Hume's Principle
exhumed

Submitted by: Ian Hulse
University College London
For the degree: Master of Philosophy

Abstract

This study is a critique of Fregean Platonism as advanced by Crispin Wright and Bob Hale. Fregean Platonism is founded upon an argument that purports to attain three objectives. The first of these is to explain why the natural numbers are not epistemologically suspect in the ways supposed by nominalists, i.e. to show how we can acquire the concepts required to form beliefs about the natural numbers, how we can refer to them and how we can know about them. The second objective is to demonstrate the existence of the natural numbers, construed as Fregean Objects. The third is to found number-theoretic logicism. This study is concerned with the success of the argument only with respect to the first two objectives. The argument makes crucial use of a certain sentence, Hume's Principle (HP). (HP) is the statement that the number of F's equals the number of G's if and only if the F's can be bijectively mapped to the G's (for any concepts F and G). Two readings of the argument are distinguished. The first takes (HP) to be a contextual explanation of the concept natural number. It is argued that the argument on its first reading fails, that (HP) cannot successfully explain the concept natural number. Derivatively, it fails to demonstrate the existence of the natural numbers. No firm conclusion is reached with respect to the argument on its second reading. It is argued that (HP) is apriori under Christopher Peacocke's analytic theory of the apriori and that this might be used to overcome objections to the analyticity (hence, apriority) of (HP) suggested by George Boolos.

We are possest of a precise standard by which we can judge of the equality and proportion of numbers; and according as they correspond or not to that standard, we determine their relations, without any possibility of error. When two numbers are so combin'd as that the one has always an unite answering to every unite of the other, we pronounce them equal; and 'tis for want of such a standard of equality in extension, that geometry can scarce be esteem'd a perfect and infallible science.

David Hume,
Treatise on Human Understanding,
I, III, para. 5.

Contents

Glossary of Named Theses

Throughout the course of this study a number of claims, theses and sentences are named and thereafter referred to by name. For the readers convenience they are collected together here.

(CPS) An adequate account of the meaning of a sub-sentential expression must make clear the contribution it makes to the meaning of the sentences in which it occurs.

(CPR) Syntactic categories are explanatorily prior to ontological categories.

(RT) All sub-sentential expression appearing in true sentences contribute to the truth-conditions of those sentences by referring.

(HP) Two concepts have the same number if and only if they can be put into one-one correspondence, i.e. if and only if there is a bijective mapping from the things that fall under one to the things that fall under the other.

In logical notation (HP) is written: $\forall F \forall G(Nx{:}Fx = Nx{:}Gx \leftrightarrow \exists R(F\ 1\text{-}1_R\ G))$.

(ET) A (Platonist) grasp of instances of the left hand side of the sentence "$\ulcorner Nx{:}Fx = Nx{:}Gx \urcorner$ means that $\ulcorner \exists R(F\ 1\text{-}1_R\ G) \urcorner$" can be derived from a (Nominalist) grasp of the corresponding instances of the right hand side.

(SS) The left hand sides of instances of "$\ulcorner Nx{:}Fx = Nx{:}Gx \urcorner$ means that $\ulcorner \exists R(F\ 1\text{-}1_R\ G) \urcorner$" are semantically complex.

(IT) The transition from L to L* is semantically non-revisionary, i.e. all symbols of L are interpreted in L* as they were in L.

(OP) The ontological commitments in both the left and right hand sides of instances of "$\ulcorner Nx{:}Fx = Nx{:}Gx \urcorner$ means that $\ulcorner \exists R(F\ 1\text{-}1_R\ G) \urcorner$" are those that are explicit on instances of the left hand side.

(IC) A content is known apriori if it is known to be true in the actual world independently of any particular experience.

Chapter 1

Fregean Platonism

1.1 Introduction

Platonism is an ontological thesis. It asserts that amongst the "ultimate furniture of the universe" we will find more than one type of entity. Some entities will be *concrete*, others *abstract*. Abstract entities are things like mathematical entities (numbers, sets, etc.), words, sentences, poems and plays (in the type, as opposed to the token, sense): concrete entities are things like figs, rabbits, the Milky Way, words and sentences etc. (in the token sense), ideas and people. In many cases, the Platonist thesis is augmented by the addition of a further claim, that at least some abstract entities enjoy their existence independently of any thought or talk about them. It is this claim that motivates the yin to Platonism's yan: Nominalism, the thesis that all entities that either have existed, will exist or exist now are concrete.[1] It is not difficult to see why Nominalism might seem attractive. Abstract entities which exist independently of any minds and which appear to exert no influence on us whatsoever seem to be the best possible example of the ineffable. Hence, the Nominalist sounds like the voice of reason when she states that, since we can't know anything about abstract entities, we might as well purge them from our ontology. I call a Nominalist who founds their Nominalism on this ground a *Conservative Nominalist*, for she claims, not that we are unable to acquire beliefs about abstracta, but that these beliefs could never acquire the exalted status of knowledge. She accepts that we can grasp the concepts required to form beliefs about abstracta but denies that we could ever be certain that the beliefs so formed were true. More radically, Nominalism can be motivated by the worry that, since abstract entities exert no influence on us, we cannot even talk about them. We cannot refer to them or acquire the concepts necessary to forming beliefs about them. I call forms of Nominalism motivated by these doubts *Radical Nominalism*.[2]

---

[1] This is contemporary Nominalism. It should not be confused with traditional Nominalism, the thesis that all existing things have been/will be/are particulars (as opposed to universals).

[2] Of course, not everyone agrees with the nominalistic assessment. Plato believed that knowledge of the abstract realm (of the forms) was the result of intellectual intuition, whereby the soul remembered things it had forgotten at birth (see Plato's *Phaedo*, *Meno* and *Republic*). More recently, Gödel has claimed that something akin to sense-perception allows us to survey the abstract realm (Gödel, 1944), Parsons has attempted to justify this claim (Parsons, 1979) and Maddy has argued that we can, and do, perceive some abstract entities in much the same way that we perceive macroscopic concrete entities (Maddy, 1980). Nonetheless, since Benacerraf published "Mathematical Truth" (Benacerraf (1973)), received opinion has had it that satisfactory epistemologies and the existence of abstract entities are not the best of friends (though see Brown (1990) for an interesting objection to Benacerraf's claim).

This study is a critique of an argument proposed by Crispin Wright. It is a bold argument and an elegant one. Its aim is threefold: to dispel the aforementioned epistemological doubts surrounding the existence of natural numbers; to show that Platonism is true by showing that the natural numbers exist as mind-independent objects and to show that some form of Logicism, the thesis that mathematics is, in some sense, reducible to logic, can be upheld, at least with respect to number-theory. The steps taken to achieve these aims result in a certain doctrine. It is called Fregean Platonism. It is an absorbing doctrine which, at first glance, is really quite convincing. The purpose of this study is to ascertain whether that *prima facie* impression is accurate, at least with respect to the Fregean's first two claims. That is to say, I aim to evaluate Fregean Platonism's claim to dispel the epistemological worries that motivate Nominalism with respect to the natural numbers and to evaluate its claim to demonstrate the objectual existence of those numbers. I will nowhere address the third claim, that some species of Logicism can be upheld with respect to number-theory.

