Mending the Master

Øystein Linnebo
University of Bristol

April 4, 2006

John Burgess, *Fixing Frege*. Princeton, NJ: Princeton University Press, 2005. Pp. xii + 257. ISBN 0-691-12231-8

1 General remarks

Fixing Frege is one of the most important investigations to date of Fregean approaches to the foundations of mathematics. In addition to providing an unrivalled survey of the technical program to which Frege's writings have given rise, the book makes a large number of improvements and clarifications. Anyone with an interest in the philosophy of mathematics will enjoy and benefit from the careful and well informed overview provided by the first of its three chapters. Specialists will find the book an indispensable reference and an invaluable source of insights and new results.

Although Frege is widely regarded as the father of analytic philosophy, his work on the foundations of mathematics was for a long time rather peripheral to the ongoing research. The main reason for this is no doubt Russell's discovery in 1901 that the paradox now bearing his name can be derived in Frege's logical system. But recent decades have seen a huge surge of interest in Fregean approaches to the foundations of mathematics. (The work of George Boolos, Kit Fine, Bob Hale, Richard Heck, Stewart Shapiro, and Crispin Wright is singled out for particular attention in the present monograph.) A variety of consistent theories have been discovered that can be salvaged from Frege's inconsistent system, and foundational and philosophical claims have been made on behalf of many of these theories.

Burgess claims quite plausibly that the significance of any such modified Fregean theory will in large part depend on how much of ordinary mathematics it enables us to develop.[1] His

[1]However, if one holds that different parts of mathematics have different sorts of foundation, then all that

book is accordingly "a survey of various modified Fregean systems, attempting to determine the scope and limits of each" (p. 2).[2] The book's agenda is thus predominantly technical, and its spirit, open-minded and experimental. The author has no strong commitments of either technical or philosophical nature to any particular neo-Fregean theory. Indeed, he confesses that "it often seems that the less one keeps of Frege, the more one gets of mathematics" (p. 2), and he ends up favoring a theory of sets that owes more to Frege's rival Cantor than to Frege himself (pp. 213-14).

Burgess characterizes *Fixing Frege* as a companion to the technical middle parts of Burgess and Rosen, 1999 (p. 2). This characterization is apt concerning both what the book offers and what it doesn't offer. As for the former, the present book does an even greater service to the neo-Fregean program—by exploring, systematizing, and improving on such approaches—than the middle parts of Burgess and Rosen, 1999 did to nominalism. As for the latter, where the technical parts of the earlier book were flanked by parts containing a careful philosophical analysis and assessment of the foundational program under scrutiny, the present book contains just a minimum of philosophical material. This decision can probably be defended on the ground that the book is sufficiently demanding as it is. (Although even its more specialized parts are in principle accessible to anyone versed in intermediate level logic.) But some more philosophically oriented readers will no doubt feel that *Fixing Frege* reads a bit like "100 theorems in search of a philosophy." Moreover, readers new to this area will have to look elsewhere for a more complete sense of the potential philosophical significance of the Fregan approaches under discussion.

In what follow I go through the book chapter by chapter, summarizing and explaining some of the technical material, and criticizing and expanding on some of the philosophical material.

2 The rise and fall of Frege's foundation

Chapter 1 opens with a survey of Frege's logic and ontology. The *grammar* of Frege's higher-order logic is nicely explained as follows. First there are sentences, which are of type S. Then there are names, which are of type N. Finally, whenever T_1, ..., T_n are types, there are n-place predicates taking arguments from precisely these types; these predicates themselves

can be required of a particular foundational theory is obviously just that it suffice for the development of the parts of mathematics for which it is intended.

[2]Unless otherwise noted, all page references are to *Fixing Frege*.

being of type $N/T_1 \ldots T_n$. For instance, there are one- and two-place first-level predicates with types (S/N) and (S/NN) respectively. There are also second-level predicates of type S/(S/N), third-level predicates of type S/(S/(S/N)), and so on.

