set of" operation. The first step is to define V_ω as the result of continuing in this way as many times as there are natural numbers. We do this by letting V_ω be the union of all of the collections V_n generated at a finite stage: $V_\omega = \bigcup_{n<\omega} V_n$. More generally, for any limit ordinal λ, we let V_λ be the union of all the collections of sets we have generated: $V_\lambda = \bigcup_{\gamma<\lambda} V_\gamma$.

The *cumulative hierarchy of sets*, V, is the union of all of the V_α. As Gödel observes (in a footnote to the passage quoted above), V isn't a set. There is no stage at which all sets are available to form a universal set. For any stage, there is a later stage containing even more sets. As a result, we ban the universal set and any other set that would lead to paradox.

Of course, this raises the question of the status of the cumulative hierarchy itself, including the question of whether "it" even exists as an object. We will encounter one appealing response to this question in Section 4.8: perhaps we can invoke plurals and simply regard the cumulative hierarchy as *all the sets* that are formed in the construction described above.

4.7 Zermelo-Fraenkel set theory

The iterative conception motivates much of today's standard set theory, Zermelo-Fraenkel set theory with the Axiom of Choice (ZFC), which is adequate for nearly all of ordinary mathematics. This is a theory of pure sets, formulated in a one-sorted language with only one non-logical predicate, '$\in$' for membership. All other set-theoretic notions are defined in terms of this single predicate. The axioms are as follows.

Extensionality: Coextensive sets are identical. That is:

$$\forall u(u \in x \leftrightarrow u \in y) \rightarrow x = y$$

Empty set: There is an empty set. That is:

$$\exists x \forall y\, y \notin x$$

Pairing: Every two objects have a pair set. That is:

$$\forall x \forall y \exists z \forall u(u \in z \leftrightarrow u = x \lor u = y)$$

Union: For every set x, there is a set y whose elements are precisely those objects that are an element of some element of x. That is:

$$\forall x \exists y \forall u[u \in y \leftrightarrow \exists z(u \in z \land z \in x)]$$

Powerset: Every set has a powerset. That is:

$$\forall x \exists y \forall u(u \in y \leftrightarrow u \subseteq x)$$

Infinity: There is an infinite set, that is, a set with $\varnothing$ as an element and such that, whenever y is an element, so too is $y \cup \{y\}$. That is:

$$\exists x[\varnothing \in x \land \forall y(y \in x \rightarrow y \cup \{y\} \in x)]$$

Separation: For any set x and any condition φ, there is a set of precisely those elements of x that satisfy φ. That is:

$$\forall x \exists y \forall u(u \in y \leftrightarrow u \in x \land \varphi)$$[14]

[14] This is an axiom *scheme*, which yields an axiom for each φ. The same goes for Replacement, stated below.

Foundation: Every non-empty set x has an element that is disjoint from x. That is:

$$\forall x(x \neq \varnothing \rightarrow \exists y(y \in x \wedge x \cap y = \varnothing))$$

Replacement: For every set x and functional condition ψ, there is a set of precisely those objects that are borne ψ by some element of x.[15] That is:

$$Func(\psi) \rightarrow \forall x \exists y \forall u[u \in y \leftrightarrow \exists z(z \in x \wedge \psi(z, u))]$$

This axiom is based on a simple and intuitive idea. Consider any set. For each of its elements, choose either to keep this element or to replace it with some other object. Then the resulting collection is also a set.

Choice: Every set x of non-empty disjoint sets has a choice set, that is, a set containing precisely one element of each element of x. An example due to Russell might be useful to understand the Axiom of Choice. Suppose you have infinitely many pairs of shoes. Then it is easy to define a set containing precisely one member of each pair, namely the set of left shoes. What if you have infinitely many pairs of socks where the two members of each pair are indistinguishable? Then we are unable to define a set containing precisely one member of each pair. The Axiom of Choice tells us that such a set exists, irrespective of our ability to define it.

As observed, ZFC is a theory of pure sets. It is easy, however, to modify the theory to make room for urelements, that is, objects that aren't sets. To do so, we first add to the language a predicate S for being a set. Using this predicate, we then formalize the axioms so as to match precisely their informal statements provided above. For example, the axiom of Extensionality is rewritten as:

$$\forall x \forall y[S(x) \wedge S(y) \;\rightarrow\; (\forall u(u \in x \leftrightarrow u \in y) \rightarrow x = y)]$$

The modified system is often known as ZFCU.

The iterative conception motivates many of the axioms of ZFC or ZFCU. The Powerset axiom provides a nice example. Suppose x is available at some stage s. Then the elements of x were available before s. Hence each subset of x is also available at s. Thus, the set of all these subsets is available at the stage immediately after s. We need not here take a stand on precisely which axioms

[15] The condition ψ is *functional* just in case, for every x there is a unique y such that $\psi(x, y)$.

of set theory are motivated by the iterative conception. Gödel appears to have taken the answer to be "all of them"; others disagree.[16]

4.8 Proper classes as pluralities

Let us use the word 'collection' in an informal way for anything that has a membership structure, such as a set, class, plurality, or indeed even a Fregean concept (where the relation between instance and concept is regarded as a membership structure). We often wish to talk about collections that are too large to form sets, such as the entire cumulative hierarchy of sets or all the ordinals. We will now explain the apparent need for such collections, why these are sometimes regarded as problematic, and finally a brilliant proposal due to Boolos, namely that plural logic provides a way to make sense of these collections.

