The Logic of Frege's Theorem

Richard G Heck Jr

Department of Philosophy, Brown University

1 Opening

As is now well-known, axioms for arithmetic can be interpreted in second-order logic plus 'Hume's Principle', or HP:

$$\begin{aligned} Nx : Fx = Nx : Gx \text{ iff } \exists R[&\forall x \forall y \forall z \forall w(Rxy \wedge Rzw \rightarrow x = z \equiv y = w) \wedge \\ &\forall x(Fx \rightarrow \exists y(Rxy \wedge Gy)) \wedge \\ &\forall y(Gy \rightarrow \exists x(Rxy \wedge Fx))] \end{aligned}$$

This result is *Frege's Theorem*. Its philosophical interest has been a matter of some controversy, most of which has concerned the status of HP itself. To use Frege's Theorem to re-instate logicism, for example, one would have to claim that HP was a logical truth. So far as I know, no-one has really been tempted by that claim. But Crispin Wright claimed, in his book *Frege's Conception of Numbers as Objects* (1983), that, even though HP is not a logical truth, it nonetheless has the epistemological virtues that were really central to Frege's logicism. Not everyone has agreed.[1] But even if Wright's view were accepted, there would be another question to be asked, namely, whether the sorts of inferences employed in the derivation of axioms for arithmetic from HP preserve whatever interesting epistemological property HP is supposed to have. Only then would the axioms of arithmetic then have been shown to have such interesting properties.

The problem is clearest for a logicist. If the axioms of arithmetic are to be shown to be logical truths, not only must HP be a logical truth, the modes of inference used in deriving axioms of arithmetic from it must preserve logical truth. They must, that is to say, be logical modes of inference. For Wright, the crucial question is less clear. It would be enough for his purposes if these modes of inference preserved HP's interesting epistemological properties, whatever these were taken to be. But Wright has, nonetheless, typically been content to claim that second-order reasoning is logical reasoning and to suppose, reasonably enough, that, if that claim is good enough for the logicist, it is good enough for his purposes, too.

The claim that 'second-order logic is logic', as it is often put, has had both defenders and detractors.[2] I am

[1] See (Boolos, 1998a) and (Wright, 2001) for one nice back-and-forth.
[2] Quine (1986) was famously skeptical. George Boolos (1998c; 1998e) was an early proponent.

not going to enter that debate here. What I want to argue here is that a neo-logicist does not need to commit herself to any claims about second-order logic.

In a typical proof of Frege's Theorem, axioms for arithmetic are derived from HP in second-order logic, but not all of the power of second-order logic is needed for the proofs of the axioms. The power of second-order logic derives from the so-called comprehension axioms, each of which states, in effect, that a given formula defines a 'concept' or 'class'—something in the domain of the second-order variables. These axioms take the form:[3]

$$\exists F \forall x[Fx \equiv A(x)].$$

In full-second order logic, one has such an axiom for every formula $A(x)$ (in which '*F*' does not occur free). At the other extreme, one could consider a system in which one had no comprehension axioms at all, but the weakest system seriously discussed is 'predicative' second-order logic, in which one has comprehension only for formulae containing no *bound* second-order variables. Predicative second-order logic is weak in a well-defined sense: Given any first-order theory Θ, adding predicative second-order logic to Θ yields a conservative extension of it.[4] Full second-order logic, on the other hand, is extremely powerful, and it is that power that underlies much of the skepticism about the appropriateness of the term 'second-order *logic*'.[5]

Between predicative second-order logic and full second-order logic are systems of intermediate strength, each admitting a different set of comprehension axioms. In principle, any set of comprehension axioms will do, and there are many that have been considered.[6] What are perhaps the most natural intermediate systems arise, though, from syntactic restrictions on the formulae appearing in the comprehension axioms. Say that a formula containing no bound second-order variables is Π^0_∞. Then where ϕ is Π^0_∞, formulae of the form $\forall F_1 \ldots \forall F_n \phi$ and $\exists F_1 \ldots \exists F_n \phi$ are Π^1_1 and Σ^1_1, respectively. If ϕ is Σ^1_n (Π^1_n), then $\forall F_1 \ldots \forall F_n \phi$ ($\exists F_1 \ldots \exists F_n \phi$) is Π^1_{n+1} (Σ^1_{n+1}). Second-order logic with Π^1_n comprehension has only those comprehension axioms in which $A(x)$ is Π^1_n (or simpler).

It is important to note that, as I have formulated the Π^1_n comprehension scheme, free second-order variables are allowed to occur in the comprehension axioms. As a result, there is no significant difference between Π^1_n comprehension and Σ^1_n comprehension. If $A(x)$ is a Σ^1_n formula, then its negation is (trivially equivalent to) a Π^1_n formula. Hence, Π^1_n comprehension delivers a concept *F* such that:

$$\forall x[Fx \equiv \neg A(x)].$$

[3] There are similar axioms for many-place predicates, of course.

[4] Given any model for Θ, let the second-order domain contain exactly the subsets of the first-order domain definable in the language of Θ so interpreted. The result is a model of Θ plus predicative comprehension, which is thus a conservative extension of Θ. (There is some need for a care here when Θ contains axiom schemata: The schemata must not come to have new instances as a result of the addition of second-order vocabulary.)

[5] Boolos expresses such a worry in (1997). Peter Koellner has developed a extremely detailed and sophisticated version of this objection in recent work.

[6] The standard reference on second-order logic is now (Shapiro, 1991), especially Chs 3–4, but see also (Feferman, 1984). There are many intermediate systems other than those we shall consider here.

But then predicative comprehension delivers a concept *G* such that:

$$\forall x[Gx \equiv \neg Fx],$$

and so we have:

$$\exists G \forall x[Gx \equiv A(x)].$$

We might as well therefore regard Π^1_n comprehension as Π^1_n-or-Σ^1_n comprehension.

More significantly, consider the formula:

$$\forall F[Fa \wedge \forall x(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fb].$$

This formula defines the so-called 'weak ancestral' of the relation *P*. It is obviously Π^1_1, so Π^1_1 comprehension delivers a concept $\mathbb{N}$ such that:

$$\mathbb{N}n \equiv \forall F[Fa \wedge \forall x(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fn].$$

If we take *a* to be 0 and *P* to be the relation of predecession—read Pxy as: *x* is the number immediately preceding *y*—then that is Frege's definition of the concept of a natural number. And Π^1_1 comprehension delivers the existence of this concept even if *P* itself has been defined by a Σ^1_1 formula, as it usually is in Fregean arithmetics:

$$Pab \equiv \exists G \exists y[b = Nx : Gx \wedge Gy \wedge a = Nx : (Gx \wedge x \neq y)].$$

The existence of the relation *P* is guaranteed by Σ^1_1—equivalently, Π^1_1—comprehension.

We may seem to be cheating here: Won't such a method end up reducing *all* comprhension to Π^1_1 comprehension? That would indeed be disastrous, but no such result is forthcoming. Chaining instances of comprehension together works in this case only because the variable *F* does not occur within the scope of the quantifier $\exists G$ that appears in the definition of *P*. The method will allow us to apply Π^1_1 comprehension twice to a formula of the form:

$$\forall F[\ldots F \cdots \rightarrow \exists G(\ldots G \ldots)],$$

but *not* to one of the form:

$$\forall F[\ldots F \cdots \rightarrow \exists G(\ldots G \ldots F \ldots)].$$

But one might still think such 'chaining' impermissible, even if coherent. Comprehension, so formulated, collapses Π^1_1 and Σ^1_1 comprehension and, moreover, fails to distinguish Π^1_1 sets from sets that are Π^1_1 in Π^1_1 sets. Is that really wise? Obviously, I am not suggesting that these distinctions do not matter, and if one wishes to use second-order logic to investigate problems to which these distinctions are relevant, then comprehension should be formulated so as to prohibit such 'chaining': One need only prohibit free second-order variables from appearing in the comprehension scheme. But it is not clear that these distinctions matter in the present context. I shall discuss the matter further below (see page 16). For the moment, I appeal to authority: Solomon Feferman formulates the comprehension axioms this way in his classic paper "Systems of Predicative Analysis" (1964).

Both the concept of predecession and the concept of natural number are thus delivered by Π^1_1 comprehension: That should make it plausible that the standard proof of Frege's Theorem requires only Π^1_1 comprehension, a conjecture that can be verified by working through the proof in detail, paying careful attention to what comprehension axioms are used.[7] There is a sense in which this result is best possible. I have mentioned several times that axioms for arithmetic can be derived from HP in second-order logic, but I have not yet said which such axioms I have in mind. There are, of course, many equivalent axiomatizations—I shall present one such axiomatization below—but what is important at the moment is that standard presentations of Frege's Theorem do not include a derivation of the usual first-order axioms for addition and multiplication. The reason is that, in a standard second-order language, the recursive defintions of addition and multiplication can be converted into explicit defintions in a way due, independently, to Dedekind and to Frege. The recursion equations themselves—and these just are the first-order axioms—can then be recovered from the definition. Unsurprisingly, however, the derivation of the recursion equations from the explicit definition needs more than predicative comprehension. The proof that addition and multiplication are well-defined and satisfy the recursion equations is by induction, and the induction is on a predicate containing the definition of addition or multiplication. The legitimacy of the induction thus presumes that the predicate in question defines a relation. Since the formula that defines addition is Π^1_1, we will need at least that much comprehension even to interpret first-order PA.

One can at least imagine a view that would regard Π^1_1 comprehension axioms as logical truths but deny that status to any that are more complex—a view that would, in particular, deny that full second-order logic deserves the name. In light of what has been said, such a view would serve the purposes of a new-logicist such as Wright. I do not expect it to be obvious at this point how such a view might be motivated, and it is in fact no part of the view I want to defend here that, say, Δ^1_3 comprehension axioms are *not* logical truths. What I am going to suggest, however, is that there is a special case to be made on behalf of Π^1_1 comprehension. Or something like it.

2 Predecession

As it happens, the only comprehension axioms one actually needs for the proof of Frege's Theorem—besides a handful of instances of predicative comprehension—are these:

$$\exists P\{Pab \equiv \exists G\exists y[b = Nx : Gx \wedge Gy \wedge a = Nx : (Gx \wedge x \neq y)]\}$$

$$\exists R\{Ran \equiv \forall F[Fa \wedge \forall x(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fn]\}$$

[7] The mentioned fact was first noted in (Heck, 2000). For a detailed proof and relevant discussion, see (Linnebo, 2004). Linnebo also proves a converse: PA with Π^1_1-comprehension is interpretable in FA with Π^1_1-comprehension. See also (Burgess, 2005), which will deservedly become the standard reference soon enough.