Each type of expression refers, according to Frege, to a particular type of entity: sentences refer to truth-values; names, to objects; and predicates of some given type, to concepts of the corresponding type. Concepts are (in Frege's mature period) understood as functions from their arguments to a truth-value and are (in Burgess's reconstruction) subject to *the law of extensionality*, which says that two concepts are identical if they yield the same value on all arguments.

Burgess next explains Frege's *higher-order logic* as a many-sorted first-order logic, whose sorts correspond to the types just mentioned and with some additional primitives and axioms (p. 12). The vexed question about the "logicality" of higher-order logic then translates into a question about the "logicality" of these additional primitives and axioms. But for technical purposes, these questions can be set aside: what matters is then just the formal theory. The higher-order theories in question have two axiom schemes governing the higher-order entities (concepts). First there is the scheme of extensionality for concepts, which I have already explained. Then there is a comprehension scheme for concepts, which says that various linguistic conditions succeed in defining concepts. For one-place, first-level concepts, these axioms look as follows:

(Comp) $$\exists X \forall x (Xx \leftrightarrow \phi(x))$$

However, higher-order logic alone doesn't given us much mathematics. So more substantive axioms are needed. Frege seeks to satisfy this need by adopting his infamous *Basic Law V*, which Burgess formalizes as

(V) $$\ddagger F = \ddagger G \leftrightarrow F \equiv G,$$

where ‡ is an operator that maps a concept to its extension, and $\equiv$ is the relation of coextensionality. Burgess prefers this axiomatic version of Basic Law V to the following schematic version

(V^*) $$\{x : \phi(x)\} = \{x : \psi(x)\} \leftrightarrow \forall x(\phi(x) \leftrightarrow \psi(x))$$

which bypasses concepts entirely. His reason is that only the axiomatic version brings out the fact that extensions are metaphysically derived from concepts and thus allows us to capture "Frege's principle of subordination," which says "that the relationship of an element to a set is derivative from the relationship of an element to a concept under which it falls" (p. 20).

However, Burgess points out that even the axiomatic version (V) is less than ideal, as it hides the existential assumptions of Basic Law V (p. 19). To make these assumptions explicit, write ExX for "x is the extension of X." (V) can then be "factored" into two components:

$$\exists x ExX \tag{V$\exists$}$$

$$ExX \wedge EyY \rightarrow (x = y \leftrightarrow X \equiv Y) \tag{V=}$$

The former component represents the existential import of Basic Law V (every concept X determines an extensions x), and the latter, its criterion of identity (any extensions must be subject to the law of extensionality). This is Burgess's preferred formalization of Basic Law V. However, since each concept F determines a unique extension, we may still use $\ddagger F$ to denote its extension. And when $\phi(x)$ determines a concept, we may still use $\{x : \phi(x)\}$ to denote the set of ϕ's.

Burgess also introduces two defined notions, ß for "is a set" and $\in$ for elementhood:

$$ß y \leftrightarrow \exists Y EyY \tag{ß}$$

$$x \in y \leftrightarrow \exists Y (EyY \wedge Yx) \tag{$\in$}$$

Note that these definitions rely on the principle of subordination, as extension-theoretic notions are defined in terms of how extensions are obtained from concepts.

As is well known, Frege showed how this simple theory of extensions allows us to define the basic notions of arithmetic and derive its basic principles (expressed using these definitions). We will look at some of the details in Section 5. For now it suffices to observe that the prospects for his logicism were looking very bright.

All that changed in June 1902 when Frege received a letter from Russell containing a simple derivation in Frege's system of (what we now know as) Russell's paradox. It is useful to break the derivation into the following three steps. First, consider the formula $\rho(x)$ that says that x is the extension of a concept under which x itself does not fall: $\exists X(ExX \wedge \neg Xx)$. By unrestricted concept comprehension, the formula $\rho(x)$ defines a concept R. Second, by (V$\exists$), this concept R has an extension r. Third, we ask whether r is a member of itself. Assume it isn't. Then by ($\in$) it is. So assume that r *is* a member of itself. Then by ($\in$) and

(V=) it isn't. We thus get the contradictory result that r is a member of itself just in case it isn't: $r \in r \leftrightarrow r \notin r$.