Let us begin with the need for a novel type of collection, in addition to sets. There are several reasons for this need. Boolos mentions two. First, collections are needed to make sense of the cumulative hierarchy V, which is the domain of set theory. For example, we would like to say that V is the subject matter of set theory and that V is well founded.

Second, collections are needed to understand and justify two axiom schemes that are part of ZFC, namely Replacement and Separation.[17] Both of these take the form of an infinite family of axioms. Consider Separation. ZFC contains an axiom

(Sep) $$\forall z \exists y \forall x (x \in y \leftrightarrow x \in z \wedge \varphi)$$

for each of the infinitely many formulas φ of its language. Behind this infinite lot of axioms lies a single, unified idea that can be expressed by reference to collections.[18] For every collection C and every set x, there is a set y of all those elements of x that belong to C. Suppose we can quantify over collections. Then the infinitely many Separation axioms could be unified as the single axiom:

(C-Sep) $$\forall C \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge z \text{ belongs to } C)$$

[16] On this topic, see e.g. Boolos 1971 and Paseau 2006.
[17] Analogous considerations apply to the arithmetical principle of induction.
[18] See also Kreisel 1967.

In the literature, the desired collections are often known as *classes*, some of which can be shown to be "too big" to be sets. These are called *proper classes*. But what would these proper classes be? Just like sets, they are collections of many objects into one. But why, then, are proper classes not sets? As Boolos (1984b, 442) nicely observes, "[s]et theory is supposed to be a theory about *all* set-like objects".

Adding proper classes to a theory of sets is just like adding yet another layer of sets on top of the sets already recognized. In light of this, why shouldn't the proper classes count as just more sets? William Reinhardt puts the point well:

> [O]ur idea of set comes from the cumulative hierarchy, so if you are going to add a layer at the top it looks like you forgot to finish the hierarchy.[19]

Plural logic seems to provide precisely what we need. A proper class does not have to be *a single object* that somehow collects together many things into one. Instead of referring in a singular way to a proper class, construed as an object, why not simply refer plurally to its many members? In this way, we eliminate singular talk about proper classes in favor of plural talk about their members. For example, the cumulative hierarchy does not have to be an object. It suffices to talk plurally about *all the sets*.

Consider now the axiom scheme of Separation. This can be turned into a single axiom using a plural formulation. Given any objects *pp* and any set x, there is a set y of precisely those elements of x that are also among *pp*:

(P-Sep) $$\forall pp \forall x \exists y \forall z (z \in y \leftrightarrow z \in x \wedge z \prec pp)$$

Let us make two final observations. To represent all of the classes that we might be interested in, we would need an unrestricted form of plural comprehension, namely:

(P-Comp) $$\exists x \varphi(x) \rightarrow \exists xx \forall x (x \prec xx \leftrightarrow \varphi(x))$$

Moreover, we need plural logic to be ontologically innocent. If plural variables commit us to new objects, using plurals in the formulation of Separation or Replacement is not essentially different from using proper classes.

[19] Reinhardt 1974, 32. For a useful elaboration of the point, see Maddy 1983, 122.

4.9 Are two applications of plural logic compatible?

We have described two very attractive applications of plural logic: as a way of giving an account of sets, and as a way of obtaining proper classes "for free". Regrettably, it looks like the two applications are incompatible. The first application suggests that any plurality forms a set. Consider any objects *xx*. Presumably, these are what Gödel calls "well-defined objects". If so, it is permissible to apply the "set of" operation to *xx*, which yields the corresponding set {*xx*}. The second application, however, requires that there be pluralities corresponding to proper classes, which by definition are collections too big to form sets. For example, there must be a plurality of all sets whatsoever to serve as the proper class *V*. But, when the "set of" operation is applied to this plurality, we obtain a universal set, which is unacceptable.

Is there any way to retain both of the attractive applications of plural logic? To do so, we would have to restrict the domain of application of the "set of" operation so that the operation is *undefined* on the very large pluralities that correspond to proper classes, while it remains defined on smaller pluralities. The obvious concern is that this restriction would be *ad hoc*. The operation does apply to vast infinite pluralities, thus forming large sets in the cumulative hierarchy. But once we allow that these infinite pluralities form sets, why are other infinite pluralities suddenly too large to do so?