The latter, of course, defines the relation that would usually be written: $\xi P^{*=}\eta$. That is, it defines the relation that is the ancestral of predecession.

Øystein Linnebo suggests that there is something seriously wrong with Frege's defintion of predecession (2004, pp. XXX-XX). It simply does not seem reasonable to suppose that a notion as simple as that of predecession should be so logically complex. Consider, for example, the proof of the familiar fact that every number other than zero has a predecessor. This proposition, $x \neq 0 \rightarrow \exists y(x = Sy)$, is one of the axioms of Robinson arithmetic, Q, but it is redundant in PA, since it is provable in PA and, in fact, in the much weaker theory known as $I\Delta_0$, which has induction only for bounded formulae: The induction can be carried out on $x \neq 0 \rightarrow \exists y < x(x = Sy)$. In Frege arithmetic, however, formalization of that proof would require Σ^1_1 comprehension, since the induction must now be on $x \neq 0 \rightarrow \exists y(Pyx)$.[8] Are we really to believe that such strong logical resources are needed for the proof of such a simple statement? The more plausible view is the one enshrined in the usual treatment of arithmetic: Predecession is a *primitive* notion.

Linnebo's concern is a sensible one, but I think it can be answered. Although the definition of predecession is undeniably Σ^1_1 in form, it is not, I want to suggest, Σ^1_1 in spirit. The definition one would really like to give is this one:

$$\text{(P-lite)} \quad P(Nx : Gx, Nx : Fx) \equiv \exists y(Fy \wedge Nx : Gx = Nx : (Fx \wedge x \neq y)).$$

To be sure, (P-lite) is not a proper definition. It does not tell us when Pab but only when $P(Nx : Gx, Nx : Fx)$: Nothing in (P-lite) tells us whether Julius Caesar, that same familiar conqueror of Gaul, precedes 0 or not. But the obvious reply is that it was supposed to be implicit in (P-lite) that *only numbers are predecessors or successors*. If Caesar is a number, then he is the number of *F*s, for some *F*, in which case (P-lite) will determine which numbers he precedes and succeeds. If he is not a number, then he does not precede or succeed any number. Hence, the question which numbers Caesar precedes and succeeds is equivalent to the question whether Caesar he is a number and, if so, which one he is. Well, if that isn't a familiar problem! Maybe it is even a serious problem. But it is a problem the neo-logicist had anyway.

Suppose that the Caesar problem has either been solved or justifiably ignored. (Maybe it isn't a serious problem, just an amusing one.) Then (P-lite) tells one everything one needs to know about predecession. How would that allow the neo-logicist to avoid appealing to Σ^1_1 comprehension? I suggest that a neo-logicist should regard predecession as *primtive* and regard both (P-lite) and

$$\text{(P-imp)} \quad Pab \rightarrow \exists F(a = Nx : Fx) \wedge \exists G(b = Nx : Gx),$$

as analytic of that notion. (P-imp makes the implicit requirement that only numbers can be or have predecessors explicit.) No appeal to comprehension is then needed to guarantee that the relation of predecession exists, any more than in the usual formulation of second-order arithmetic.

One might worry that this strategy makes everything too easy. Why can't the neo-logicist just regard the ancestral as primitive and take the usual definition of the concept of natural number to be analytic of it?

[8] A different proof can be given that would not require comprehension at all, but there are other examples of this same form.

Then no appeal to comprehension would be needed! It will become clear that I am in a way sympathetic with that suggestion, but the arguments just offered on behalf of the claim that (P-lite) is analytic do not generalize to the case of the ancestral. Those arguments apply only to certain sorts of explicit definitions, namely, those that can be resolved into something of the form:

$$Ra_1 \dots a_n \stackrel{df}{\equiv} \exists F_1 \dots \exists F_n [a_1 = \Phi x : F_1 x \wedge \dots \wedge a_n = \Phi x : F_n x \wedge \mathcal{R}_x(F_1 x, \dots, F_n x)],$$

which is equivalent to the conjunction of

$$\text{(R-lite)} \quad R(\Phi x : F_1 x, \dots, \Phi x : F_n x) \equiv \mathcal{R}_x(F_1 x, \dots, F_n x)$$

and

$$\text{(R-imp)} \quad Ra_1 \dots a_n \rightarrow \exists F_1(a_1 = \Phi x : F_1 x) \wedge \dots \wedge \exists F_n(a_n = \Phi x : F_n x).$$

The arguments presented above purport to show that (R-lite) is already an adequate definition of *R*, *modulo* an instance of the Caesar problem. But they apply only to this sort of case.

The case of the ancestral is not such a case,[9] but there is a different such case that is important. Consider the so-called predicative fragment of *Grundgesetze*, which consists of predicative second-order logic plus a schematic form of Frege's Basic Law V:

$$\hat{x}Fx = \hat{x}Gx \equiv \forall x(Fx \equiv Gx).$$

This theory is known to be consistent (Heck, 1996). What saves the system from inconsistency is the fact that membership is defined in terms of a Σ^1_1 formula:

$$a \in b \equiv \exists F(b = \hat{x}Fx \wedge Fa),$$

and we do not have comprehension for such formulae in the predicative framgment. So, crucially, we cannot prove naïve comprehension:

$$a \in \hat{x}(x \notin x) \equiv a \notin a,$$

whence the paradox that threatens to arise when we take *a* to be $\hat{x}(x \notin x)$ is averted. But it is averted only at the cost of our inability to prove the formula just displayed, and that has always seemed to me to be deeply counterintuitive. I can now give some content to the intuition thus countered.

The definition of membership is of precisely the form we have been discussing. The definition of membership one would really like to give is this one:

$$(\in\text{-lite}) \quad a \in \hat{x}Fx \equiv Fa.$$

That is not a proper definition. It does not tell us when $a \in b$ but only when $a \in \hat{x}Fx$, and so on and

[9] It is not such a case because we know that the ancestral cannot be defined by a Σ^1_1 formula.

so forth. But *modulo* the Caesar problem, or so I would argue, ($\in$-lite) is a perfectly good definition. Any neo-Fregean who is prepared to countenance Basic Law V ought to regard membership as primitive, and characterized by ($\in$-lite) and

$$(\in \text{-imp}) \quad a \in b \rightarrow \exists F(b = \hat{x}Fx).$$

But then Russell's paradox reappears. And that seems to me an intuitively satisfying result. There is nothing *truly* impredicative about the definition of membership. The substitution of $x \notin x$ for Fx in Basic Law V *ought* to be permitted. The predicative fragment of *Grundgesetze* may be consistent, then, but it is not really *coherent*.[10]

3 Ancestral Logic

As noted above, the only impredicative instances of comprehension needed for the proof of Frege's Theorem are these:

$$\exists P\{Pab \equiv \exists G\exists y[b = Nx : Gx \wedge Gy \wedge a = Nx : (Gx \wedge x \neq y)]\}$$

$$\exists R\{Ran \equiv \forall F[Fa \wedge \forall x(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fn]\}$$

The arguments of the last section purported to establish that the former has no significant epistemological costs. If that is accepted, then we may draw the following intermediate conclusion: As far as the logic used in the proof of Frege's Theorem is concerned, the question whether it is epistemologically innocent reduces to the question what our attitude should be to Frege's definition of the ancestral.

The assumption that the ancestral of an arbitrary relation exists is much weaker than full Π^1_1 comprehension. In fact, there is a logic known as *ancestral logic* which formalizes the logic of the ancestral in an otherwise first-order language.

We may characterize ancestral logic semantically as follows (Shapiro, 1991, p. 227).[11] We begin with an ordinary first-order language $\mathcal{L}$ and form a new language $\mathcal{L}_*$ by adding an operator $*_{xy}$ which forms a relational expression from a formula with two free variables, these being bound by the operator. So we have formulae of the form: $*_{xy}(\phi xy)(a, b)$, where ϕxy is a formula. Let an interpretation of $\mathcal{L}$ be given. We expand it to an interpretation of $\mathcal{L}_*$ as follows:[12] Suppose that ϕxy is satisfied by exactly the ordered pairs in some set Φ; then $*_{xy}(\phi xy)(a, b)$ is true if, and only if, there is a finite sequence $a = a_0, \ldots, a_n = b$ such that each $< a_i, a_{i+1} >\in \Phi$. Less formally: It is true just in case *a* can be linked to *b* by a finite sequence of ϕ-steps. We require that there should be at least one such step: $*_{xy}(\phi xy)$ is therefore the *strong* ancestral of ϕ, so-called because we do not, in general, have: $*_{xy}(\phi xy)(a, a)$. The weak ancestral of ϕ, denoted $*^{=}_{xy}(\phi xy)$, may be defined in the usual way as: $*_{xy}(\phi xy)(a, b) \vee a = b$. We shall use the more familiar notation $\phi^* ab$ and $\phi^{*=}ab$, omitting the bound variables when there is no danger of confusion.

[10] Note that this argument does not even purport to show that predicativity restrictions are not otherwise justified. It is entirely specfic to the case of Basic Law V.

[11] See also (Avron, 2003) for recent work on such logics.

[12] This specification is less precise that it would really need to be, since it does not allow for additional free variables in ϕ. But let us not be too pedantic.

It is easy to see that ancestral logic is not completely axiomatizable: It permits the formulation of a categorical theory of arithmetic (Shapiro, 1991, p. 228).[13] But, of course, that need not prevent us from partially axiomatizing the logic. One way to proceed would be to take as introduction rules

$$\phi ab \vdash \phi^* ab$$
$$\phi^* ab, \phi bc \vdash \phi^* ac$$

and as an elimination rule:

$$\phi^* ab \vdash \forall x(\phi ax \rightarrow A(x)) \wedge \forall x \forall y(A(x) \wedge \phi xy \rightarrow A(y)) \rightarrow A(b),$$

Call this system *weak* ancestral logic. Its introduction rules reflects the 'inductive' character of the ancestral: Taking ϕab to mean: *b is a's parent*, they tell us that one's parents are one's ancestors and that any parent of an ancestor is an ancestor. The elimination rule is a principle of induction, in schematic form.