1.2 Fregean Platonism and the Context Principle

Fregean Platonism finds its first recent expression in Wright's *Frege's Conception of Numbers as Objects* (Wright (1983)) but it has also received a spirited defense at the hands of Hale (Hale (1987). Perhaps its most outspoken critics have been Field (1984), Dummett (1991) and Boolos (1990). The doctrine is derived from a suggestion Frege makes in his *Grundlagen der Arithmetik* as to how we might ascertain the objectivity of the natural numbers through the contextual definition of number vocabulary. It has three defining characteristics over and above the description of Platonism given in Section 1.1. Firstly, it is concerned, specifically, with the natural numbers and the concept natural number.[3] This latter is understood to be a sortal concept, i.e. a concept with which is associated a criterion of identity and distinctness for its instances. Secondly, there is the argument given to attain the Fregean's three aims. Thirdly, and finally, there is this arguments reliance on Frege's Context Principle, the famous dictum that we should:

"never ... ask for the meaning of a word in isolation, but only in the context of a proposition." (*Grundlagen*, Introduction)

To begin with the last, Wright finds two strands of interpretation in the Context Principle. An interpretation under which the "content" Frege alludes to is taken to be (something like) Fregean sense, and an interpretation under which it is taken to be reference.[4] The former I call the Context Principle for Sense. It is this thesis:

(CPS) An adequate account of the meaning of a sub-sentential expression must make clear the contribution it makes to the meaning of the sentences in which it occurs.

---

[3] So, for example, whether or not the Fregean approach can be applied to other systems of mathematical objects, such as the reals, is not decided upon, though Wright (1997a) contains some tentative suggestions concerning this issue. See also Parsons (1997) for investigations into the possibility of a Fregean approach to sets.

[4] It might be objected here that Wright is imposing ideas from Frege's middle period, i.e. the sense-reference distinction, onto an assertion from his early period and that this is not a justifiable procedure. But that would be to misunderstand the point of Fregeanism. Fregeanism aims to attain the three goals already mentioned. If it is successful, then whether or not Frege proposed it in its successful form is an interesting exegetical question. But it does not concern us here. Perverse as it may sound, we are interested in Wright's conception of Fregeanism, not Frege's.

Here, of course, a "sub-sentential expression" should be taken to be a semantically significant unit and not just any part of a sentence. For example, singular terms and predicates are sub-sentential expressions, syncategorematic expressions and incomplete symbols, in Russell's sense, are not (Russell, 1919). In an older vocabulary, we might say that the expressions (CPS) is concerned with are those which correspond to constituents of the propositions expressed by the sentences in which they occur. That such a constraint is necessary to an adequate account of meaning is plain. If I am presented with a sentence the sub-sentential constituents of which I understand with but one exception, then an explanation of the meaning of that expression must put me in a position to grasp the meaning of the sentence. This is the case *whatever* sentence I am presented with. It follows that the explanation of the unfamiliar expressions meaning must have put me in that position, it must have taught me the expressions contribution to the meaning of the sentences in which it occurs.[5]

The second strand of interpretation in Wright's account of the Context Principle, that which takes the principle to be concerned with content as reference, I call the Context Principle for Reference. It is this claim:

(CPR) Syntactic categories are explanatorily prior to ontological categories.

Thus distinctions between types of entity should be drawn with respect to distinctions between the syntactic role of different classes of expressions. Ontological categories are *defined* by the formal properties of the symbols used to represent them. So, for example, the class of objects is defined as the class of those entities apt to be referred to by a singular term, the class of first-level concepts is defined as the class of those entities apt to be referred to by a predicate whose arguments are singular terms etc.. It follows that, if the definitions of ontological categories are not to be circular, the semantic role of an expression, e.g. functioning as a singular term, must be characterisable on purely syntactic grounds. Hence, the Fregean is committed to providing syntactic criteria for membership of the different classes of expression. This

---

[5] It may be objected that, if I do not know the compositional structure of the sentence I am given, then an adequate account of the meaning of the unfamiliar expression need not put me in a position to understand it. That is certainly true. However, the compositional structure of the sentence - its logical form - can itself be understood to be a constituent of the proposition expressed by the sentence, even though there is no sign in the sentence to signify it. See section 2.3.

is a non-trivial task. However, in no part of this study will I assess the Fregean's claim to have achieved it. For the sake of argument, I assume that such criteria can be, and have been, formulated.[6] Neither will I evaluate Wright's interpretation of the Context Principle. This means that I am not going to assess whether Frege meant to assert either (CPS) or (CPR) (or both) when he asserted the context principle and that I am not going to evaluate them in their own right. Both (CPR) and (CPS) are assumed.[7]

Let us now make clear certain of Wright's other assumptions. Firstly, our understanding of a sentence, our grasp of its meaning, is at least partially explained by our knowledge of its truth-conditions.[8] So by (CPS), the meaning of a sub-sentential expression must make clear at least its contribution to the truth-conditions of the sentences in which it appears. It follows that the Fregean will have taken the first step towards attaining her first objective if she can give an epistemologically sound account of the truth-conditions of sentences containing arithmetical vocabulary. This is because such an account will dispel the Conservative Nominalist's worry about how we can know about the numbers and will show the Radical Nominalist how we can begin to acquire number-theoretic concepts. Secondly, there is a thesis that I call the "Reference Thesis". This is the following claim:

(RT) All sub-sentential expressions appearing in true sentences contribute to the truth-conditions of those sentences by referring.

It follows that if a sub-sentential expression appears in a true sentence, then its referent exists. So, by (CPR), if a singular term appears in a true sentence, its referent exists and is an object. Thus, the Fregean will have achieved her second objective, to show the objectual and mind-independent existence of the natural numbers, if she can exhibit some uncontentiously true sentences which contain singular terms which refer to numbers. She will also have shown the Radical Nominalist why reference to numbers is not problematic which will complete the attainment of her first objective. Therefore, the questions we must ask are these: Is there an epistemologically sound account of

[6] See Dummett (1973), Hale (1987).

[7] I will say this much, though: (CPR) seems to be correct as part of an interpretation of Frege's middle period philosophy of logic; (CPS) seems to be correct. With respect to the former, even Wright admits that (CPR) may seem counter-intuitive and does not receive any positive support at his hands (1983, p52). Instead, he tries to show that accepting (CPR) is more attractive that rejecting it.

[8] See Section 2.1 and 2.2 for an explanation of this somewhat cagey expression.

the truth-conditions of number-theoretic sentences and, if so, what is it? Are number-theoretic sentences ever true and, if so, how can we know them?

The Fregean's answer to both of these questions makes reference to a certain principle: Hume's Principle. It is as follows:

(HP) Two concepts have the same number if and only if they can be put into one-one correspondence, i.e. if and only if there is a bijective mapping from the things that fall under one to the things that fall under the other.

In fact, the principle contains two "sub-principles", that the number of any one concept equals the number of any other if and only if the concepts are *equinumerous*, i.e. there are *as many* of the one as of the other, and that two concepts are equinumerous if and only if they can be put into one-one correspondence. In logical notation, Hume's principle may be written:

(HP) $\forall F \forall G(Nx{:}Fx=Nx{:}Gx \leftrightarrow \exists R(F\ 1\text{-}1_R\ G))$,

where "Nx:Fx" is the cardinality operator which should be read "The number of F's" and "$\exists R(F\ 1\text{-}1_R\ G)$" abbreviates a second-order logical formula expressing "there exists a bijective mapping from F to G".