3 Fixing Frege's foundation

A variety of attempts have been made to fix Frege's foundation. The oldest attempt is due to Frege himself, who suggested that Basic Law V be modified such that the extensions of two concepts can be identical regardless of whether they agree on their extensions. More precisely, assume x and y are the extensions of the concepts X and Y respectively. Then Frege suggests that $x = y$ just in case X and Y agree on all objects with the possible exception of x and y. Although this amendment blocks the derivation of Russell's paradox, Burgess nicely shows how another contradiction can be derived on the plausible assumption that there are at least two objects.

Next Burgess gives a very accessible explanation of Russell's incredibly complicated ramified theory of types, which is Russell's most developed attempt at a fix. Russell rejected any sort of extension operator ‡ converting higher-order entities into objects. He also rejected Frege's choice of higher-order entities, preferring to Fregean concepts what he called *propositional functions*. Unlike Fregean concepts, propositional functions are not subject to the law of extensionality. But more importantly, each level of propositional function is subdivided into what Burgess calls *rounds* (but are more standardly called *orders*). A *first-round* propositional function is one definable by means of an open formula containing no quantifiers ranging over concepts. A *second-round* propositional function allows the defining formula to contain quantifiers ranging over first-round propositional functions but not over other such functions. Finally, Russell's ramified theory of types allows propositional functions of any finite round.

All these definitions of higher-order entities are *predicative* in the sense that they do not quantify over any totality to which the defined entity belongs. For a definition to be predicative is for it to be innocent of a certain kind of circularity—which may or may be vicious. Burgess does not take a stand on the nature of this circularity or analyze the question in any detail.

Russell's ramified theory of types (with an axiom of infinity) is known to be consistent. But it is also a paradigm example of a theory so hampered by philosophically motivated restrictions as to be incapable of interpreting much of classical mathematics. For instance, the axiom scheme of induction ends up being severely restricted in this theory; in particular, we lose induction on any formula containing the natural number predicate! To deal with this problem, Russell is forced to introduce the poorly motivated axiom of reducibility, which effectively undoes the effect of the ramification. Burgess also shows how, in order to obtain the laws of arithmetic, Russell needs to add an axiom of infinity. Burdened with such additional axioms, Russell's theory lost any right to count as "pure logic." As a result of this failure, logicians and philosophers in the middle decades of the last century generally turned to set theory as their foundational theory, leaving more Fregean approaches to languish.

However, recent decades have seen an explosion of interest in Fregean approaches to the foundations of mathematics. A variety of attempts have been made to find new and philosophically more appealing ways to fix Frege. To categorize these attempts, recall the derivation of Russell's paradox. The first step took us from the formula $\rho(x)$ to the concept R. The second step took us from the concept R to its extension r. The third step observes that r is a member of itself just in case it isn't. Now, this third step can hardly be denied, as it relies on nothing but a definition and the law of extensionality for sets. The recent attempts to fix Frege therefore fall into two groups, according as they reject the first step or the second. To reject the first step is to deny that all formulas define concepts. To reject the second step is to deny that all concepts have extensions. (Russell's theory is an overkill that rejects both steps.) Burgess describes a variety of attempted fixes from groups one and two in respectively Chapters 2 and 3.

In accordance with the main aim of the book—to determine how much of classical mathematics can be done in these various theories—Burgess defines a precise notion of mathematical strength and describes a scale on which this strength can be measured. The notion of mathematical strength has two ingredients. First, a theory T_2 is said to be *interpretable in* another theory T_1 if there is a translation of the formulas of the language of T_2 into formulas of the language of T_2 such that (a) the axioms of T_2 are mapped to theorems of T_1, and (b) the translation preserves logical structure (p. 50). It follows that every theorem of T_2 is mapped to a theorem of T_1. Second, T_2 is said to be *conservative over* T_1 if every formula in the language of T_1 that is provable in T_2 is also provable in T_1 (p. 52). Finally, we say that T_2 is *reducible to* T_1 if T_2 is interpretable in a conservative extension of T_1. One easily sees that any theory reducible to a consistent theory is itself consistent. Relative to a meta-theory, reducibility is therefore a measure of *consistency strength*.