Weak ancestral logic incorporates, in its elimination rule, one half of Frege's definition of the ancestral. But it does not incorporate the other half of Frege's definition, and that is, to my mind, an important weakness. Consider the following argument.

> Suppose that *b* is *a*'s ancestor and that *c* is *b*'s ancestor. Suppose further (i) that all of *a*'s parents are blurg and (ii) that blurghood is hereditary—that is, that any parent of someone who is blurg is also blurg. Since *b* is *a*'s ancestor, *b* is blurg, by the elimination rule. But then, by (ii), all of *b*'s parents are blurg and so, since *c* is *b*'s ancestor, *c* is blurg, again by the elimination rule. That is, if (i) and (ii), then *c* is blurg. And so, by Frege's definition of the ancestral, *c* is *a*'s ancestor.

That, obviously, is an argument for the transitivity of the ancestral and, so far as I can see, nothing like it can be formalized within weak ancestral logic. That is not to say that the transitivity of the ancestral cannot be proved in weak ancestral logic. It can be, though in a different way, namely, in roughly the way Frege proves it in *Begriffsschrift*.[14] But that is a different argument, one whose formalization in standard second-order logic requires the use of Π^1_1 comprehension. No comprehension at all is needed for the formalization of the argument just given (Boolos and Heck, 1998, p. 319). That we can formalize the more complicated argument in weak ancestral logic but not the less complicated one suggests to me that it has things upside down.

[13] Add to a first-order formulation of *PA* the axiom: $\forall n[*^{=}_{xy}(y = Sx)(0, n)]$.

[14] Frege's proof is by induction on $\phi^* a\xi$. The elimination rule yields

$$\phi^* bc \rightarrow [\forall x(\phi bx \rightarrow \phi^* ax) \wedge \forall x \forall y(\phi^* ax \wedge \phi xy \rightarrow \phi^* ay) \rightarrow \phi^* ac]$$

The second conjunct follows immediately from the second introduction rule; the first follows from $\phi^* ab$ and the second introduction rule. Hence, $\phi^* bc \wedge \phi^* ab \rightarrow \phi^* ac$.

It would be cleaner if I had a nice example of a theorem whose proof is easily formalized using Frege's definition of the ancestral but which cannot be formalized in weak ancestral logic. There must be some, but I haven't given the matter enough thought to identify one.

One may have been wanting to ask what the nonsense term ‘blurg’ is doing in the above argument, and that is a perfectly reasonable question. But such reasoning is very common. At least it is very common for me to engage in such reasoning, especially when I am teaching logic to undergraduates. Perhaps that would be more obvious if I were to replace ‘blurg’ with ‘*F*’, but one does not have to use letters to engage in such reasoning. And there is no reason to dismiss it out of hand. In many cases, such reasoning can be understood as tacitly semantic. We may take ‘blurg’ to be a variable that ranges over expressions and construe the argument as a whole as tacitly invoking semantic notions, such as truth. A related proposal would construe the argument substitutionally. On either construal, however, this particular argument would only establish something about concepts we can *name*, whence it is surely invalid. But it seems to me a perfectly good argument, so some other way of understanding such reasoning is needed.

A fan of second-order logic might suggest that ‘blurg’ is a second-order variable and that the argument as a whole tacitly involves second-order quantification, its intuitive force revealing the extent to which second-order reasoning is intuitively compelling (Boolos, 1998e, pp. 59–60). But there are two aspects to this suggestion, and they can be disentangled: We can interpret ‘blurg’ as the natural language correlate of a *free* second-order variable and simultaneously deny that second-order *quantification* is involved in the argument at all.

Given a first-order language, add to it a stock of second-order variables. We do not permit these variables to be bound by quantifiers: They occur only free. Thus, there are formulae in the language such as:

$$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb,$$

where *F* is a free second-order variable. An interpretation of such a formula is simply a first-order interpretation. Free second-order variables are treated just as predicate-letters are in first-order logic: They are assigned subsets of the domain. Implication is then defined as usual: A set of formulae Γ implies a formula A if, and only if, every interpretation that makes all formulae in Γ true also makes A true. A formula is valid if it is implied by the empty set of formulae.[15]

The proof-theory is also straightforward. I shall take us to be working in a system of natural deduction. Such a system will have some mechanism or other for keeping track of the premises used in the derivation of a given formula. I assume that we have some natural set of rules for first-order logic already in place. No special rules that govern free second-order variables are being introduced at this point. Call the resulting system minimal *schematic logic* (minimal SL). It should be clear that minimal SL is sound with respect to the semantics mentioned above. It would also appear to be complete, since the free second-order variables hardly differ from predicate-letters.

We can now reformulate ancestral logic. The elimination rule, which we may call (*–), remains one direction of Frege’s explicit definition of the ancestral, though it now need not be formluated as a schema but can be formulated using a free second-order variable:

$$\phi^* ab \vdash \forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb$$

[15] A formula of schematic logic is therefore valid only if its universal closure is a valid second-order formula.

The introduction rule (*+) is the other direction of Frege's definition of the ancestral:

$$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb \vdash \phi^* ab,$$

where *F* may not be free in any premises on which the premise of this inference itself depends.[16] Call the resulting system (first-order) *minimal schematic ancestral logic* (minimal SAL). It should again be clear that this logic is sound if the ancestral is interpreted as indicated above. If, of course, is not complete with respect to that semantics, since no recursive axiomatization can be.

What we have done is to transcribe Frege's explicit definition of the ancestral into the framework of schematic logic. Why does the transcription work? Consider the introduction rule, (*+). If we can prove

$$(\dagger)\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb,$$

then, in standard second-order logic, we can use universal generalization to conclude that

$$(\dagger\dagger)\ \forall F[\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb]$$

and use Frege's definition of the ancestral to conclude that $\phi^* ab$. Similarly in the case of the elimination rule: $\phi^* ab$ and Frege's definition together imply (††) which in turn implies (†). But we can now see that (††) is just a rest stop and that Frege's explicit definition is a ladder we can kick away.[17] Goodbye, ladder.

The transitivity of the ancestral can be proven in minimal SAL, thus:[18]

$[1](1)$	$\phi^* ab$	Premise
$[2](2)$	$\phi^* bc$	Premise
$[3](3)$	$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy)$	Premise
$[1](4)$	$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb$	$(1, *-)$
$[1,3](5)$	Fb	$(3, 4)$
$[1,3](6)$	$\forall x(\phi bx \rightarrow Fx)$	$(3, 5)$
$[2](7)$	$\forall x(\phi bx \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fc$	$(2, *-)$
$[1,2,3](8)$	Fc	$(3, 6, 7)$
$[1,2](9)$	$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fc$	$(3, 8, \rightarrow +)$
$[1,2](10)$	$\phi^* ac$	$(9, *+)$

So minimal SAL does not suffer from the same problem that plagues weak ancestral logic.

But minimal SAL is still a very weak logic. The transitivity of the ancestral can be proven in minimal SAL because it can be proven in second-order logic without any appeal to comprehension. But the proof of

[16] It is here, of course, that the additional expressive power provided by the presence of free second-order variables makes itself felt: No such rule could possibly be formulated in a purely first-order language.

[17] Thanks to Stewart Shapiro for the allusion.

[18] I have not included all the steps that would be required to make this argument formally precise, only enough to make it clear that it could be.

theorem (124) of *Begriffsschrift*:

$$\phi^{*}ab \wedge \forall x \forall y \forall z(\phi xy \wedge \phi xz \rightarrow y = z) \wedge \phi ac \rightarrow \phi^{*=}ac,$$

breaks down, as a little experimentation will show. The reason is that the proof requires Π^1_1 comprehension, and there is nothing in minimal SAL that gives us the power of Π^1_1 comprehension.[19]

How are we to get that power without second-order quantifiers? Easily. There are no explicit comprehension axioms in the formal systems of *Begriffsschrift* and *Grundgesetze*. Rather, Frege has a rule of substitution: Given a theorem of the form $\ldots F \ldots$, infer $\ldots \phi \ldots$, for any formula ϕ (subject to the usual sorts of restrictions). The substitution rule is, as is well-known, equivalent to comprehension.[20] What we need here is thus a rule of substitution: Suppose we have derived *A* from the premises in Γ, and let $A_{F/\phi}$ be the result of replacing all occurences of *F* in *A* by the formula ϕ (subject to the usual sorts of restrictions, again). Then, if *F* is not free in Γ, we may infer $A_{F/\phi}$. As a special case, of course, if *A* is provable (from no assumptions), then we may infer any substitution instance of it. And, given this rule, theorem (124) of *Begriffsschrift* can now be proven. See the appendix for the proof.

The substitution rule is clearly sound given the semantics sketched above, but it is another question whether we should regard it as *justified* and, if so, on what sort of ground. This question will be considered below, in section 7.

We thus have two kinds of systems: There are the systems without the substitution rule—minimal schematic logic and minimal schematic ancestral logic—and there are the systems with the substitution rule—what we may call full schematic logic and full schematic ancestral logic. In fact, there are further distinctions to be drawn, since the substitution rule can be restricted in various ways. We might, for example, require the formula that replaces *F* not to contain *. Adding this rule to minimal schematic ancestral logic would give us the effect of predicative comprehension, so we might call the resulting system *predicative* schematic ancestral logic.

4 Schemata in Schematic Logic: A Digression

One major reason second-order languages are so appealing is that principles that have to be formulated, in first-order languages, as axiom-schemata can be formulated in second-order languages as single axioms. It would be nice if schematic languages had a similar appeal, if, for example, the axiom of separation could be expressed by the single axiom:

$$\text{(Sep)} \quad \forall z \exists y \forall x(x \in y \equiv x \in z \wedge Fx),$$

[19] Zoltan Gendler-Szabó has proved that comprehension is, in fact, required. It is also required for the proof of the famous theorem (133):

$$\phi^{*=}ab \wedge \phi^{*=}ac \wedge \forall x \forall y \forall z(\phi xy \wedge \phi xz \rightarrow y = z) \rightarrow \phi^{*}bc \vee b = c \wedge \phi^{*}cb,$$

which thus cannot be proven in minimal SAL, either.