Now, the Fregean makes three claims with respect to (HP). Firstly, she claims that the terms formed when the cardinality operator binds predicates are singular terms.[9] To justify this claim she appeals to the aforementioned syntactic criteria for singular termhood. I will not challenge this claim at any point in this study. Secondly, she claims that (HP) can be stipulated to be true. So understood, it fixes truth-conditions for sentences containing the cardinality operator. The Fregean infers that it can be used as a *contextual definition* of that operator. As such it is supposed to introduce and explain the sortal concept natural number. This is *all* we need in order to understand the sentences of arithmetic.[10] So, if the Fregean is correct to assert that (HP), taken as a definition, can fix our concept natural number, then part of her first

[9] The astute will observe that this commits the Fregean to a referential treatment of at least some uses of some definite descriptions. An interesting question to ponder is what becomes of the Fregean's argument if all definite descriptions are Russellian. I shall not be dealing with this question.

aim will have been attained. She will have convinced the Radical Nominalist that there is no epistemological problem with our grasp of arithmetical concepts. Thirdly, the Fregean claims that the right hand side of instances of (HP) are often true and that we are often in a position to see that they are. For example, we are able to verify that "$\exists R$(x is Frege $1\text{-}1_R$ y is Frege)". Since (HP) is true, it follows that we are able to verify that "Nx:x is Frege = Ny:y is Frege" is true. So (HP) also yields true sentences containing singular terms that refer to numbers and we can tell that some of these are true. Thus, given (RT) and (CPR), numbers are independently existing objects. Hence, the Fregean achieves her second aim, demonstrating the independent, objectual existence of the numbers, and completes the attainment of her first by showing the Radical Nominalist that there is no problem with referring to numbers. So, given all her other assumptions, the Fregean claims that the stipulation of (HP) is sufficient to put us in a position to attain her first two objectives. I call the Fregean's argument when (HP) is stipulatively true "the Argument I". Sadly, it looks as if the Argument I fails. I will explain why in Chapter 2.

Assuming that I am correct that the Argument I fails, what is the Fregean to do? In Chapter 3, I suggest a weakened version of the Fregean's argument, call it "the Argument II". It is weaker than the first for it assumes that we already possess and deploy the concept natural number and that this concept is sortal. (HP) is therefore invoked only to provide true sentences in which singular terms for numbers occur (as above). However, since (HP) is not stipulated to be true, we must find some other explanation of its truth and of our knowledge of it. Otherwise, it will not yield true sentences which require the existence of numbers to explain their truth and, even if it did, we would not necessarily know them. My suggestion is that (HP) may be knowable *apriori*, that it may be known, in some sense, independently of experience.[11] Note that the Fregean who proposes the Argument II cannot hope to achieve the first aim of Fregeanism, to dispel *all* the epistemological doubts motivating nominalism. This is because the Argument II will not show the Radical Nominalist how we acquire a grasp of arithmetical concepts. Nonetheless, it might convince the Conservative

[10] See Section 2.1, paragraph 1.

[11] I do not mean to suggest that this suggestion is uniquely mine. Frege sometimes says that (HP) might represent the "correct analysis" of our pre-existing concept of number. Wright, too, envisages this weakened form of the argument under which (HP) is an apriori truth concerning concepts we already possess. However, for both, the proposal they really want to advance is the Argument I.

Nominalist by showing her how we can attain knowledge of numbers. This would still be a noteworthy result. Whether the Argument II can prove it is the subject of Chapter 3.[12][13]

[12] Note that there is a large class of sentences structurally similar to (HP) which could be used in arguments analogous to the Fregean's arguments concerning numbers. Some of these are first-order logical sentences of the form ⌈#a=#b ↔ E(a, b)⌉ where ' #' denotes a second-level function mapping distinct equivalence classes of objects onto distinct objects and 'E(a, b)' expresses an equivalence relation on objects. I call these A-Sentences. Others are second-order logical sentences of the form ⌈#F=# ↔ GE(F, G)⌉ where '#' denotes a third-level function mapping distinct equivalence classes of concepts onto distinct objects. I call these B-Sentences. When the distinction between the two is not important they are merely called "Sentences".

[13] terminological note: throughout the rest of this study the expressions "number" and "natural number", and their cognates, are used interchangeably.

Chapter 2

The Argument I

2.1 Introduction

Having made clear the background and fundamental assumptions of Fregeanism, it is time to consider in greater detail how its three aims are to be achieved when (HP) is supposed true by stipulation. Recall (from Section 1.1) that the aims are as follows: to dispel the epistemological doubts surrounding the existence of natural numbers; to show that Platonism is true by demonstrating the independent, objectual existence of the natural numbers and to show that some species of Logicism can be upheld, at least with respect to number theory. In Wright's most recent account of Fregeanism (1998b), he states that the position involves four claims:

"(i) that the vocabulary of higher-order logic plus the cardinality operator, 'Nx:...x...', provides a sufficient definitional basis for a statement of the basic laws of arithmetic [the Dedekind/Peano axioms];
(ii) that when they are so stated, $N^{=}$ [(HP)] provides for a derivation of those laws within higher-order logic;
(iii) that someone who understood a higher-order language to which the cardinality operator was added would learn, on being told that $N^{=}$ is analytic of that operator, all that it is necessary to know in order to construe any of the new statements that would then be formulable.
(iv) Finally and crucially, that $N^{=}$ may be laid down without *significant epistemological obligation*: that it may simply be stipulated as an explanation of the meaning of statements of numerical identity, and that - beyond the issue of the satisfaction of the truth-conditions it thereby lays down for such statements - no competent demand arises for an independent assurance that there *are* objects whose conditions of identity are as it stipulates." (1998b, p389)

Claims (i) and (iii) are concerned with what it is we grasp when we understand the propositions of arithmetic. Claims (ii) and (iv) are concerned with how we recognise their truth. I shall not, in this study, be concerned to dispute the first and second claims. In fact, I accept them. As Wright observes (p390), this is to recognise two quite significant achievements. First, what he calls the "analytic reduction" of arithmetic to a higher-order logic (in fact, second-order will do) whose lexicon contains only one non-logical primitive, the cardinality operator. Second, the proof of "Frege's Theorem", the statement that the Dedekind/Peano axioms for arithmetic can be derived from (HP) alone.[14] Wright is quite correct to label these "substantial achievements". However, with one proviso, I regard their interest to be mainly mathematical, not philosophical. The proviso is that (i) and (ii) demonstrate that the only concept required to understand arithmetic is the concept natural number (see n10).

The philosophical work, then, must involve claims (iii) and (iv). It is easy to see that it does. Given (i), the truth of (iii) entails that there can be no epistemological problem concerning our grasp of arithmetical sentences. Given (i) and (ii), the truth of (iv) entails that there can be no epistemological problem with our recognition of a range of objects (the natural numbers) which satisfy the Dedekind/Peano axioms. Thus, the epistemological doubts motivating Nominalism (identified in Section 1.1) are dispelled. That was the first aim of Fregeanism. Furthermore, the second aim, too, is achieved if (i)-(iv) are true. (i) and (iii) ensure that anyone who understands second-order logic is in a position to understand the sentences of arithmetic. (ii) and (iv) ensure that such a person will recognise that these axioms are true and that they are therefore committed to the existence of the natural numbers. To convert this commitment into a commitment to the numbers *as objects*, one need only note that, according to the syntactic criteria for singular termhood mentioned in Section 1.2, the cardinality operator binds predicates to form singular terms. Hence, by definition according to (CPR), their referents are objects. To see that they exist *independently* of any thought or talk about them invoke (RT) and note that the truth of arithmetical sentences is, on this account of their truth-conditions, as independent a matter as the truth of sentences concerning equinumerousity on concepts. Having assumed that (i) and (ii) are true, it remains to explain what justification is sufficient for (iii) and (iv). Mimicking Rosen (1993), we may construct a useful little story to identify the main points of contention.

Suppose we have on our hands a die-hard nominalist (in fact, a Radical Nominalist) who speaks a language, L, which includes the resources of first order logic interpreted classically and the resources of second-order logic interpreted in some nominalistically acceptable way.[15] Suppose, further, that she accepts a Quinean standard of ontological commitment, under which she is committed to any objects she quantifies over and refers to (Quine, 1958), and believes in a univocal, realist account of the truth of sentences of L.[16] In particular, assume that she takes the existential

[14] Boolos introduced the label "Frege's Theorem" for this statement since Frege outlined its proof in *Grundlagen* (Boolos, 1990). I adopt the name throughout this study.