Next, Burgess points out that virtually all theories of actual foundational interest have comparable consistency strengths (p. 54). Relying on this striking fact, he describes a scale on which consistency strength can be measured. The scale takes us from Robinson arithmetic at the bottom, through fragments of primitive recursive arithmetic, to Peano arithmetic and various higher-order arithmetics, onwards to weak systems of set theory, and all the way up to set theories with various large cardinal assumptions. The description of this scale is unrivalled: it is both very detailed and quite accessible.

Despite the indisputable technical importance of the notion of reducibility, this reviewer finds its philosophical significance somewhat unclear. Although the translation underlying an interpretation of one theory in another preserves theoremhood, this does not automatically guarantee that it also preserves philosophically important properties such as epistemic status or ontological commitments. For instance, the fact that geometry is interpretable in analysis doesn't automatically show that the former enjoys the same epistemic status as the latter. To show that an interpretation preserves philosophically important properties, we will in general have to show that the translation preserves linguistic meaning (up to some suitable equivalence). Moreover, as Burgess mentions on p. 52, an analogous restriction applies to the notion of conservativeness. Assume one theory is conservative over another theory in which we justifiably believe. This only entitles us to *instrumental use* of the former theory, not to outright belief in it.

4 Cutting down on concepts

The first class of attempts to fix Frege impose restrictions on which formulas define concepts. The most natural such restriction allows concept comprehension only on predicative formulas. This restriction has been defended by Michael Dummett, who blames the inconsistency of Frege's system not on Basic Law V itself but rather on the impredicativity of the background second-order logic.[3] Richard Heck has since proved that Basic Law V is indeed consistent when only predicative comprehension is allowed.[4] His result holds even for the schematic form (V*) of Basic Law V, which in the context of restricted comprehension is stronger than the axiomatic form (V). Moreover, Heck's result has since been strengthened to so-called Δ^1_1-comprehension by Fernando Ferreira and Kei Wehmeier.[5] These results, which provide some partial support for Dummett's controversial analysis, receive nice proofs in the book.

However, for Dummett's analysis to be convincing, a certain amount of mathematics has to

[3]See Dummett, 1991, pp. 217-222.
[4]See Heck, 1996.
[5]See Ferreira and Wehmeier, 2002.

be interpretable in these theories; otherwise the contradiction will be avoided not by excising a precisely circumscribable "source of inconsistency" but more bluntly by rendering our theories impotent. Burgess's meticulous examination of whether this is so yields a negative verdict: It turns out that very little mathematics can be interpreted in such theories. Burgess shows that Basic Law V with predicative monadic second-order logic interprets Robinson arithmetic Q (and therefore also the system known as $I\Delta_0$), and that adding one more "round" of predicatively defined concepts allows us to define exponentiation (and thus interpret $I\Delta_0(\exp)$).[6] Extrapolating, one might think that allowing more rounds of predicative concepts should allow us to interpret stronger arithmetical theories. This hope is quashed by a proof in Section 2.5 that functional arithmetic with so-called "superexponentiation" (what is known as $I\Delta_0$(superexp)) suffices to prove the consistency of Basic Law V in a monadic second-order logic with *any finite number of rounds* of predicatively defined concepts.[7] By Gödel's second incompleteness theorem, this means that no matter how many rounds of predicatively defined concepts we allow, we will never be able to define superexponentiation. Since this is a very serious limitative result, a blanket ban on impredicative second-order comprehension à la Dummett appears unattractive.[8]

To be defensible, any call to restrict the second-order comprehension scheme thus has to be more nuanced. One such claim is made by the present reviewer in Linnebo, 2004, namely that regardless of the permissibility of impredicative comprehension in general, *the most basic concepts and axioms of arithmetic* should not be made to depend on principles of such great strength and sophistication. A technical result of Burgess's that bears on this claim will be discussed in the next section.