[20] Substitution implies comprehension: Trivially, we have $\forall x(Fx \equiv Fx)$, so existential generalization yields: $\exists G \forall x[Gx \equiv Fx]$; by substitution: $\exists G \forall x[Gx \equiv \phi x]$, for each formula ϕ. The proof of the converse is messier but not difficult: It is by induction on the complexity of the formula $\ldots F \ldots$.

from which its various instances could then be inferred by substitution. But it can't be, not if we characterize the rule of substitution as we did above. A *theory*, after all, is a set of formulae, and its theorems are the formulae that are deducible from the sentences in that set.[21] Such a deduction would assume some of the sentences of the theory as premises and then derive a theorem from them. If (Sep) is taken as a premise in a deduction, however, the variable *F* will obviously be free in that premise, whence it cannot be substituted for. Indeed, it is easy to see that (Sep) does not imply all (or even most) other instances of separation, not if 'implies' is defined as it was above. So it is a very good thing that they cannot all be deduced from it.

There is really a more basic problem here: I've yet to say what it might mean to assert something like (Sep); I've yet, that is, to say what the truth-conditions of (Sep) are. One might reasonably want to deny that (Sep) *has* truth-conditions: Since it contains a free second-order variable, one cannot speak of it as being true or false absolutely but only as being true or false under this or that assignment of a value to *F*. But there is an alternative: One can give free second-order variables the so-called 'closure interpretation', effectively taking (Sep) to be true just in case its universal closure is true. We do not actually need to consider the universal closure, of course, for we can define truth for formulae of schematic languages directly: A formula is true if, and only if, it is true under all assignments to its free variables.

This definition of truth would solve our problem concerning separation. Unfortunately, however, it brings a whole host of other problems with it.[22] To make further progress, we need to distinguish two sorts of assumptions that occur in argument. Sometimes, one makes an assumption 'for the sake of argument'. For example, one might assume the antecedent of a conditional and try to prove its consequent in order to prove a conditional. So, for example, if one were trying to prove $F0 \rightarrow G0$, one might begin by assuming $F0$: "Suppose 0 is blurg", one might say. One is not to continue with, "So 0 is odd. And prime. And, for the matter, even." One is not, that is to say, expect to be understood as having assumed that zero has every property there is. If we call an assumption made 'for the sake of argument' a *supposition*, then what has been shown is that suppositions are not to be understood in terms of their universal closures, the reason being that a supposition is relevantly like the antecedent of a conditional: No-one would suppose that $F0 \rightarrow G0$ should be understood as $\forall F(F0) \rightarrow \forall G(G0)$, however tempted they were by the thought that it should be understood as $\forall F \forall G(F0 \rightarrow G0)$.

But this observation does not make the closure interpretation of (Sep) any less available. The reason is that suppositions are mere tools of argument. They are not put forward as true in their own right. In a sense, that is obvious, since one sometimes makes a supposition only for *reductio*, but I am suggesting something stronger, that suppositions are not even *assumed* to be true: That is not the role they play in argument. Still, it ought nonetheless to be possible to assume that something is true, if only to investigate its consequences.

[21] Sometimes the term 'theory' is used in a different sense—a theory is a deductively closed set of sentences, and its theorems are just the members of that set—but that sense is not relevant here.

[22] Shapiro considers a logic $L2K-$, which is similar to the systems we have been discussing. Free second-order variables are given the closure interpretation. It turns out, however, to be surprisingly difficult to define a notion of implication with respect to which any reasonable set of deductive principles is sound (Shapiro, 1991, p. 81). In the end, Shapiro does define such a notion, but it is not really consistent with the closure interpretation. On Shapiro's definition, Fx does not imply Gx. But if the former really means $\forall F(Fx)$ and the latter really means $\forall G(Gx)$, it should. If one wanted to say that it is perhaps best if Fx doesn't imply Gx, I'd happily agree, but that intuition isn't consistent with the closure interpretation.

Let us reserve the term *hypothesis* for an assumption of this kind. Then there is no bar to our understanding hypotheses in terms of their universal closures.

Formally, then, we distinguish within an inference's premises between its suppositions and its hypotheses. An inference is valid if every interpretation that makes all of its suppositions and hypotheses true—where truth for hypotheses is understood in terms of the closure interpretation—also makes its conclusion true. The distinction must be tracked in the proof-theory as well, and rules of inference that discharge premises, such as *reductio* and conditional proof, will need some modification: The premise discharged must be a supposition, not an hypothesis. But the crucial observation, for our purposes, is that the rule of substitution can now be relaxed: One can infer $A_{F/\phi}$ from A so long as F is not free in any supposition on which A depends; it may be free in hypotheses on which A depends. This rule is clearly sound: Since, semantically speaking, hypotheses are treated as if they were universally closed, it is as if there aren't any free variables in the hypotheses, at least as far as the definition of implication is concerned.

There are other technical issues we could discuss, but let me set them aside: How they are resolved does not really bear upon the philosophical issue that opened this discussion.[23] The question with which we started was whether we can regard (Sep) as a formulation of separation. The answer is that we can if we regard a theory as a set of *hypotheses* from which theorems are to be deduced. As an *hypothesis*, (Sep) does imply all other instances of separation, including those containing other free variables.

One might wonder why I chose separation as my example rather than induction, since the same issues will, of course, arise with respect to induction in the context of schematic logic. They do not, however, arise in the context of *ancestral* logic, since we can formulate induction in ancestral logic as the single sentence:

$$\forall z *^{=}_{xy} (y = Sx)(0, z).$$

As we shall see, this point can be generalized: Any principle expressed in first-order logic by an axiom schema can be expressed by a single sentence in Arché logic, to be introduced next.

5 Arché Logic

The methods used in section 3 allowed us to transcribe Frege's explicit definition of the ancestral into schematic logic. A brief review will reveal, however, that the methods used presume only that the formula defining the ancestral is Π^1_1: They presume nothing about what formula it is. We can thus generalize that construction.

[23] Of course it matters that they *can* be resolved. I'll leave that as an exercise. One nice thing to do is add structural rules that allow a hypothesis freely to be converted to a supposition and a supposition to be converted to a hypothesis so long as none of its free variables are free in any of the suppositions on which it depends.

The distinction I am drawing here is very close to the distinction between rules of inference and rules of deduction that I drew, for an ostensibly quite different purpose, in (Heck, 1998). In fact, however, the formal situation is almost identical, since modal formula are there interpreted as if they were always proceeded by a universal quantifier over accessible worlds. The techniques developed there can therefore be used here.

Consider first the simplest case. Let $\phi_x(Fx, y)$ be an arbitrary formula containing no free variables other than those displayed. In standard second-order logic, we can explicitly define a new predicate $\mathcal{A}_\phi$ as follows:

$$\mathcal{A}_\phi(y) \equiv \forall F \phi_x(Fx, y).$$

The definition is licensed, in effect, by Π^1_1 comprehension, which guarantees that $\mathcal{A}_\phi$ exists. But the trick used with the ancestral can also be used here. We have an introduction rule, ($\mathcal{A}_\phi$+):

$$\phi_x(Fx, y) \vdash \mathcal{A}_\phi(y),$$

where *F* again may not be free in any premises on which $\phi_x(Fx, y)$ depends, and an elimination rule, ($\mathcal{A}_\phi$−):

$$\mathcal{A}_\phi(y) \vdash \phi_x(Fx, y).$$

In effect, these rules define $\mathcal{A}_\phi(y)$ as equivalent to $\forall F \phi_x(Fx, y)$ without using an explicit universal quantifier to do so. The extension of the new predicate $\mathcal{A}_\phi$ is then the set of all those *y* such that $\phi_x(Fx, y)$ is true for every assignment to the *F*.

More generally, we allow more than one predicate variable to occur in ϕ; we allow the predicate variables to be of various adicities; and we allow additional free first-order variables, in which case what is defined is a relation rather than just a predicate. So, in general, ϕ may be of the form $\phi_{x_1 \ldots x_{max_{1 \leq i \leq n}(k_i)}}(F_1(x_1, \ldots, x_{k_1}), \ldots, F_n(x_1, \ldots, x$ and, simplifying notation, we introduce a new predicate $\mathcal{A}_\phi(\overline{y})$ subject to the rules:

$$\phi_{\overline{x}}(\overline{F}, \overline{y}) \vdash \mathcal{A}_\phi(\overline{y})$$

$$\mathcal{A}_\phi(\overline{y}) \vdash \phi_{\overline{x}}(\overline{F}, \overline{y})$$

It is a more serious question whether we wish to allow additional free *second-order* variables to occur in ϕ, in which case these methods would allow us to define what Frege would have called a relation of 'mixed level', subject to the rules:

$$\phi_{xy}(Fx, Gy, z) \vdash \mathcal{A}_\phi y(Gy, z)$$

$$\mathcal{A}_\phi y(Gy, z) \vdash \phi_{xy}(Fx, Gy, z)$$

To take this step would force an expansion of the language to allow predicate variables to occur as arguments of predicates of mixed level. Such a step might reasonably be regarded as momentous or, at least, as involving new ideas. Fortunately, we shall not need this extension here. I mention only because it is natural and could allow one to motivate logics of greater stength than the ones we shall be considering.

Call what was just described the *scheme of schematic definition*. It allows us to transcribe what we would normally regard as an explicit definition of a new predicate or relation in terms of a Π^1_1 formula into schematic logic: If we can prove $\psi_x(Fx, a)$, then we can, in standard second-order logic, use universal generalization to conclude that $\forall F \psi_x(Fx, a)$ and so that $\mathcal{A}_\psi(a)$; if we have $\mathcal{A}_\psi(a)$, then by definition, $\forall F \psi_x(Fx, a)$ and so, by universal instantiation, $\psi_x(Fx, a)$. But the explicit definition of $\mathcal{A}_\psi(a)$ in terms of

$\forall F\psi_x(Fx, a)$ simply mediates the transitions between $\psi_x(Fx, a)$ and $\mathcal{A}_\psi(a)$. It can be eliminated in favor of an schematic definition of $\mathcal{A}_\psi(a)$ in terms of those same transitions.

If we add the scheme of schematic definition to (minimal) SL, we thus get a system in which new predicates co-extensional with Π^1_1 formulae can be introduced by schematic definition. Here again, however, the scheme of schematic definition is, by itself, deductively very weak: To exploit its power, we need a rule of substitution. We thus have three sorts of systems, depending upon the strength of the substitution principle we assume. Call the system without substitution *minimal Arché logic* (minimal AL). *Predicative Arché logic* contains a restricted substitution principle: The formula replacing *F* may not contain any of the new predicates $\mathcal{A}_\phi$. *Full Arché logic* allows unrestricted substitution.[24] These systems obviously include minimal, predicative, and full schematic ancestral logic, respectively, and the logical strength of miminal, predicative, and full AL should be close to that of second-order logic with no comprehension, predicative comprehension, and Π^1_1 comprehension, respectively. But it should be equally clear that the language in which these systems are formulated has *nothing like* the expressive power of a second-order language.