[15] Whether there is such an interpretation of second-order logic is not a trivial matter. For one suggestion for such an account see Boolos (1994, 1995).

[16] The assumption that she believes in realistic truth for L is actually stronger than necessary. The point of the Argument II is to show truth-in-L* is the same concept as truth-in-L. So number-theoretic truth is the same as truth-in-L. Hence, there is no room for a nominalist construal of the

quantifier to be univocal and *always* indicative of ontological commitment.[17] Suppose, finally, that she denies any commitment to abstract objects. The Fregean suggests that she extends L to a new language L* by stipulating that for any predicates F and G:

(1) ⌜Nx:Fx=Nx:Gx⌝ means that ⌜$\exists$R(F 1-1_R G)⌝

and asserts that from (1) it is possible to infer:

(HP) $\forall F \forall G((Nx{:}Fx=Nx{:}Gx) \leftrightarrow (\exists R(F\ 1\text{-}1_R\ G))$.

The Fregean then exhibits her syntactic criteria for singular termhood, which will verify that the cardinality operator "Nx:...F..." binds predicates to form singular terms. Call these "N-terms". She also shows that claims (i) and (ii) above can be verified on the basis of (HP). I assume that the nominalist accepts these results.

In support of claim (iii) above, the Fregean asserts the following epistemic thesis:

(ET) A (Platonist) grasp of instances of the left hand side of the sentence "⌜Nx:Fx=Nx:Gx⌝ means that ⌜$\exists$R(F 1-1_R G)⌝" can be derived from a (Nominalist) grasp of the corresponding instances of the right hand side.

Doubtless it is plausible that whatever the nominalist understands by the LHS of an instance of (1) is derived from her grasp of the corresponding RHS, since the nominalist understands L prior to understanding L*. However, (ET) asserts much more than that. From a grasp of the RHS of an instances of (1), the nominalist is expected to derive the concept number, and so acquire the ability to (platonistically) construe all sentences of arithmetic, consonant with claim (iii). However, it is not obvious that this is possible because (1) does not allow for the elimination of the cardinality operator from all contexts in which it can appear in L*. For example, it

---

number-theoretic fragment of L*, given all the Fregean's assumptions and claims concerning the transition from L to L*. The Argument I is supposed to show that numbers exist in the same sense as everything else.

[17] The reasons for this precise characterisation of the nominalist's attitude to ontology will become clear in Section 2.3.

does not allow for its elimination from sentences like "Jenny the Radical Nominalist is Nx: x is a chair", i.e. mixed identity sentences where only one term flanking the sign for identity is, or is definitionally equivalent to, a term formed by the cardinality operator. Neither does (1) allow for the elimination of the cardinality operator when it is applied to predicates which already contain it, e.g. it fails to eliminate it from the sentence "Nx:x is Big Ben is Nx:x is [Nx: x is a chair]". This tells us two things about (1). Firstly, it is not, strictly speaking, a definition. A definition in the strict sense allows for the elimination of the vocabulary it defines from any sentence in which that vocabulary appears. Secondly, it is not immediately obvious how our nominalist can derive an understanding of all arithmetical sentences from L* alone. But a demonstration to that effect is required for the justification of (iii). Thus, the question to address is what criterion should be placed upon the cognitive capacities of our nominalist in order that (iii) be deemed true. Note that the nominalist must be able to derive from (1) the *sortal* concept number which the Fregean already claims to possess and deploy. Recall (from Section 1.2) that a sortal concept is one with which there is associated a criterion of identity and distinctness. This will distinguish instances of the concept from each other and from instances of other sortals. If it can be shown that the nominalist derives such a concept from (1) there can surely be no block to her understanding all the sentences of arithmetic, given that (i) is true. The suggestion, then, is that the nominalist must be able to evaluate all statements of the form ⌜q=Nx:Fx⌝, where q is any singular term in L*, and that this ability should flow from her knowledge of (1). This, of course, is Frege's Caesar Problem for definitions of number, that a putative definition of number must provide a basis for deciding all identity statements involving singular terms for numbers (Frege, 1950, p67-68). Not surprisingly, the Fregeanism proposed by Wright and Hale must answer the Caesar Problem (a fact of which they are well aware). Only if it can will (iii) be justified.

However, solving the Caesar Problem, although necessary for the Fregean, is rather an advanced point at which to start when arguing for (iii). Prior to that, one might feel that the LHS's of instances of (1) must be shown not to be merely unstructured labels for their corresponding RHS's but to be semantically complex symbols. In fact, the Fregean must assert more even than that. In order to give substantial content to (ET), she must advance two further claims, both concerning the semantics of the sentences of L*. These I call the Semantic Stipulation and the Interpretation Thesis:

(SS) The left hand sides of instances of "$\ulcorner$Nx:Fx=Nx:Gx$\urcorner$ means that $\ulcorner \exists R(F\ 1\text{-}1_R\ G)\urcorner$" are semantically complex.

(IT) The transition from L to L* is semantically non-revisionary, i.e. all symbols of L are interpreted in L* as they were in L.[18]

Observe that it would be a mistake to conflate these two theses. Although it seems reasonable to suppose that (IT) entails (SS), the converse does not hold. One could adhere to (SS) without accepting (IT), for example by saying that, for given predicates 'F' and 'G', "Nx:Fx=Nx:Gx" is of the logical form $\ulcorner$aRb$\urcorner$ but by denying that '=' expresses identity. To see how (SS) and (IT) fleshes out (ET), observe that the Fregean assertion of (ET) should understood to *assume* (SS) and (IT). (ET) is the claim that, given (SS) and (IT), one can derive from a grasp of instances of the RHS of (1) the sortal concept number. So it is the claim that one takes the meaning of the RHS's of instances of (1) and, by imposing a different, partially interpreted logical form upon them, acquires a grasp of the meaning of the corresponding LHS's.[19]

(SS) and (IT) have a further job to do, viz. to justify (iv). (iv) entails, via (ii), the claim that from a proof of the truth of the Dedekind/Peano axioms one can infer a commitment to the existence of the natural numbers. To justify this claim, the Fregean infers from (SS) and (IT) the following claim, which I call Ontological Priority:

(OP) The ontological commitments in both the left and right hand sides of instances of "$\ulcorner$Nx:Fx=Nx:Gx$\urcorner$ means that $\ulcorner \exists R(F\ 1\text{-}1_R\ G)\urcorner$" are those that are explicit on instances of the left hand side.

This inference is justified by the assumptions made in Section 1.2. By the syntactic criteria for singular termhood, the N-terms are singular terms. According to (CPR) and (RT), such terms have objectual reference when they appear in true sentences. So,

[18] That (SS) is something that can be stipulated is explained in Section 2.3 where it is also argued that (IT) cannot be a matter of stipulation.

if the RHS of an instance of (1) is satisfied, then the corresponding LHS commits us to the existence of the referents of its N-terms. The only way to block this inference is to deny either (SS) or (IT). So, given (1), (i) and (ii), the truth of (OP) is enough to show that we are committed to a countable infinity of objects. The only way for the Nominalist to deny that this was a commitment to the natural numbers would be for her either to assert that the objects required to model the Dedekind/Peano axioms are concrete objects or for her to abandon her nominalism, saying that those objects are abstract, but deny that they are the numbers. I can see no reason whatsoever for her to adopt the latter course. With respect to the former, I think such a response would be at best dubitable. This is because it commits the nominalist to claiming, apriori, that the concrete world contains infinitely many objects. But how could that be known apriori?[20]

It is pertinent to observe that (OP) does not just say that speakers of L* are committed to the existence of numbers but also that *so are the speakers of L*. The Argument I is greatly attractive because it seeks to show, not just that Platonism is true, but that Nominalism is incoherent. Given the truth of (OP), the nominalist speaker of L is already committed to the existence of numbers, it's just that until the move to L* is effected this commitment can go unobserved.