Another potential source of restrictions on the second-order comprehension scheme is the idea that individuation has to be well-founded: For each type of object, there must be a criterion of identity that doesn't presuppose anything about the identity of the object to be individuated. This idea gives rise to different requirements when applied to objects that are individuated in different ways. Sets are typically taken to be individuated in terms of

[6] $I\Delta_0$ is a weak theory of second-order arithmetic which instead of an induction scheme has an axiom stating that induction holds for all concepts, but which allows concept comprehension only on so-called Δ_0-formulas, that is, on formulas with only bounded quantifiers. $I\Delta_0(\exp)$) adds to $I\Delta_0$ a symbol for exponentiation and recursion equations that define it. Although $I\Delta_0$ is interpretable in Q, $I\Delta_0(\exp)$) is not.

[7] Superexponentiation $\Uparrow$ is defined by the recursion equations $x \Uparrow 0 = x$ and $x \Uparrow (y+1) = x^{(x \Uparrow y)}$.

[8] However, as Burgess notes at the end of Chapter 2, some loose ends still exist. In particular, it is not known how much arithmetic is interpretable in the strongest consistent system of the sort in question: the schematic version (V^*) of Basic Law V in a monadic second-order logic with Δ^1_1-comprehension. However, Burgess conjectures that this will not take us very much further up his scale of mathematical theories.

their elements: same elements, same set. If so, then the well-foundedness requirement gives rise to the axiom of Foundation, which says that the $\in$-relation is well-founded. Concepts, on the other hand, appear to be individuated in an intensional way: same condition of application, same concept. If so, then for a condition of application to succeed in individuating a concept, it must not presuppose anything about the identity of the concept it attempts to individuate. But such presuppositions are easily incurred when the defining condition quantifies over a range of concepts containing the concept to be individuated. So the well-foundedness requirement may well give rise to a restriction on the concept comprehension scheme of a broadly predicativist character.[9]

5 Cutting down on extensions

The second class of attempts to fix Frege deny that all concepts have extensions. The investigation of such attempts consists in large part in an investigation of so-called *abstraction principles*, which are principles of the form

$$(*) \qquad \S\alpha = \S\beta \leftrightarrow \alpha \sim \beta$$

where α and β are variables of some type and $\sim$ is an equivalence relation on the entities over which these variables range. Basic Law V is the strongest possible abstraction principle on concepts (assuming, as Burgess does, the law of extensionality for concepts). For given extensions, we can interpret any other abstracts on concepts; for instance, interpreting $\S X$ as $\{\ddagger Y : X \sim Y\}$ makes the abstraction principle with respect to $\sim$ true. This raises the hope that, even if Basic Law V has to be given up, there may be some natural class of weaker abstraction principles all of which are acceptable. The program of looking for such a class is called *abstractionism*.

As it turns out, little is lost if in studies of the second class of attempts to fix Frege we restrict our attention to abstraction principles. For the most interesting alternative way of cutting down on extensions can be adequately modeled by means of abstraction principles. This alternative is to assign an extension to a concept X only if X satisfies some condition $\phi(X)$. But if we assign some "dummy abstract" to all and only those concepts that don't

[9]Both Fine, and Linnebo, 2006 explore this idea (although both bypass concepts altogether in favor of (what are here called) extensions). This results in theories of both sets and extensions, where the former are individuated from below via their elements, and the latter are individuated from above via their condition of application. Such theories have some interesting foundational applications.

satisfy this condition, then this proposal corresponds to adopting the abstraction principle

$$\ddagger X = \ddagger Y \leftrightarrow X \equiv Y \vee (\neg\phi(X) \wedge \neg\phi(Y)).$$