The axiom scheme of separation can be expressed in full AL by a single axiom. Let $\sigma_x(Fx, w)$ be the formula:

$$\forall z\exists y\forall x(x \in y \equiv x \in z \wedge Fx) \wedge w = w.$$

Then the scheme of schematic definition gives us a new predicate $\mathcal{A}_\sigma(w)$ subject to the rules:

$$\forall z\exists y\forall x(x \in y \equiv x \in z \wedge Fx) \wedge w = w \vdash \mathcal{A}_\sigma(w)$$

$$\mathcal{A}_\sigma(w) \vdash \forall z\exists y\forall x(x \in y \equiv x \in z \wedge Fx) \wedge w = w$$

or equivalently:

$$\forall z\exists y\forall x(x \in y \equiv x \in z \wedge Fx) \vdash \forall w\mathcal{A}_\sigma(w)$$

$$\forall w\mathcal{A}_\sigma(w) \vdash \forall z\exists y\forall x(x \in y \equiv x \in z \wedge Fx)$$

So separation is expressed by the single sentence: $\forall w\mathcal{A}_\sigma(w)$. A similar technique plainly applies to any axiom schema.[25]

The scheme of schematic definition also allows us to define new predicates that are co-extensional with Σ^1_1 formulae. Let $\phi_x(Fx, a)$ be a formula. We want to define a new predicate that is equivalent to $\exists F\phi_x(Fx, a)$.

[24] In full AL, the elimination rule ($\mathcal{A}_\phi-$) effectively takes the form

$$\mathcal{A}_\phi(a) \vdash \phi_x(B(x), a),$$

where $B(x)$ is an arbitrary formula (subject to the usual restrictions). In predicative AL, $B(x)$ is not permitted to contain new predicates of the form $\mathcal{A}_\phi$.

[25] A more elegant way to proceed is to extend the scheme of schematic definition to allow a new zero-place predicate (i.e., a sentential variable) to be defined in terms of a formula $\phi_x(F_1x, \ldots, F_nx)$, in which case we have:

$$\phi_x(F_1x, \ldots, F_nx) \vdash \mathcal{A}_\phi$$

$$\mathcal{A}_\phi \vdash \phi_x(F_1x, \ldots, F_nx)$$

Then separation is expressed by the zero-place predicate thus defined when we take $\phi_x(Fx)$ to be (Sep).

But $\exists F\phi_x(Fx,a)$ is equivalent to $\forall F\neg\phi_x(Fx,a)$, so we may use the scheme of schematic definition to introduce a new predicate $\mathcal{A}_{\neg\phi}$, subject to the rules:

$$\neg\phi_x(Fx,a) \vdash \mathcal{A}_{\neg\phi}(a)$$
$$\mathcal{A}_{\neg\phi}(a) \vdash \neg\phi_x(Fx,a)$$

So $\mathcal{A}_{\neg\phi}(a)$ is equivalent to $\forall F\neg\phi_x(Fx,a)$. We now regard $\neg\mathcal{A}_{\neg\phi}(a)$ as degeneratively of the form $\psi_x(Fx,a)$ and introduce a new predicate $\mathcal{A}_{\neg\mathcal{A}_{\neg\phi}}$, which I shall write: $\mathcal{A}^{\phi}$, subject to the rules:

$$\neg\mathcal{A}_{\neg\phi}(a) \vdash \mathcal{A}^{\phi}(a)$$
$$\mathcal{A}^{\phi}(a) \vdash \neg\mathcal{A}_{\neg\phi}(a)$$

So $\mathcal{A}^{\phi}(a)$ is equivalent to $\forall F(\neg\mathcal{A}_{\neg\phi}(a))$, that is, to $\neg\mathcal{A}_{\neg\phi}(a)$, that is, to $\neg\forall F\neg\phi_x(Fx,a)$ and so to $\exists F\phi_x(Fx,a)$, as wanted.

This argument obviously depends upon our allowing predicates defined using the scheme of schematic definition to appear in formulae used to define yet further new predicates using that same scheme. Formally speaking, we could consider restricting the scheme so as not to allow such iteration. Such a restriction would correspond to our not allowing free variables in the comprehension scheme. But this restriction has no motivation in the context of this investigation. The scheme of schematic definition formalizes a certain mode of *concept-formation*. Once one has used it to form a certain concept, one has that concept, and there is simply no reason one cannot iterate the process of concept-formation in the way we have allowed.

That said, there is a more elegant way to define predicates that are equivalent to Σ^1_1 formulae. As we have seen, the scheme of schematic definition, as currently formulated, in effect characterizes $\mathcal{A}_{\phi}$ in terms of introduction and elimination rules that mirror those for the universal quantifier that appears in its explicits second-order definition. That suggets that we should characterize $\mathcal{A}^{\phi}$ in terms of introduction and elimination rules that mirror those for the existential quantifier that appears in *its* explicit second-order definition. So we may take the introduction rule for $\mathcal{A}^{\phi}$, ($\mathcal{A}^{\phi}$+), to be:

$$\phi_x(A(x),a) \vdash \mathcal{A}^{\phi}(a),$$

where, in this case, variables free in $A(x)$ may be free in premises on which $\phi_x(A(x),a)$ depends. In full AL, $A(x)$ may be any formula; in predicative AL, it may not contain schematically defined predicates; in minimal AL, it must be an atomic formula.

The elimination rule, ($\mathcal{A}^{\phi}-$), is more complex, but only because the elimination rule for the existential quantifier is itself more complex, involving as it does the discharge of an assumption. Suppose we have derived a formula *B* from formulae in some set Δ together with $\phi_x(A(x),a)$, where none of the free variables occurring in $A(x)$, other than *x* itself—which is actually bound in $\phi_x(A(x),a)$—occur free in *B* or in Δ. Suppose further that we have derived $\mathcal{A}^{\phi}(a)$ from the formulae in some set Γ. Then we may infer *B*, discharging $\phi_x(A(x), a)$, so that *B* depends only upon Γ and Δ.[26] Symbolically:

$$\begin{array}{ccc} [\phi_x(A(x),a)] & \Delta & \Gamma \\ \ddots & \vdots & \vdots \\ & B & \mathcal{A}^{\phi}(a) \\ & \ddots & \vdots \\ & & B \end{array}$$

This rule simply parallels the relevant instance of the usual elimination rule for the second-order existential quantifier, except that we have replaced $\exists F \phi_x(Fx, a)$ with $\mathcal{A}_\phi(a)$. It is convenient to expand the scheme of schematic definition to allow schematic definitions of this form, too, since doing so adds no additional strength to the logic.

I intend the term 'scheme of schematic definition' to be taken seriously: I propose to regard the introduction and elimination rules ($\mathcal{A}_\phi$+) and ($\mathcal{A}_\phi-$) as *defining* the new predicate $\mathcal{A}_\phi$ and therefore regard the rules themselves as effectively self-justifying, since they are consequences of (because components of) a definition. In particular, then, I am proposing that we should regard (*+) and (*–)[27] as defining the ancestral. Perhaps that would be a reason to regard the ancestral as a logical notion and to regard these rules are logical rules. I am not sure, because I am not sure what the word 'logical' is supposed to mean here.[28] But the crucial issue for the neo-logicist is epistemological. The proof of Frege's Theorem makes heavy use of the ancestral and of inferences of the sort (*+) and (*–) describe. A neo-logicist must therefore show that she is entitled both to a grasp of the concept of the ancestral and to an appreciation of the validity of (*+) and (*–), and this entitlement must be epistemologically innocent in the sense that it does not itself import epistemological presuppositions that undermine the neo-logicist project: It must not, for example, presuppose a grasp of the concept of finitude, and it is a common complaint that our grasp of the concept of the ancestral presupposes precisely that. But if we regard the ancestral as schematically defined by (*+) and (*–), we may dismiss this complaint.

To be sure, one cannot simply introduce a new expression and stipulate that it should be subject to whatever introduction and elimination rules one wishes: Inconsistency threatens, as Arthur Prior famously showed (1960). A complete defense of the position I am developing here would thus have to contain an answer to the question when such stipulations are legitimate,[29] and to many others besides. But my purpose here is more modest. I am trying to argue that a certain position is available and worth considering. Whether it is true is a question for another day.

[26] If one wants to make the distinction between suppositions and hypotheses here, then $\phi_x(A(x), a)$ must be a supposition.

[27] In schematic ancestral logic, * is an operator, and so its logic can be characterized by the pair of rules (*+) and (*-). In Arché logic, there is no such operator. Rather, we have to define the ancestral of each relation separately, using the scheme of schematic definition. But this point does not affect the present discussion, so I shall ignore it.

[28] At the end of his life, George Boolos claimed no longer to understand the question whether second-order logic is logic, stated so baldly.

[29] For some recent discussion, see (Hale and Wright, 2000).

6 Frege's Theorem

If we are to prove Frege's Theorem in some form of schematic logic, we must be able to formalize HP in schematic logic. It is quite easy to do so. HP is, of course, neither a definition nor an instance of comprehension, but the techniques developed above may nonetheless be applied to it: We may represent HP as a pair of rules. Let $Fx \approx^{Ryz}_{xyz} Gx$ abbreviate:

$$\forall x \forall y \forall z \forall w(Rxy \wedge Rzw \rightarrow x = z \equiv y = w) \wedge \forall x(Fx \rightarrow \exists y(Rxy \wedge Gy)) \wedge \forall y(Gy \rightarrow \exists x(Rxy \wedge Fx)),$$

so that $Fx \approx^{Ryz}_{xyz} Gx$ says: *R* correlates the *F*s one-one with the *G*s. Then the introduction rule (N+) is easy enough to state:

$$Fx \approx^{Rxy}_{xyz} Gx \vdash Nx : Fx = Nx : Gx$$

The elimination rule (N–) is more complicated, but it simply parallels ($\mathcal{A}^{\phi}-$):

$$\begin{array}{ccc} \Delta \quad & [Fx \approx^{Rxy}_{xyz} Gx] & \Gamma \\ \quad \ddots & \vdots & \vdots \\ & B & Nx : Fx = Nx : Gx \\ & \ddots & \vdots \\ & & B \end{array}$$

That is: If we have derived a formula *B* from assumptions in some set Δ together with the assumption that $Fx \approx^{Rxy}_{xyz} Gx$ and we have derived $Nx : Fx = Nx : Gx$ from the assumptions in some set Γ (with *R* free in neither Δ nor *B*), then we may infer *B*, discharging $Fx \approx^{Rxy}_{xyz} Gx$. *Arché arithmetic* is full Arché logic plus these two rules.