This is the Fregean's side of our story, then. But the nominalist is not going to accept it all without a whimper. There are, in fact, a great many objections to the Argument I. Field compares the whole process to the Ontological Argument for the existence of God and states that Fregeanism is just as suspect.[21] Heck and Boolos have argued that it is not clear that (1) can be known to be an acceptable "definition" (Boolos 1990, 1997, Heck, 1990). Dummett has asserted that it only establishes a "parochial platonism" for the language in question and argues that this shows that even if the Argument I is accepted, then the existence of numbers must be mind-dependent (1973). He has also attacked (SS), (IT) and the use of (HP) as a "definition".

---

[19] (SS) and (IT) are what is required to fix the meaning of instances of the LHS beyond the fixing of their truth-conditions (see Section 1.2, n8).

[20] There is one notable nominalist who is prepared to suggest that the world contains infinitely many concrete objects: Field suggests that there exists an infinity of (concrete) space-time points (Field, 1980). However, this is not known apriori as it must be in this context for the nominalist speaker of L.

[21] This can be found in Field (1984). Field's paper precipitated an exchange which is not yet over. See, in chronological order, Wright (1988, 1990), Hale (1990, 1992), Field (1993), Hale (1994).

(Dummett, 1973, 1991, 1998)[22]. This to mention just a few. There is not space in a study of this length to address all the objections to be found in the literature.[23] What I propose to do is to focus on those that would occur first to the nominalist whom the Fregean is trying to convince. The first objection I shall deal with (Section 2.2) is a demand for clarification and justification: Under the assumption of (SS) and (IT), is (ET) true? Clearly, if it is not, the Argument I fails to get off the ground. Next (Section 2.3), I shall consider a very natural response to the move from L to L*. This response is embodied in the assertion that the correct approach to the extension of L is a *reductive* approach, an approach that denies (OP). This sort of approach is motivated by the very plausible intuition that the linguistic posturing that takes us from L to L* can really be no more than that: linguistic posturing. Then (Section 2.4), I consider the Fregean's response to the Caesar Problem. By the Fregean's own admission, a solution to this problem is crucial to the success of the Argument I. Finally (Section 2.5), the results of the previous investigations are summarised and a conclusion reached: the Argument I, as proposed by Wright and Hale, fails.

[22] Wright and Hale respond to Dummett's criticisms in Wright (1983, 1997, 1998a, 1998b), Hale (1984, 1994, 1997).
[23] It is a methodological assumption of mine that all the objections to the Argument I that I do not consider can be rebutted. In particular, any challenges to the acceptability of (1) as a stipulation are assumed to have been met (for such challenges see the papers by Heck and Boolos cited in the text).

2.2 Equivalence and Explanation

To determine whether (ET) is true it is first necessary to determine the sense of "means that" as it occurs in (1). The sense of "semantic equivalence" in which (1) expresses an equivalence must be fixed. This is because, until we have a more developed understanding of (1), it is not clear how one might begin to derive a grasp of instances of the LHS of (1) from a grasp of instances of its RHS, let alone whether this is actually possible. In addition, it is not a foregone conclusion that all synonymies yield material equivalences. So it is not a foregone conclusion that from (1) it is possible to infer (HP).[24] But the Fregean needs (HP) to establish her claims (i) and (ii). This section therefore has two objectives: to determine the import of (1) and to evaluate (ET) given that (SS) and (IT) are supposed to hold.

We do well to return to the source of the argument. In a much publicised passage, Frege states that:

> "The judgement 'line a is parallel to line b', or, using symbols,
>
> a//b
>
> can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'the direction of line a is identical with the direction of line b'. Thus we replace the symbol // by the more generic symbol =, through removing what is specific in the content of the former and dividing it between *a* and *b*. We carve up the content in a way different from the original way, and this yields us a new concept." (*Grundlagen*, §64)

Frege took this case of directions to be analogous to the case of (1). Of course, the analogous case is rather too metaphorical to be of use as an elaboration of (ET). However, it does give us at least a suggestion for the sort of equivalence asserted by (1). This is identity of content. What we must decide in order to achieve our first objective is how to conceive of this "content".

A somewhat anachronistic suggestion would be to take content as Fregean sense.[25] However, there is a difficulty with this suggestion first made clear by Dummett (1991). Sense is a technical notion which brings along with it much theoretical baggage, most notably that the senses of expressions are composed of the senses of their constituents and that to grasp the former one must have a prior grasp of

[24] This may sound rather surprising. For an argument to the effect that it is true see Tappenden (1993).
[25] See n4.

the latter (and the compositional structure of the former, more on this below). Suppose now that (1) is taken to express identity of sense. (ET) states that our grasp of instances of the LHS of (1) is to be explained on the basis of our prior grasp of the corresponding instances of the RHS. Recall, also, that the LHS's are to be considered semantically complex and that the vocabulary of L that they involve is not reinterpreted in the move to L*, i.e. (SS) and (IT) are assumed. Dummett's point is that these four claims are inconsistent with the compositionality of sense.[26] Consider the sentence, "Nx:x is a cat = Ny:y is a dog". By compositionality, its sense is partly composed of the sense of "Nx:x is a cat". Furthermore, my grasp of "Nx:x is a cat" involves grasp of the sense of "Nx:...x...", again by compositionality. Hence, under the assumption that (1) expresses identity of sense, it follows that my (compositional) grasp of the sentence "∃R(x is a cat $1\text{-}1_R$ y is a dog)" involves grasp of that concept. But then it cannot be the case that (1) introduces a new concept via a prior grasp of its RHS, for, to grasp the RHS, I must already grasp the only concept that it is supposed to introduce, that expressed by the cardinality operator. Hence, Fregeanism is inconsistent with the compositionality of sense.

Obviously, the Fregean cannot deny (ET), (SS) or (IT). Each denial results in abandoning the Argument I. Hence, she is faced with the following choice: deny the compositionality of sense or deny the claim that (1) expresses identity of sense. However, as Hale observes (1997), working within a broadly Fregean framework there is only one alternative left for our explication of "content" if we opt for the latter: to claim that (1) expresses identity of reference. He also points out that if the referent of a sentence is taken (as Frege urged it should be) to be a truth-value, it is very difficult to see how to make sense of (ET). That is, if (1) expresses the material equivalence of its left and right hand sides we are left with the confusing claim that the stipulation of (1) allows us to "carve up" anew the truth-value of the RHS. But how does one "carve up" a truth-value? Matters are made even worse if we note that *any* necessary truths will be equivalent in this way. In the light of that fact, what sense can be made of the intimate link between the left and right hand sides of (1), the link that allows (1) to define or explain a new concept, when "2+2=4" and "Christmas is the celebration of Christ's birth." are equivalent in the same way?