Abstractionism has enjoyed some preliminary success with *Hume's Principle*, which says that the number of Fs is identical to the number of Gs just in case the Fs and the Gs can be one-to-one correlated. This abstraction principle can be formalized as

$$\text{(HP)} \qquad \#F = \#G \leftrightarrow F \approx G$$

where $\#F$ denotes the number of Fs and $F \approx G$ abbreviates the second-order claim that there is a relation that one-to-one correlates Fs and the Gs. Burgess describes (with unsurpassed attention to historical detail) how Hume's Principle has been found to have two very desirable properties. Let *Frege Arithmetic* (or *FA* for short) be the second-order theory whose sole non-logical axiom is (HP). Then firstly, FA is consistent. And secondly, Frege's own definitions and derivations show how FA gives rise to all the axioms of ordinary Peano arithmetic—a result known as *Frege's theorem.*

The previous section alluded to the question how strong concept comprehension axioms are needed for Frege's theorem to go through. Linnebo, 2004 proves that, given Frege's own definitions of the primitive signs of arithmetic, predicative comprehension is insufficient to prove that every number has a successor. In Section 2.3 Burgess cleverly bypasses this limitative result by adopting an alternative, non-Fregean definition of natural number predicate, which allows him to interpret Robinson arithmetic Q (which of course says that every number has a successor) in FA with only predicative comprehension. If one imposes no requirements (beyond the obvious technical ones) on one's definitions of the primitive signs of arithmetic, then Burgess's result will dispel any worries caused by Linnebo, 2004. If, on the other hand, one requires that these definitions should correspond to our ordinary arithmetical thought and practice, and if one thinks that Frege's definitions are implicit in such thought and practice, then the result will be of no immediate philosophical significance.[10]

Setting aside all questions about impredicative comprehension, it is clear that any viable abstractionism must be able to delineate a much wider class of acceptable abstraction principles than just Hume's Principle. Burgess summarizes some of the efforts towards this goal

[10]Heck, gives a different predicative proof of Frege's theorem, using Frege's own definitions transposed to the un-Fregean setting of ramified second-order logic.

undertaken by "the Scottish school" of abstractionism, which "proceeds piecemeal, adding specific abstraction principles one by one" (p. 170). Proceeding in this way, the Scottish school has successfully extended abstractionism to cover real analysis (Section 3.2), but has had only partial success in their attempted extensions to set theory (Section 3.4).

The most influential objection to abstractionism is, according to Burgess, the *bad company objection* (p. 164). According to this objection, we are not entitled to belief in Hume's Principle and other useful abstraction principles because they are surrounded by unacceptable principles. There are, for instance, inconsistent abstraction principles dangerously close to (HP). An example is the "dyadic counterpart" of (HP), which says that the order type of a dyadic relation R is identical to the order type of another such relation S just in case R and S are isomorphic. Although this principle only does to dyadic relations what (HP) does to concepts, it falls victim to the Burali-Forti paradox of the largest ordinal.

One of the most promising-looking responses to the bad company objection is that abstraction is only acceptable on second-level equivalence relations $\mathbf{R}$ that are *non-inflating* in the sense that there are no more $\mathbf{R}$-equivalence classes than there are objects. But Burgess points out two problems with this response. Firstly, on this account an independent guarantee of the infinity of the universe would be needed. But the infinity of the universe is supposed to follow from (HP), not be a presupposition for it. Secondly, there remains a problem of *hyperinflation.* For any concept Z, consider the equivalence relation $\mathbf{R}_Z$ which X bears to Y just in case *either* both X and Y coincide with Z *or* neither does. Abstraction on each $\mathbf{R}_Z$ is clearly non-inflating, as this relation divides all concepts into just two equivalence classes. Nevertheless, on the plausible assumption that two abstracts can be identical only if their corresponding equivalence classes of concepts are identical, any theory that allows abstraction on each $\mathbf{R}_Z$ is inconsistent. For in any such theory we can interpret the inconsistent Basic Law V by letting $\ddagger X$ be the abstract of X with respect to the relation $\mathbf{R}_X$.

To deal with the problem of hyperinflation, Fine, 2002 develops a *general* theory of abstraction principles. The core idea of this theory is that abstraction cannot be allowed on equivalence relations such as $\mathbf{R}_Z$ on the ground that they are non-logical: for almost all concepts Z, the reference to Z will be a breach of topic-neutrality. To rectify this Fine extends Tarski's characterization of logical notions in terms of permutation invariance to second-level concepts. He then shows that the theory which allows abstraction on all non-inflating *and logical* second-level equivalence relations $\mathbf{R}$ is consistent. Fine shows that $(n+1)$'st-order Peano arithmetic is interpretable in the n'th order version of the resulting general theory of abstraction.