I should emphasize before continuing that I am *not* claiming that (N+) and (N–) schematically define the cardinality operator. I am not even claiming that the possibility of formulating HP in schematic logic should do anything to ease any concerns one might have had about its epistemological status. My point here concerns only the logic needed for the proof of Frege's Theorem.

Frege's definitions of arithmetical notions can all be formalized in Arché arithmetic. Zero may be defined in the usual way:

$$0 = Nx : x \neq x$$

Frege's definition of predecession becomes an schematic definition of a new relation-symbol *P* subject to the rules (P+):

$$\exists y[Nx : Gx = b \wedge Gy \wedge Nx : (Gx \wedge x \neq y) = a] \vdash Pab$$

and (P–):

$$\begin{array}{ccc} \exists y[Nx : Gx = b \wedge Gy \wedge Nx : (Gx \wedge x \neq y) = a] & \Delta & \Gamma \\ \ddots & \vdots & \vdots \\ & B & Pab \\ & \ddots & \vdots \\ & \multicolumn{2}{c}{B} \end{array}$$

As usual, G must not be free in Δ or in B. The weak ancestral of this relation is schematically defined as subject to the two rules:

$$(\mathrm{P}*+) \quad \forall x(Pax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fb \vdash P^*ab$$

$$(\mathrm{P}*-) \quad P^*ab \vdash \forall x(Pax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge Pxy \rightarrow Fy) \rightarrow Fb,$$

subject to the usual restrictions. The definition of the concept of natural number is then:

$$\mathbb{N}n \equiv P^*0n \vee 0 = n.$$

Frege's proofs of axioms of arithmetic can then be formalized straightforwardly.

We may take the axioms of arithmetic to be as follows:[30]

1. $\mathbb{N}0$
2. $\mathbb{N}x \wedge Pxy \rightarrow \mathbb{N}y$
3. $\forall x \forall y \forall z(\mathbb{N}x \wedge Pxy \wedge Pxz \rightarrow y = z)$
4. $\forall x \forall y \forall z(\mathbb{N}x \wedge \mathbb{N}y \wedge Pxz \wedge Pyz \rightarrow x = y)$
5. $\neg \exists x(\mathbb{N}x \wedge Px0)$
6. $\forall x(\mathbb{N}x \rightarrow \exists x(Pxy))$
7. $A(0) \wedge \forall x \forall y(\mathbb{N}x \wedge A(x) \wedge Pxy \rightarrow A(y)) \rightarrow \forall x(\mathbb{N}x \rightarrow A(x))$

For convenience, induction has been formulated as a schema. As we saw above, this can be avoided, but let us work with the schema, for simplicity.[31]

As in Frege arithmetic—second-order logic plus HP—axioms (1) and (2) follow easily from the definition of $\mathbb{N}$: (1) is immediate, and (2) follows from the transitivity of the ancestral. The proofs of axioms (3), (4), and (5) in Frege arithmetic all appeal to HP, but they use only predicative comprehension and so are easily formalized in Arché arithmetic.

[30] Arithmetic and multiplication can be defined using the scheme of schematic definition, so we need no special axioms governing them.

[31] Consider the formula: $F0 \wedge \forall x \forall y(\mathbb{N}x \wedge Fx \wedge Pxy \rightarrow Fy) \rightarrow Fa$. The scheme of schematic definition yields a new predicate $\mathcal{S}a$ subject to the two rules:

$$F0 \wedge \forall x \forall y(\mathbb{N}x \wedge Fx \wedge Pxy \rightarrow Fy) \rightarrow Fa \vdash \mathcal{S}a$$
$$\mathcal{S}a \vdash F0 \wedge \forall x \forall y(\mathbb{N}x \wedge Fx \wedge Pxy \rightarrow Fy) \rightarrow Fa$$

Then induction is: $\forall x(\mathbb{N}x \rightarrow \mathcal{S}x)$.

Axiom (7) is stronger than what the definition of the ancestral by itself delivers. Simple manipulations give us:

$$A(0) \wedge \forall x \forall y (A(x) \wedge Pxy \rightarrow A(y)) \rightarrow \forall x (\mathbb{N}x \rightarrow A(x)).$$

But (7) is stronger, since the second conjunct of its antecedent is weaker in virtue of its containing the conjunct $\mathbb{N}x$ in its antecedent. But the instances of (7) can be proven, as usual, by induction on $\mathbb{N}\xi \wedge A(\xi)$.[32]

Axiom (6) can be derived from axioms (3), (4), and (5), using only the following very weak consequence of HP (Boolos, 1998d):

$$(\text{Log}) \quad \forall x(Fx \equiv Gx) \vdash Nx : Fx = Gx.$$

The proof uses only Π^1_1 comprehension, so it can be formalized in full Arché logic. See the appendix for the details.

A form of Frege's Theorem can thus be proven in Arché arithmetic. Exactly how strong the resulting fragment of second-order arithmetic is, I do not know. It is a natural conjuecture that it is equivalent to second-order arithmetic with Π^1_1 comprehension.[33] But, at the very least, it is certainly stronger than first-order PA. The explicit definition of satisfaction for the language of first-order arithmetic is Π^1_1, so it can be converted into an schematic definition in Arché arithmetic. The usual induction will then establish the consistency of first-order PA, so Arché arithmetic is certainly a non-conservative extension of first-order PA.

7 Philosophical Considerations

A close examination of the proofs in the appendix will show that, if one regards predecession as primitive and subject to the rules (P+) and (P–), as suggested in section 2, then Frege's Theorem can be proven in full schematic ancestral logic. If one regards the ancestral too as primitive and subject to the rules (*+) and (*–), then Frege's Theorem can be proven in *predicative* schematic ancestral logic. Predicative systems are generally regarded as epistemologically innocent. So if both predecession and the ancestral could be regarded as primitive, the mentioned rules being analytic of these notions, the logic needed for the proof of Frege's Theorem would be epistemologically innocent.

But the question which notions are primitive does not seem to me to be the right question to ask here: It is too slippery. It is better, I think, to regard both predecession and its ancestral as defined by means of the scheme of schematic definition and to regard (P+), (P–), (P*+), and (P*–) as analytic on the ground that they are consequences of, because components of, those definitions. The proof of Frege's Theorem, in that case, needs full Arché logic. In particular, the proof needs the unrestricted rule of substitution. A philosopher with principled concerns about impredicativity might therefore be tempted to say—and might, indeed, long

[32] Even if we had convinced ourselves that $\mathbb{N}\xi$ could be taken as primitive, this argument, formalized in second-order logic, would still need Π^1_1 comprehension. That is another reason one should not expect to get by with much less if one is trying to derive the axioms of PA from HP. As it happens, however, if one is willing to forego induction and interpret a weaker theory, such as Robinson arithmetic, then one can do so in predicative second-order logic: See (Heck, 2007).

[33] Is there a natural extension of Arché arithmetic that admits Δ^1_2 comprehension? that is, one that is equivalent to predicative analysis?

have been wanting to say—that, however amusing the foregoing may be, it is largely beside the point if the question at issue is whether the logic required for the proof of Frege's Theorem is epistemologically innocent. The unrestricted substitution rule is impredicative, and the only question, really, was where the impredicativity would ultimately surface. The bump has been pushed around a fair bit, but the rug is no flatter now than it was before.

I disagree. The scheme of schematic definition allows us to introduce a new predicate $\mathcal{A}_\phi$ subject to the rules:

$$(\mathcal{A}_\phi+) \quad \phi_x(Fx, a) \vdash \mathcal{A}_\phi(a)$$

$$(\mathcal{A}_\phi-) \quad \mathcal{A}_\phi(a) \vdash \phi_x(Fx, a)$$

in the presence of the unrestricted substitution rule, the elimination rule is equivalent to:

$$(\mathcal{A}_\phi\text{sub}) \quad \mathcal{A}_\phi(a) \vdash \phi_x(B(x), a),$$

where $B(x)$ is now a formula rather than a variable. In particular, in the case of the ancestral, the elimination rule is equivalent, in the presence of unrestricted substitution, to:

$$(\phi * \text{sub}) \quad \phi^* ab \vdash \forall x(\phi ax \to B(x)) \wedge \forall x \forall y(B(x) \wedge \phi xy \to B(y)) \to B(b).$$

Such a rule can be understood in two ways.[34] One way takes the set of formulae $B(x)$ that can appear in the rule to be determined by reference to some fixed language. That is how axiom schemata, such as the induction scheme in PA:

$$A(0) \wedge \forall x(A(x) \to A(Sx)) \to \forall x A(x),$$

are usually understood. The induction scheme is usually regarded as abbreviating an infinite list of axioms, one for each formula of the language of arithmetic. When we consider expansions of the language of PA, then—say, the result of adding a truth-predicate *T*—that expansion does not, in itself, result in any new axioms' being added to the original theory. The sentence

$$T0 \wedge \forall x(Tx \to T(Sx)) \to \forall x Tx,$$

in particular, and other sentences containing *T*, do not automatically become axioms of the new theory, though such sentences do have the form of induction axioms. That is why adding a truth-predicate to PA, and even adding the Tarskian clauses for the truth-predicate, yields a conservative extension: One can't do much with the truth-predicate if it doesn't occur in the induction axioms.