[26] In fact, Dummett does not distinguish between (IT) and (SS), hence he takes it that there are three

In reaction to these thoughts, Hale suggests that we might take the fact or state of affairs (actual or possible) expressed by a sentence as the content identity of which is stated by (1). However, this suggestion can be rejected out of hand. The conception of content-as-state-of-affairs seems overtly metaphysical, a conception of content under which the content of "It is raining" is the fact that it is raining, conceived of as an independently existing entity (or, if not an entity, then at least an independently existing part of reality). Hence, to state that we are capable of grasping this sort of content is in direct and flat contradiction to Fregeanism's fundamental assumptions. As Wright puts it:

"We should thus have to credit ourselves with the capacity to recognise, quite independently of the syntactic form in which we give expression to a particular fact, just what objects there are in the portion of reality which constitutes that fact. Such recognition would be possible only if we had some understanding of the notion of an object which could be exercised independently of our understanding of syntactic categories and our capacities to recognise truth-value." (1983, p35-6)

But (CPR) asserts that our conception of ontological categories is dependent upon our prior grasp of syntactic categories and (RT) asserts that our ability to perceive the independent existence of objects is just our ability to determine that sentences containing singular terms are realistically true. Wright's gloss on the import of these two theses is that we cannot peer behind language to grasp states of affairs without mediation. The only things that we can so grasp are senses. Picturesquely, one might say that language builds a wall between us and the world. We can see without any problem (grasp without mediation) which bricks in which structural arrangement constitute part of that wall (what the sense of a sentence is, on the basis of the senses of its parts and its compositional structure). But the world and its states of affairs are behind the wall, we can only perceive it and them by looking at the wall. So content-as-state-of-affairs will *not* do for the equivalence expressed by (1) because we have to grasp that content *without mediation* to see that it can be carved up anew (the reason why we have to grasp it so will become clear below).

How, then, is the Fregean to cope with Dummett's objection to the suggestion that (1) expresses equivalence of sense? Hale observes that Dummett's objection derives its force, not just from the claim that sense is compositional, but from the claim that it is *strongly* compositional. As he puts it:

Fregean claims that are jointly inconsistent with the compositionality of sense.

"What the argument [Dummett's objection] relies on is the further claim ... that the sense of a complex expression *is actually composed of* the sense of its constituents in such a way that one cannot grasp the sense of the whole without grasping those of its parts." (1997, p249)

Now, Hale is not going to deny that Fregean sense is strongly compositional. Rather, he suggests that if a notion of sense can be made out which is not strongly compositional then it will be possible to block Dummett's argument. Call strongly compositional sense "strong sense" and call Hale's notion (whatever it is) "weak sense". Then, to grasp the weak sense of "Nx:x is a cat = Ny:y is a dog", I do not need to grasp the concept expressed by "Nx:...x...", so I do not need to grasp that concept in order to grasp the (strong or weak) sense of "∃R(x is a cat $1\text{-}1_R$ y is a dog)". Hale concludes that if a suitable notion of "weak sense" can be circumscribed, then it will be plausible to say that the equivalences express identity of weak sense, that this weak sense is what is carved up anew by our acceptance of (1) and that the strong sense of instances of its LHS can be derived from the carving up that results from the imposition on the appropriate weak sense of a certain partially interpreted logical form (that required by (SS) and (IT), i.e. the form ⌜aRb⌝ where the relation is construed as identity). So what characterisation does he suggest for the notion of "weak sense"?

The characterisation is given by supplying a criterion of identity for weak senses.[27] The criterion, which I call (CWS), is as follows:

"Two sentences have the same truth-condition [weak sense] iff anyone who understands both of them can tell, without determining their truth-values individually, and by reasoning involving only compact entailments, that they have the same truth-value." (1997, p257)

(CWS) is based upon Frege's criterion of identity for thoughts, which states, more or less, that two sentences express the same thought if someone who understands both of them can immediately tell that they have the same truth-value. The complications that Hale has added to Frege's criterion are as follows: that a thinker who understands two sentences need not be able to tell immediately that they have the same truth-value, i.e. that some reasoning upon the two sentences is to be permissible; that she must not

[27] Hale in fact uses the term "truth-condition" rather than weak sense but this is because his notion of weak sense corresponds to what others have understood by "truth-condition". (Hale, 1997, p252). In the text I retain the term "weak sense" in order not to blur the distinction between the sort of content we might be able to grasp (some sort of sense) and the sort of thing we cannot grasp (truth-conditions, i.e. states of affairs) given our assumption of (CPR) and (RT).

determine their truth-values individually and that she must only reason using compact entailments. The first condition is plausible because to disallow *any* inference at all is simply not plausible. The sentences "Today is Tuesday"" and "It will be Tuesday in 441 days time" share truth-conditions (and so, weak sense), but it takes some working out to see that that is the case. The second condition is to block all apriori truths from having the same weak sense. A thinker could tell that "Christmas is the celebration of Christ's birth" and "2+2=4" were both true (as was noted above) yet these sentences are hardly in the market for sharing weak sense and so being, in the context of the Argument I, "recarvings" of each other. The third condition is the most subtle. Hale defines compact entailment thus:[28]

> "A_1 ,...., A_n compactly entail B iff (i) A_1 ,....., A_n entail B and
> (ii) for any non-logical constituent E occurring in A_1 ,...., A_n there is some substitution E'/E which applied uniformly through A_1 ,....., A_n , yields A_1' ,......, A_n' which do not entail B." (p256)

The condition that only compact entailments be permissible in an investigation into whether two sentences have the same weak sense builds upon the second. Until it is imposed, there remain apriori sentences whose truth-values need not be determined individually for a thinker to judge them materially equivalent (Hale cites "$P\vee\neg P$" and "$Q\vee\neg Q$" as examples, p256) and aposteriori sentences for which the same holds ("P" and "$(P\wedge Q)\vee(P\wedge\neg Q)$"). Assuming we wish to deny the identity of these sentences weak senses, we must find some way of excluding them from satisfaction of putative criterion of identity. This is the job of compact entailment. What the condition actually does is to restrict the inferences allowable in determining whether sentences have the same truth-condition by disallowing inferences from premises which have redundant non-logical parts. The meaning of each non-logical constituent of a set of sentences must be crucial to any deductive consequence of that set if the deduction is to be compact.[29] It is easily seen that restricting the reasoning permissible to compact entailments blocks satisfaction of (CWS) by the pairs of sentences recently mentioned.

[28] In fact, he took the definition straight from Wright (1989).

[29] In fact, there are further niceties to be discerned with compact entailment and its presence in (CWS). However, these are not of interest here, as I shall not be objecting to the criterion on the grounds that it employs the notion of compact entailment.

For example, P can be inferred from $(P\wedge Q)\vee(P\wedge\neg Q)$ no matter how "Q" is interpreted. Hence, the inference is not compact.

This is Hale's proposed (CWS). It is a criterion which individuates weak senses. An instance of (1) asserts the identity of the weak sense of its left and right hand sides. Thus we have achieved our first objective, i.e. to determine the import of "means that" as it appears in (1). Note, also, that this explanation of (1) justifies the inference from (1) to (HP) (see n23).

However, before turning to our second task, to determine whether (ET) is true, it is worth asking if there are any other objections which may give us just cause to doubt the utility of weak sense in the Argument I or, indeed, anywhere else. One problem that springs to mind is that, since the (CWS) involves *thinkers*, it leaves open the possibility that weak senses are thinker-relative. To elaborate, suppose two people, A and B, both understand two sentences, S_1 and S_2. Assume, further, that A can see that S_2 is compactly entailed by S_1, whereas B cannot. It then follows that S_1 and S_2 have the same weak sense for A but different weak senses for B. So the identity of weak senses is subjective (i.e. thinker-relative). This does not present an immediate problem for the Fregean, since the weak senses they are interested in are those of instances of the LHS of (1) and these are the weak senses of the RHS's by stipulation. Hence no-one could fail to see that they were the same. However, it will transpire (see below) that ontological commitments go hand-in-hand with weak senses. So if weak senses are subjective, so are ontological commitments.[30] This is a result which I assume the Fregean can not accept.