Burgess gives a very compact presentation of Fine's theory in Section 3.3. Next he advances the discussion considerably by proving the converse of Fine's result. Burgess has thus pinpointed the exact mathematical strength of Fine's general theory of abstraction: that of third-order arithmetic, on the assumption that our logic is (merely) second-order. What does this tell us? On the positive side, Fine has given us a theory of abstraction on concepts which is both consistent and suitably general. On the negative side, although Fine's theory is strong enough for most practical purposes, Burgess's result shows that it stops far short of contemporary set theory. Moreover, the theory fails to address the bad company objection in its full generality, as it only applies to abstraction on concepts, not on polyadic relations.

6 What remains of Frege's foundation?

Towards the end of the book Burgess changes tack, leaving Fregean approaches behind in favor of a novel and extremely elegant development of the old set-theoretic idea of limitation of size. According to this idea, a concept X defines a set just in case the things falling under X are not "too many" to form a set. On the usual analysis, some things are "too many" just in case they can be put in one-to-one correspondence with everything there is. This analysis is known to yield some but not all of set theory. Burgess develops a new and different analysis, starting with the idea that some things are "too many" just in case they are "undefinably many" (p. 192): that any attempt to characterize how many they are fails because this characterization is also true of some things that are not "too many." Now, one concept that has "too many" objects falling under it is (on pain of paradox) the universal concept. Applied to this special case, Burgess's guiding idea gives rise to the principle that "anything that is true when said about all objects remains true when said about just some objects," few enough to form a set (p. 192). He gives a precise formalization of this principle in terms of relativization of quantifiers. Let ϕ^t be the result of relativizing all quantifiers in ϕ to t.[11] Then the above principle amounts to the following so-called *reflection principle*:

(Refl) $$\phi \rightarrow \exists t\, \phi^t$$

This analysis of the idea of limitation of size gives rise to an amazing theory of sets,

[11] To do so is to replace each first-order universal quantifier $\forall x$ in ϕ with $\forall x(x \in t \rightarrow \ldots)$, to replace each first-order existential quantifier $\exists x$ with $\exists x(x \in t \wedge \ldots)$, and likewise for the second-order quantifiers.

which Burgess calls *Fregeanized Bernays set theory* or *FB* for short. FB uses just *monadic* second-order logic; that is, it has no variables or quantifiers for polyadic relations. This logic is just what is needed for the unique Fregean element of FB, namely the principle of subordination, according to which the notions of *set* and *element* are derivative from the notions of *concept* and *falling under*. So Burgess adopts the principles (ß) and ($\in$) as axioms (rather than as definitions, which was their status in Chapter 1 of the book, as discussed in Section 2 above). He also adopts the axiom (V=) of extensionality for set, which arguably is an analytic truth, and a second-order axiom of separation for sets, which will be true on any version of the doctrine of limitation of size. Finally there is the reflection principle (Refl). Drawing on earlier work by Paul Bernays, Burgess shows that "All the existence axioms of second-order Zermelo-Frankel set theory as well as inaccessibles, hyperinaccessibles, ... are deducible in" FB (p. 196). He also shows how the remaining axioms of (strong) extensionality and Foundations can be obtained in a natural way by relativizing the quantifiers to the pure well-founded sets.