Formally, of course, one can proceed how one likes, but this way of thinking of the induction scheme is not obviously best. Even if our theory of arithmetic is formulated in a first-order language, one would have thought the induction scheme should be regarded as one that *does*, as it were, automatically import

[34] See (Feferman, 1991) for discussion of, and applications of, this distinction. Regarding Feferman's historical remarks in §1.5, it is perhaps worth noting that the first accurate formulation of the sort of substitution rule that is needed here is due to Frege: See Rule 9 in section 48 of *Grundgesetze*. As Feferman notes, citing Church, such a rule is missing from *Begriffsschrift*.

new formulae of the appropriate form into our theory as our language expands. In ordinary, everyday mathematics, we do not so much as ask, as our language expands, whether the new instances of induction that become available should be accepted as true. To regard the induction scheme in this way, then, is to regard it as expressing an open-ended commitment to the truth of all sentences of a certain form, both those we can presently formulate and those we cannot. The rule ($\phi *$ sub) is to be understood in this way, too: It expresses an open-ended commitment to the validity of all inferences of a certain form.[35]

The unrestricted substitution rule thus expresses the open-ended nature of the commitments we undertake when we schematically define a new predicate *via* an instance of the scheme of schematic definition. And that, simply, is how the scheme is intended to be understood. One whose understanding of the ancestral is completely constituted by her grasp of the rules (ϕ^*+) and (ϕ^*-) would, it seems to me, be quite surprised to hear that these rules do not license the sort of inference required for the proof of theorem (124) of *Begriffsschrift*. I am not saying that it would be *incoherent* to refuse to accept that inference. I am simply saying that it would not be a natural reaction.

One might object that this justification of the substitution rule, if it is defensible at all, ought to apply just as well in the context of full second-order logic. In fact, I think it does so extend. In the case of second-order logic, however, there is another and more fundamental problem with which we must contend: We must explain the second-order quantifiers. Absent such an explanation, we do not so much as understand second-order languages, and the question how the substitution rule should be justified doesn't arise. Now, to understand the second-order universal quantifier, one must understand what it means to say that *all* concepts are thus-and-so. But to understand that sort of claim, or so it is often argued, one must have a conception of what the second-order domain comprises. One must, in particular, have a conception of (something essentially equivalent to) the full power-set of the first-order domain, and many arguments have been offered that purport to show that we simply do not have a definite conception of $\wp(\omega)$. It is not my purpose here to evaluate such arguments. Maybe they work, and maybe they do not. My purpose here is to identify an epistemologically relevant difference between second-order logic and Arché logic. Here it is: Since there are no second-order quantifiers in schematic languages, the problem of explaining the second-order quantifier simply does not arise in that context.

One might object that the problem arises nonetheless. The thought would be that our understanding of the introduction rule for the ancestral essentially involves just such a conception. How else are we to understand

$$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb,$$

as it occurs in the premise of the rule (*+), except as involving a tacit initial second-order quantifer? Does it not say, explicit quantifier or no, that all concepts *F* that are thus-and-so are so-and-thus? Does understanding that claim not require the disputed conception of the power-set? No, it does not. A better reading would be: *A* concept that is thus-and-so is so-and-thus. What understanding this claim requires is not a capacity to conceive of *all* concepts but simply the capacity to conceive of *a* concept: to conceive of an arbitrary concept, if you like. The contrast here is entirely parallel to that between arithmetical claims like $x + y = y + x$,

[35] I borrow the term 'open-ended' from Vann McGee (1997).

involving only free variables, and claims involving explicit quantification over all natural numbers. Hilbert famously argued that our understanding of claims of the former sort involves no conception of the totality of all natural numbers, whereas claims of the latter sort do, and that there is therefore a significant conceptual and epistemological difference between these cases. I am making a similar point about claims involving only free second-order variables as opposed to claims that quantify over concepts.

But one might insist that, nonetheless, if we do not have a definite conception of the full power-set—if, in particular, there is nothing in our understanding of free second-order variables that guarantees that they range over the full power-set of the first-order domain—then the meanings of the predicates we introduce by schematic definition will be radically underdetermined, at least. It was stipulated earlier that $\mathcal{A}_\phi(a)$ is true if, and only if, $\phi_x(Fx, a)$ is true for every assignment of a subset of the first-order domain to *F*. But why? If we have no conception of the full power-set, why not take the domain of the second-order variables to be smaller? Why not restrict it to the definable subsets of the domain? Surely none of the axioms and rules of Arché logic require the second-order domain to contain every subset of the first-order domain.

Obviously, there is a technical point here that is incontrovertible: The existence of non-standard models is a fact of mathematics, and a very useful one at that. But the philosophical significance of this technical point is not so obvious. It seems to me that there *is* something about the axioms and rules of Arché logic that requires the second-order domain to be *unrestricted*, and that is the crucial word. The difference between the standard model and the various non-standard models is to be found not in what the standard model *in*cludes but in what non-standard models *ex*clude: A non-standard model of necessity excludes certain concepts from the domain of the second-order variables. That, however, is incompatible with the nature of the commitments we undertake when we introduce a new predicate using the scheme of schematic definition. Those commitments are themselves *unrestricted* in the sense that we accept no restriction upon what formulae may replace $B(x)$ when we infer $\phi_x(B(x), a)$ from $\mathcal{A}_\phi(a)$. One might be tempted to object that, if so, we must somehow conceive of the totality of all such formulae in advance. But that would simply repeat the same error: No such conception of the totality of all formulae is needed; what is needed is just the ability to conceive of *a* formula—an arbitrary formula, if you like.

It would not be unreasonable to claim that, at this point, we are essentially at stalemate, although both sides have moves remaining. But that is enough for my purpose here. What I am trying to do is not to convince the reader of any particular position. I am trying, rather, to convince the reader of the *interest* of a certain position, namely, the position that full Arché logic is epistemologically innocent in whatever way such positions as the neo-logicist's need logic to be. It is no part of my position that full second-order logic is not epistemologically innocent in that sense. But it is my position that there is *enough of a difference* between Arché logic and second-order logic that it would not be unreasonable to regard them as epistemologically unequal in this same sense. I take myself to have accomplished that much. The resources deployed above in the defense of full Arché logic against the predicativist skeptic are not resources that are obviously sufficient to defend full second-order logic. Perhaps they can be built upon for that purpose. I don't necessarily say otherwise. But perhaps they cannot be.

The critical difference between Arché logic and second-order logic thus turns out to lie not so much in the logical principles that distinguish them but, rather, in the expressive power of the underlying languages. What we have seen is not just that the full deductive strength of second-order logic is not needed for the proof of Frege's Theorem: That has been known for some time. What we have seen is that not all of the expressive power of second-order languages is needed, either. Quine's view (1986) was that second-order quantification is not even a logical *notion*: If second-order variables range over sets, then second-order quantification is quantification over sets, and second-order logic is set-theory in sheep's clothing, quite independently of its proof-theoretic strength. Even if one interprets second-order quantifiers in terms of plurals, as suggested by George Boolos (1998e), however, one might have other reasons to suppose that plural quantifiers are non-logical constants (see e.g. Resnik, 1988), perhaps reasons connected with the expressive power of plural quantifiers. Again, it is no part of my view that plural quantifiers are *not* logical constants. What I am claiming is that the question whether axioms for arithmetic can be derived, purely logically, from HP does not depend upon how that issue might be resolved: The language of Arché arithmetic has a stronger claim to be a logical language than the language of second-order logic does.[36]

Appendix

Here, as above, I am not including all the steps that would be necessary to make the argument completely formal. In particular, standard first-order moves will be repressed for the most part and only briefly indicated where they are not. My intention is simply to make it clear that these results can be proven in the relevant systems.

Proof of *Begriffsschrift*, Theorem (124)

The proof is in full schematic ancestral logic. It could also, of course, be carried out in full Arché logic.

[36] This material was presented to the Mathematics Workshop at Arché, the AHRC Research Centre for the Philosophy of Logic, Language, Mathematics and Mind, at the University of St Andrews, in February 2005. Thanks to Arché for its support, which is much appreciated, and to Crispin Wright for arranging another visit. I thank everyone who attended for their comments, but special gratitude is due to Crispin and to Stewart Shapiro for their enthusiasm about this material, which is what convinced me to write it down.

Earlier versions of these ideas were presented in a graduate seminar given at Harvard University in Fall 2004. Thanks to the members of that seminar for their reaction. Thanks too to Øystein Linnebo for his comments on an early draft.

It should be obvious that my work owes a great deal to Crispin's. That is true not only of my work on Frege's philosophy of mathematics, but also of my work on vagueness and of my work on philosophy of language. But my debt to him is far greater than that. Although I first met Crispin in the summer of 1993, at a conference on philosophy of mathematics organized by Matthias Schirn, the proceedings of which were published as (Schirn, 1998), he had already been generous with his time, discussing philosophy over email with a distant graduate student. Since then, I have many times had the privilege of spending time with Crispin, whether in St Andrews or elsewhere, discussing philosophy, football, and our families, and I am honored now to call him a friend. He has been a reliable supporter, both of my work and of me, so much so that I am quite certain that my career would have been far different if not for his presence in my life.

Thank you, Crispin, for everything. And long live Arché.

$[1]$	(1)	$\phi^* ab$	Premise
$[2]$	(2)	ϕac	Premise
$[3]$	(3)	$\forall x \forall y \forall z(\phi xy \wedge \phi xz \rightarrow y = z)$	Premise
$[1]$	(4)	$\forall x(\phi ax \rightarrow Fx) \wedge \forall x \forall y(Fx \wedge \phi xy \rightarrow Fy) \rightarrow Fb$	$(1, *-)$
$[1]$	(5)	$\forall x(\phi ax \rightarrow \phi^{*=} cx) \wedge \forall x \forall y(\phi^{*=} cx \wedge \phi xy \rightarrow \phi^{*=} cy) \rightarrow \phi^{*=} cb$	$(4, subst)$
$[6]$	(6)	ϕax	Premise
$[2, 3, 6]$	(7)	$x = c$	$(2, 3, 6)$
$[2, 3, 6]$	(8)	$\phi^{*=} cx$	def $\phi^{*=}$
$[2, 3]$	(9)	$\forall x(\phi ax \rightarrow \phi^{*=} cx)$	$(6, 7)$
$[]$	(10)	$\forall x \forall y(\phi^{*=} cx \wedge \phi xy \rightarrow \phi^{*=} cy)$	transitivity
$[1, 2, 3]$	(11)	$\phi^{*=} cb$	$(5, 9, 10)$

The substitution rule, applied at line (5), is essential to this proof, which therefore collapses in minimal schematic ancestral logic. Since the substituted formula contains *, it cannot be replicated in predicative schematic ancestral logic either.