To respond to this challenge, we might seek inspiration from Gödel, who points to a restriction when he requires that the "set of" operation be applied to "well-defined objects". How should this restriction be understood? One option is to understand Gödel as requiring that the objects in question be *properly circumscribed*. Perhaps a collection corresponding to a proper class fails to satisfy this requirement. We explore this idea in Chapter 12 and find that there are indeed "collections" that fail to be properly circumscribed. However, we also argue that every plurality is (in the appropriate sense) properly circumscribed and can thus figure as an argument of the "set of" operation. Thus, if our argument succeeds, the two applications of plurals remain incompatible, and we must choose between them. We recommend retaining the first application of using plurals to give an account of sets, while looking elsewhere for an interpretation of talk about proper classes that aren't properly circumscribed and therefore cannot figure as arguments of the "set of" operation. A natural option is to look to second-order logic. We discuss this in Section 12.8.

Appendices

4.A Defining the translations

We wish to define a translation τ from the language of our simple set theory SST to that of plural logic. The central idea is obvious: let us replace talk about a set with talk about the objects that are elements of the set. Thus, instead of saying that x_i is an element of the set X_j, we say that x_i is one of xx_j. So we adopt:

$$\tau(x_i \in X_j) = x_i \prec xx_j$$

Identity statements involving set terms are translated as the corresponding plural coextensionality statements. For example, '$X_i = X_j$' is translated as:

$$\forall x_0(x_0 \prec xx_i \leftrightarrow x_0 \prec xx_j)$$

Atomic predications concerning individuals are left unchanged by the translation. Next, the translation commutes with the logical connectives. For example, the translation of a negated formula is the negation of the translation of the formula:

$$\tau(\neg\varphi) = \neg\tau(\varphi)$$

Finally, we need to translate existentially quantified formulas. (For simplicity, we may treat universal quantifiers as abbreviations in the usual way.) The individual existential quantifier poses no problem: here too we let the translation commute with the logical operator.

The set existential quantifier is slightly harder. Suppose we let the translation commute, setting '$\tau(\exists X_j\, \varphi)$' to be '$\exists xx_j\, \tau(\varphi)$'. This does not quite work. For we want to have an empty set but no empty plurality. Boolos (1984b, 444) proposes a trick to iron out this wrinkle. Let τ translate '$\exists X_j \varphi$' as

$$\exists xx_j \tau(\varphi) \vee \tau(\varphi') \tag{4.5}$$

where φ' is the result of substituting '$x_i \neq x_i$' everywhere for '$x_i \in X_j$'. The second disjunct simulates an expansion of the range of quantification, thus accommodating the possibility that a set is empty. (To see how this works, suppose X_j is empty. Then '$x_i \in X_j$' always has the same truth value as '$x_i \neq x_i$', namely *false*.) By induction on formal derivations, one can easily prove that each theorem of SST is mapped to some theorem of PFO.

It is easy to define a "reverse" translation that maps formulas of the language of plural logic to formulas of our two-sorted set-theoretic language. As expected, one can prove that this translation maps theorems of the former to theorems of the latter. So we can translate in both directions between PFO and SST while preserving theoremhood. Analogous results can be obtained for PFO+ and SST+.

4.B Defining the interpretation

The two translations we have just encountered illustrate an important general notion, which will provide a useful conceptual tool in subsequent discussions. So let us make explicit the relevant properties of the translations.

Suppose we are comparing two theories, T_1 and T_2, which are formulated in two multi-sorted languages $\mathcal{L}_1$ and $\mathcal{L}_2$, respectively. (Note that all the formal languages we consider in this book can be viewed as languages of this kind.) And suppose we have specified a translation τ from $\mathcal{L}_1$ to $\mathcal{L}_2$ such that

(i) τ is recursive, that is, there is an effective algorithm that specifies how to translate any given formula of $\mathcal{L}_1$;
(ii) τ commutes with the logical connectives (for instance, $\tau(\neg\phi) = \neg\tau(\phi)$);
(iii) τ maps every theorem of T_1 to a theorem of T_2.

We wish to make some remarks about the translation of quantified formulas. First, the translation should permit a change in the type of variables. In particular, we sometimes want to map plural variables to set variables and vice versa. Second, a quantified formula is usually translated as a restricted quantification:

$$\exists v\,\phi \overset{\tau}{\longmapsto} \exists u(\theta(u) \land \tau(\phi))$$

But this requirement is unnecessary. In fact, to accommodate Boolos's trick, which simulates an expansion of the range of a quantifier, we must refrain from requiring that every quantified formula be translated in this way.

A translation that satisfies these three properties is said to provide *an interpretation of* T_1 *in* T_2. When there are such translations in both directions—as in the examples mentioned in the previous section—the two theories are said to be *mutually interpretable*.

As just defined, the notion of interpretability is entirely proof-theoretic: it is concerned with syntax, not semantics. However, by the soundness of the proof systems we use for our logic, the notion has a semantic upshot as well. Suppose τ is an interpretation of T_1 in T_2. Then any model of T_2 allows us to define, in a recursive manner, a model of T_1. The basic idea is simply to interpret each predicate of $\mathcal{L}_1$ in accordance with its τ-translation into $\mathcal{L}_2$ and to let the domain(s) of $\mathcal{L}_1$ be interpreted in accordance with how its quantifiers are translated into $\mathcal{L}_2$.