Despite this success, it is worrisome how extremely sensitive FB is on the choice of primitive vocabulary. It would for instance be very natural to regard the subordination principles (ß) and ($\in$) as definitions rather than as axioms. But then Burgess's proof no longer goes through. If on the other hand we allow notions that *can* be defined to instead be primitives governed by additional axioms, then inconsistency threatens. For the adoption of a new primitive $\equiv$ governed by the axiom $X \equiv Y \leftrightarrow \forall z(Xz \leftrightarrow Yz)$ would then be just as natural as Burgess's adoption of ß and $\in$ as primitives governed by (ß) and ($\in$) respectively. But the theory with this new primitive $\equiv$ is inconsistent.[12]

Fixing Frege ends with a critical evaluation of second-order logic. While there is much to be learnt from this discussion (especially its earlier parts), there is also much to disagree with. Burgess develops an argument against any conceptual interpretation of higher-order logic (such as Frege's). This argument has great destructive potential; for if it succeeds, it will pull the rug from under almost all of the book, the only serious exception being Burgess's own Fregeanized Bernays set theory, which can survive on Boolos's legitimate interpretation of monadic second-order logic in terms of ordinary English plural quantifiers.[13] The argument has three parts. First, Burgess claims that the meanings of the higher-order quantifiers have

[12]The trick needed to see this is to apply the reflection principle to the theorem $\exists Y(X \equiv Y)$ to get $\exists t \exists Y[\forall z(Yz \rightarrow z \in t) \wedge Y \equiv X]$, which ensures that there is a set t containing as elements absolutely all the objects that fall under X. I elaborate on this criticism and suggest a response to it in unpublished work.

[13]See Boolos, 1984.

to be explained in some natural language such as English or German. Second, he claims that any such explanation will turn the concepts over which the higher-order variables are said to range into objects. His argument, borrowed from Frege, is that phrases such as 'the concept *horse*' are singular terms and as such will denote objects, not concepts, if they denote at all. Third, Burgess asserts that a concept cannot also be an object, as this is disallowed by the inconsistency of Basic Law V.

Each step of this argument can be challenged. As this is not the place to assess these challenges, I will simply state them and observe that Burgess does nothing to addresses them. Against the first step it can be objected that higher-order languages can be learnt directly, without going via some natural language.[14] Against the second step it can be objected that English *does* have the resources to explain quantification over at least first-level relations. For instance, '$\forall R(Rab \rightarrow Rcd)$' can be read as "however a stands to b, so c stands to d."[15] Against the third step we recall from Section 4 that Basic Law V is perfectly in order provided the concept comprehension scheme is suitably restricted. We also observe that it is extremely natural to compare and contrast one's favorite higher-level entities with (ordinary) objects in a way that requires the use of *one* kind of variable to range over both kinds of entity.[16]

So for all we know, more than just monadic second-order logic may be available to Frege and neo-Fregeans. This means that more of Frege's foundation may remain than Burgess admits. However, if I am right about this, it will only make this wonderful little book more important, not less.[17]

References

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

Burgess, J. P. and Rosen, G. (1999). *A Subject with No Object.* Oxford University Press, Oxford.

[14] See Williamson, 2003, p. 459.

[15] See Rayo and Yablo, 2001.

[16] See Linnebo, 2006.

[17] Thanks to John Burgess and Richard Heck for valuable comments on the penultimate version of this review.

Dummett, M. (1991). *Frege: Philosophy of Mathematics*. Harvard University Press, Cambridge, MA.

Ferreira, F. and Wehmeier, K. (2002). On the Consistency of the Δ^1_1-CA Fragment of Frege's *Grundgesetze*. *Journal of Philosophical Logic*, 31:301–311.

Fine, K. Class and membership. Unpublished manuscript.

Fine, K. (2002). *The Limits of Abstraction*. Oxford University Press, Oxford.

Heck, R. G. Ramified Frege Arithmetic. Unpublished manuscript.

Heck, R. G. (1996). The Consistency of Predicative Fragments of Frege's *Grundgesetze der Arithmetik*. *History and Philosophy of Logic*, 17:209–220.

Linnebo, Ø. (2004). Predicative Fragments of Frege Arithmetic. *Bulletin of Symbolic Logic*, 10(2):153–74.

Linnebo, Ø. (2006). Sets, Properties, and Unrestricted Quantification. In Rayo, A. and Uzquiano, G., editors, *Unrestricted Quantification: New Essays*, pages x–y. Oxford University Press, Oxford.

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

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