Proof of Axiom (6)

This proof is in full schematic ancestral logic plus the following restricted form of HP, which George Boolos dubbed **Log**:

$$\forall x(Fx \equiv Gx) \rightarrow Nx : Fx = Nx : Gx$$

We will need the 'roll-back theorem': $P^* ab \rightarrow \exists y(P^{*=} ay \wedge Pyb)$. Its proof is straightforward. We will show that axiom (6) follows from the other axioms of arithmetic, to which we freely appeal.

We start by proving Theorem (145) of *Grundgesetze*: $P^{*=} 0x \rightarrow \neg P^* xx$.

$[]$	(1)	$P^{*=} 0n \wedge \neg P^* 00 \wedge \forall x \forall z(P^{*=} 0x \wedge \neg P^* xx \wedge Pxz \rightarrow \neg P^* zz) \rightarrow \neg P^* nn$	Axiom (7)
$[]$	(2)	$P^* 00 \rightarrow \exists y(P^{*=} 0y \wedge Py0)$	roll-back
$[]$	(3)	$\neg \exists y(P^{*=} 0y \wedge Py0)$	Axiom (5)
$[]$	(4)	$\neg P^* 00$	$(2, 3)$
$[5]$	(5)	$P^{*=} 0x \wedge \neg P^* xx \wedge Pxz$	Premise
$[6]$	(6)	$P^* zz$	Premise
$[6]$	(7)	$\exists y(P^{*=} zy \wedge Pyz)$	roll-back
$[8]$	(8)	$P^{*=} zy \wedge Pyz$	Premise
$[5, 8]$	(9)	$x = y$	$5, 8,$ Axiom (4)
$[5, 8]$	(10)	$Pxz \wedge P^{*=} zx$	$5, 8, 9$
$[5, 8]$	(11)	$P^* xx$	$10,$ transitivity
$[5, 6]$	(12)	$P^* xx$	$11, [8] \exists -$
$[5]$	(13)	$\neg P^* zz$	$5, 12; [6] \neg +$
$[]$	(14)	$\forall x \forall z[P^{*=} 0x \wedge \neg P^* xx \wedge Pxz \rightarrow \neg P^* zz]$	$13, [5] \rightarrow +; \forall +$
$[]$	(15)	$P^{*=} 0n \rightarrow \neg P^* nn$	$1, 4, 14$

We now prove the existence of successor by proving: $P^{*=}0n \rightarrow P(n, Nx : P^{*=}xn)$. We shall need the following simple fact twice:

$$Fa \rightarrow P[Nx : (Fx \wedge x \neq a), Nx : Fx].$$

That is Theorem 102 of *Grundgesetze* and shall be cited as such. It follows immediately from (P+).

[]	(1)	$P^{*=}0n \wedge P(0, Nx : P^{*=}x0) \wedge$ $\forall y \forall z[P^{*=}0y \wedge P(y, Nx : P^{*=}xy) \wedge Pyz \rightarrow P(z, Nx : P^{*=}xz)] \rightarrow$ $P(n, Nx : P^{*=}xn)$	Axiom (7)
[]	(2)	$P^{*=}00 \rightarrow P[Nx : (P^{*=}x0 \wedge x \neq 0), P^{*=}x0)$	Gg 102
[]	(3)	$P[Nx : (P^{*=}x0 \wedge x \neq 0), P^{*=}x0)$	2, def $P^{*=}$
[]	(4)	$P^{*=}x0 \wedge x \neq 0 \rightarrow P^{*}x0$	def $P^{*=}$
[]	(5)	$P^{*}x0 \rightarrow \exists u(P^{*=}0u \wedge Pu0)$	roll-back
[]	(6)	$\neg\exists u(P^{*=}0u \wedge Pu0)$	Axiom (5)
[]	(7)	$\neg(P^{*=}x0 \wedge x \neq 0)$	4, 5, 6
[]	(8)	$\forall x[(P^{*=}x0 \wedge x \neq 0) \equiv x \neq x]$	7∀+
[]	(9)	$Nx : (P^{*=}x0 \wedge x \neq 0) = Nx : (x \neq x)$	8, **Log**
[]	(10)	$Nx : (P^{*=}x0 \wedge x \neq 0) = 0$	9, def 0
[]	(11)	$P(0, Nx : Nx : P^{*=}x0)$	9, 10

So that establishes the basis step. We now prove the induction step to complete the proof.

[12]	(12)	$P^{*=}0y \wedge P(y, Nx : P^{*=}xy) \wedge Pyz$	Premise
[]	(13)	$P^{*=}zz \rightarrow P[Nx : (P^{*=}xz \wedge x \neq z), Nx : P^{*=}xz]$	Gg 102
[]	(14)	$P[Nx : (P^{*=}xz \wedge x \neq z), Nx : P^{*=}xz]$	13, def $P^{*=}$
[12]	(15)	$z = Nx : P^{*=}xy$	12, Axiom (3)
[12]	(16)	$Nx : (P^{*=}xz \wedge x \neq z) = Nx : P^{*=}xy \rightarrow P(z, Nx : Nx : P^{*=}xz)$	14, 15
[12]	(17)	$\forall x[(P^{*=}xz \wedge x \neq z) \equiv P^{*=}xy] \rightarrow$ $Nx : (P^{*=}xz \wedge x \neq z) = Nx : P^{*=}xy$	**Log**
[12]	(18)	$\forall x[(P^{*=}xz \wedge x \neq z) \equiv P^{*=}xy] \rightarrow P(z, Nx : Nx : P^{*=}xz)$	16, 17

We now need only establish the antecedent of (18) to complete the proof. First, right-to-left:

[19]	(19)	$P^{*=}xy$	Premise
[19]	(20)	$P^{*=}xz$	12, 19, transitivity
[21]	(21)	$x = z$	Premise
[12, 21]	(22)	$P^{*=}xy \wedge Pyx$	12, 19, 21
[12, 21]	(23)	$P^{*=}0y \wedge P^{*}yy$	12, 22, transitivity
[12]	(24)	$x \neq z$	23, Gg 145; [12]¬+
[12]	(25)	$P^{*=}xy \rightarrow P^{*=}xz \wedge x \neq z$	20, 24; [19] → +

Now left-to-right:

[26]	(26) $P^{*=}xz \wedge x \neq z$		Premise
[26]	(27) $P^{*}xz$		def $P^{*=}$
[26]	(28) $\exists u(P^{*=}xu \wedge Puz)$		27, roll-back
[29]	(29) $P^{*=}xu \wedge Puz$		Premise
[12, 29]	(30) $u = y$		12, 29, Axiom (4)
[12, 29]	(31) $P^{*=}xy$		29, 30
[12, 26]	(32) $P^{*=}xy$		28, 31; [29]∃−
[12]	(33) $P^{*=}xz \wedge x \neq z \rightarrow P^{*=}xy$		32, [26] → +
[12]	(34) $\forall x[(P^{*=}xz \wedge x \neq z) \equiv P^{*=}xy]$		25, 33, ∀+
[12]	(35) $P(z, Nx : Nx : P^{*=}xz)$		18, 34
[]	(36) $\forall y \forall z[P^{*=}0y \wedge P(y, Nx : Nx : P^{*=}xy) \wedge Pyz \rightarrow P(z, Nx : Nx : P^{*=}xz)]$		35, [12] → +, ∀+

That completes the proof of the induction step. So we may conclude:

[] (37) $P^{*=}0n \rightarrow P(n, Nx : P^{*=}xn)$ 1, 11, 36

References

Avron, A. (2003). Transitive closure and the mechanization of mathematics. In Kamareddine, F., editor, *Thirty Five Years of Automating Mathematics*, pages 149–71. Kluwer Academic Publishers, New York.

Boolos, G. (1997). Is Hume's principle analytic? In Heck, R., editor, *Language, Thought, and Logic: Essays in Honor of Michael Dummett*, pages 245–261. Clarendon Press, Oxford.

Boolos, G. (1998a). Is Hume's principle analytic? In Boolos (1998b), pages 301–314.

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

Boolos, G. (1998c). On second-order logic. In Boolos (1998b), pages 37–53.

Boolos, G. (1998d). On the proof of Frege's theorem. In Boolos (1998b), pages 275–291.

Boolos, G. (1998e). To be is to be a value of a variable (or to be some values of some values of some variables). In Boolos (1998b), pages 54–72.

Boolos, G. and Heck, R. (1998). *Die Grundlagen der Arithmetik* §§82–83. In Boolos (1998b), pages 315–338.

Burgess, J. P. (2005). *Fixing Frege*. Princeton University Press, Princeton NJ.

Feferman, S. (1964). Systems of predicative analysis. *Journal of Symbolic Logic*, 29:1–30.

Feferman, S. (1984). Towards useful type-free theories, i. In Martin, R., editor, *Recent Essays on the Liar Paradox*, pages 237–288. Oxford University Press, New York.

Feferman, S. (1991). Reflecting on incompleteness. *Journal of Symbolic Logic*, 1991:1–49.

Hale, B. and Wright, C. (2000). Implicit definition and the a priori. In Boghossian, P. and Peacocke, C., editors, *New Essays on the A Priori*, pages 286–319. Clarendon Press, Oxford.

Heck, R. (1996). The consistency of predicative fragments of Frege's *Grundgesetze der Artithmetik*. *History and Philosophy of Logic*, 17:209–20.

Heck, R. (1998). That there might be vague objects (so far as concerns logic). *The Monist*, 81:277–99.

Heck, R. (2000). Counting, cardinality, and equinumerosity. *Notre Dame Journal of Formal Logic*, 41:187–209.

Heck, R. (2007). Ramified frege arithmetic. *Journal of Philosophical Logic*.

Linnebo, Ø. (2004). Predicative fragments of Frege arithmetic. *Bulletin of Symbolic Logic*, 10:153–74.

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

Prior, A. (1960). The runabout inference-ticket. *Analysis*, 21:38–9.

Quine, W. (1986). *Philosophy of Logic*. Harvard University Press, 2d ed edition.

Resnik, M. (1988). Second-order logic still wild. *Journal of Philosophy*, 85:75–87.

Schirn, M., editor (1998). *Philosophy of Mathematics Today*. Oxford University Press, Oxford.

Shapiro, S. (1991). *Foundations without Foundationalism: A Case for Second-order Logic*. Oxford University Press, Oxford.

Wright, C. (1983). *Frege's Conception of Numbers as Objects*. Aberdeen University Press, Aberdeen.

Wright, C. (2001). Is Hume's principle analytic? In *The Reason's Proper Study*, pages 307–332. Clarendon Press, Oxford.