So how might the Fregean respond? Firstly, she may suggest that A and B do not both understand S_1 and S_2. But this is not plausible. Understanding is a vague notion. It is not possible to give a strict criterion for when a thinker understands a sentence.[31] As a result, there are certainly cases in which we would say that two people both understood a sentence or set of sentences, even though they were not able to draw the same inferences from them. For example, many people would understand the axioms of Group Theory. Not all of these would be able to infer Sylow's

[30] And by this I mean not that what people *think* they are committed to will differ but that what they are *actually* committed to will differ, depending upon what compact entailments between sentences they are capable of following.
[31] See, for example of such a view of understanding, Wittgenstein (1968).

Theorems from that basis. This goes for compact entailments, too, since the property of compactness is transitive over deductive inference.

A second, more plausible response, may be to "idealise" (CWS), i.e. to individuate weak senses with a similar criterion beginning, "Two sentences have the same weak sense iff an *ideal thinker* who understands both...". Of course, the notion of an ideal thinker itself requires explanation. However, I suspect that such an explanation is possible. Something along these lines could be made to work. Since, for the Fregean, it *has* to work, I'm going to assume that it does.[32]

Having assumed that the notion of weak sense is acceptable on its own merits, we can now determine whether (ET) is true. To begin with, note that Hale's proposal of taking (1) to express identity of weak sense under the characterisation of weak sense given above, does provide a picture with which to understand Frege's metaphor in the *Grundlagen*. (1) expresses identity of weak sense. In grasping an instance of the RHS of (1), "∃R(x is a dog $1\text{-}1_R$ y is a cat)" say, we grasp its strong (compositional) sense. From this we derive a grasp of its weak sense. We then impose upon this weak sense a compositional structure, in accordance with (SS). This is the logical form ⌜aRb⌝. We construe the relation in that form to be identity, according to (IT), and construe 'a' and 'b' as singular terms. In so doing, we arrive at the strong sense of the LHS, i.e. we arrive at the sense ordinarily understood by "N:x is a dog = Ny:y is a cat". This explains how (ET) might be true since it shows us how (1) allows us to attain a grasp of new concepts on the basis of old ones.[33] Furthermore, it has the added advantage of rationalising (OP). Recall that this thesis states that both speakers of L and speakers of L* are committed to the ontology required by a face-value reading of the LHS's of instances of (1). We are now in a position to see why. One is committed to whatever objects are required for a face-

---

[32] A word of explanation as to why we could not take all mention of thinkers out of (CWS): Hale is concerned to ensure that weak senses are not susceptible to Slingshot arguments. These are arguments which demonstrate, of a given class of entities which might be taken to be the content expressed by a sentence, that there are far fewer of them than one might think; normally two, one corresponding to true sentences, one corresponding to false ones. It is necessary to avoid Slingshots for weak senses because, if there are not enough of them, (ET) is no longer plausible. Given one weak sense for all true sentences and assuming (ET), for example, the weak sense of "Marmalade is made using oranges" can be recarved to yield the strong sense of "The integral of one divided by tan(x), over the real number interval $[0,\infty)$, is π divided by two." This makes a mockery of (ET). However, Slingshots tend to affect only extensional characterisations of sentence contents. Hence, the reason why thinkers are included in (CWS). For more on Slingshots see Neale (1995, 1997).

value reading of the semantics of a sentence whose strong sense is the result of any acceptable carving up of a weak sense. Implicit ontological commitments correlate with whatever weak senses we believe. Explicit ontological commitments correlate with how we express these weak senses (i.e. they correlate with strong senses). However, the commitments of the former do not coincide with the latter. It sometimes requires a "reconceptualisation"[34] of a weak sense to make the implicit commitments explicit. One might object that this is little better as an elucidation of (ET) than Frege's metaphor in the *Grundlagen*. However, I do not think that would be fair. We now have an account of the sense in which (1) expresses an equivalence together with a criterion of identity to individuate the things of which it asserts identity. We have an explanation for a nominalist speaker of L as to why they should not find (OP) so surprising. Additionally, we have the beginnings of an explanation of how we move from a grasp of the (strong) sense of the RHS of an instance of (1) to a grasp of the strong sense of the corresponding LHS. This is a lot better than Frege's metaphor.

However, it is not clear that this notion of weak sense will not collapse into a metaphysical, rather than a semantical, notion. By this I mean that it is not clear that (CWS) will individuate some sort of content rather than some sort of fact, where the latter are understood to be mind-independent constituents of reality. Hale argues convincingly (1996) that truth-conditions are distinct from strong (compositional) senses. Let S_1 and S_2 be any two sentences which share truth-condition but now their compositional form. Even though it is a thesis of Frege's that fixing the truth-conditions of S_1 and S_2 is sufficient to fix their (strong) senses, it is a mistake to identify their strong senses with truth-conditions. This is because strong sense also determines reference and does so in a particular way. S_1 and S_2 determine their reference in the same way if they involve the same constituent senses united in the same compositional structure. But, by assumption, S_1 and S_2 do not share compositional form. Hence they do not share strong sense, so identity of truth-condition is not sufficient for identity of strong sense.

Of course, this argument fails if sentences cannot be found which fulfill the conditions placed upon S_1 and S_2. Hale gives an example of some such which nicely

[33] The mental process I have described seems to be the only cogent way to read Hale's account of (ET) in Hale (1994).

[34] One of Wright's favourite modes of expression when explaining the Argument I.

illuminate the argument in the general case (1997, p253-4). It is plausible to suppose that "There are husbands" and "There are wives" have the same truth-condition, viz. "$\exists x \exists y$(x is female $\wedge$ y is male $\wedge$ x is married to y)". However, it is not plausible to say they have the same sense. "There are husbands" is true because something satisfies the predicate "$\exists x$(x is male x $\wedge$ x is married to ξ)", "There are wives" is true because something satisfies the predicate "$\exists y$(y is female $\wedge$ y is married to ξ)". That is they determine their truth-value in different ways. Since they are composed of the same constituent senses, viz. "ξ is male", "ζ is female" and "ξ is married to ζ", the only way to explain this difference is to admit that the compositional structure of a sense is just as essential to its identity as the concepts of which it is composed.[35]

The problem for the Fregean is that there is no obvious analogue of this argument for weak senses. This is because compositional structure is *not* essential to the identity of weak senses. Furthermore, weak sense is not compositional *by design*, for its definition was suggested to escape Dummett's objection to taking the "content" alluded to by Frege as Fregean (or something like Fregean) sense. I do not, at the moment, see a way in which the Fregean might respond to this challenge. If weak sense does indeed collapse into a metaphysical notion, then it is subject to the decisive objection described above for taking content as "state-of-affairs-expressed" or something similar. So if weak sense does collapse into a metaphysical notion, the Argument I is inconsistent. This is because, according to the explanation of (ET) recently proposed, we grasp the weak sense of a sentence in order to derive a grasp of the strong sense of a different sentence. If weak senses are (metaphysical) states of affairs, then suggesting that we can grasp weak senses contradicts (CPR) and (RT). On the assumption that (CPR) and (RT) are true, assumptions that the Fregean must make to motivate the Argument I, it follows that (ET) is false.

Of course, the challenge may be met. It is therefore pertinent to ask whether the situation is any better if weak sense does not collapse into a metaphysical notion. Suppose our view of the world is staggered, as it is according to the Fregean. Our grasp of different states of affairs is via our grasp of strong Fregean senses. But in between a state of affairs and a strong sense which represents it there is another sense, a weak sense, distinct from both the state of affairs and the strong sense. Different

[35] This explains the comment n6, where I said that I thought it plausible to take the compositional

carvings of this weak sense yield different strong senses which correlate with the same state of affairs. In grasping a weak sense we are supposed to see that it can be carved up in these different ways. But that surely requires that we can grasp it without the help of any of the strong senses which it yields. We have to have an unmediated (i.e. pre-linguistic, even pre-representational) view of a weak sense in order to see that it yields a certain strong sense. The problem is that one cannot consistently assert (CPR), (RT) and the claim that we have a pre-representational grasp of weak senses. If Wright is correct to urge that (CPR) conjoined with (RT) entails that our understanding is always mediated by the logical properties of our representational system (see above), then I do not see how we can grasp weak senses at all, or, at least, not in the way the Fregean requires. For, to repeat, the only way that we have found to explain Frege's metaphor of carving up contents is this: we grasp a strong sense, *from this we derive an unmediated grasp of a weak sense*, from this we derive a different strong sense by carving up the weak sense in a new way. The italicised clause expresses something which, it seems to me, Wright's attitude to (CPR) and (RT) does not legitimate. A grasp of a weak sense is a grasp of an unstructured content, by definition of weak sense. According to (CPR) and (RT), the only things we can grasp without mediation are strong senses (as explained above). I can see no reason why this should not preclude an unmediated grasp of an (unstructured) weak sense in precisely the way in which it precludes such a grasp of a state of affairs. It follows, if I am right, that (ET) is inconsistent with (CPR) and (RT) as construed by Wright and Hale. So, under the assumption that they are true, it follows that (ET) is false. Hence, even if weak senses can be distinguished from states of affairs, they cannot do the job for which they were designed.

How might the Fregean respond to this objection? A first, somewhat naive, retort would be the claim that, actually, Frege's metaphor was precisely that, a metaphor, and it's not surprising that it should break down at some point when we try to refine it. This is a very poor answer. What we are doing is trying to provide a philosophico-theoretical representation of a practical ability, viz. the ability to derive new concepts from our grasp of old ones.[36] We need to provide a consistent such

structure of a sense as a constituent of that sense.

[36] This notion of a theoretical representation of an ability is not something I've just made up. See, Dummett (1976).

representation in order to be sure that we have such an ability. It seems to me that the Fregean has only one option: to deny that the truth of (ET) requires that we are able to grasp non-compositional weak senses in a non-compositional way. That is, to give a different account of (ET) which does not require that we are able to unmediatedly grasp weak senses.

There is a hint of this strategy in Hale (1997). He states, with respect to Frege's *Grundlagen* metaphor, that:

"... what Frege there envisages being carved up anew is not (directly anyway) the content of a complete sentence, but that of the symbol for parallelism." (1997, p245)

However, the suggestion is not developed into an account of (ET). Presumably it would go something like this (in very vague terms, I admit): we do not carve up the weak sense of "a is parallel to b" but rather the sense of "x is parallel to y" (which we can grasp, by compositionality, because we can grasp the strong sense of "a is parallel to b"); we find that it divides into the sense of identity and "another bit"; this other bit can be taken as the sense of the direction-operator and so we have all the conceptual materials required for construing some sentences as statements of directional identity. Furthermore, there was no need to grasp a weak sense during this process. So the objection I described above does not apply.

This account of (ET), even prescinding from its vagueness, is of no more use to the Fregean than the first. To begin with, note that the stipulation of weak sense equivalence between statements of directional identity and statements of parallelism amongst lines seems to have dropped out as irrelevant. Similarly with the stipulation of (1). On this account of (ET), there is no need to stipulate (1) in order to acquire, by a little conceptual surgery, the concept of number. We just tamper with the sense of "x is equinumerous to y" and find that it factorises into the sense of '=' and another bit. We then pick two predicates and put their senses together with the sense of '=' and the other bit, in a different form to that in which they are combined in a statement of equinumerousity, to yield a new strong sense. The Fregean will no doubt respond that the stipulation is necessary to fix the truth-conditions of statements of numerical identity. That, no doubt, is true, provided the sentences introduced on the LHS of instances of (1) can be interpreted as expressing the new strong senses we have learnt to form and to grasp. However, it is by no means clear that, on this account of (ET), such a procedure can be justified.

Consider the sentence:

(a) $\exists x$(x is a bird) $\wedge$ $\exists y$(y is a horse)

Unarguably, this contains the constituent concepts "ξ is a bird" and "ζ is a horse". We now start carving up concepts. "ξ is a bird" is seen to be composed of the concepts "ξ is winged" and "ζ is a creature". Now, we put these concepts together in a compositional structure different from that of (a) to yield the strong sense of:

(b) $\exists x$(x is winged $\wedge$ x is a creature $\wedge$ x is a horse).

Now suppose we stipulate the following:

(c) "Pegasus exists" shares weak sense with "$\exists x$(x is a bird) $\wedge$ $\exists y$(y is a horse)".

From (c), by the definition of weak sense, it follows that:

(d) "Pegasus exists" is true $\leftrightarrow$ $\exists x$(x is a bird) $\wedge$ $\exists y$(y is a horse).

The question I want the Fregean to answer is this: Is the strong sense of "Pegasus exists" that expressed by (b)? If it is, then we have shown that a winged horse exists if a bird exists and a horse exists. This is absurd. So the Fregean must assert that the strong sense of "Pegasus exists" is not the strong sense of (b). But now there is no reason to suppose the contrary of an instance of (1). The relationship between the stipulation (c) and the strong sense expressed by (b) is precisely the same as the relationship between the stipulation of (1) and a strong sense formed in the way suggested above (in accordance with the proposed account of (ET)). There is none. To elucidate, the suggestion is that we grasp the strong sense of "Cow is equinumerous to Horse", say, we divide up the sense of "equinumerous" to yield the sense of '=' and the sense of "Nx:Fx" and we unite the various components to form a new strong sense which we choose to express by "Nx:x is a cow = N:y:y is a horse". In addition, we stipulate that "Nx:x is a cow = N:y:y is a horse" shares weak sense with "Cow is equinumerous to Horse". But how can we stipulate that? The Fregean has shown us how to grasp a new strong sense, and she has stipulated, in accordance with (IT) and (SS), that a certain partially interpreted formula with a determinate logical form, i.e. $\ulcorner a=b \urcorner$, is to be equivalent in weak sense to a given assertoric sentence. She has not shown that that formula should be interpreted so as to express the new strong sense which she has chosen to express by "Nx:x is a cow = N:y:y is a horse". In effect, because this account of (ET) is independent of the stipulation of (1), the link between fixing the truth-conditions of sentences containing the cardinality operator and fixing their meaning has been broken. So, whilst it may be true to say that we can derive a fully compositional understanding of numerical identity statements by transforming our grasp of old concepts into a grasp of new ones, it is not true to say that Fregeanism is thereby vindicated. A compositional grasp of "Nx:x is a cow = Ny:y is a horse" will tell us that the truth-condition of that sentence is the existence, and identity, of the number of cows and the number of horses. But the identification of that truth-condition with the truth-condition for "Cow is equinumerous to Horse" can no longer be a matter of stipulation. According to the Fregean (1) is a stipulation. Therefore either (1) or (ET) must be rejected given this account of (ET).

Now, if (1) is rejected the Fregean no longer has any explanation of the truth-conditions of any arithmetical sentences. An explanation of a sentences truth-conditions should be taken to be a non-trivial specification of those truth-conditions, i.e. the identification of those truth-conditions with those of a different sentence. Part of the aim of the Argument I is to give such an explanation which links the truth-conditions for number-theoretic sentences to those of sentences which the Radical and Conservative Nominalists find epistemologically unproblematic. Furthermore, the rejection of (1) requires the rejection of (HP), since the Fregean infers (HP) from (1), and that is crucial to the proof Frege's Theorem which is, in turn, crucial to the explanation of number-theoretic truth. So the Fregean cannot reject (1). On the other hand, if (ET) is rejected then the Fregean has no explanation of how we grasp arithmetical concepts. It seems, therefore, that this account of (ET) lands Fregeanism in a dilemma neither horn of which it can grasp.

Perhaps it may be suggested that this dilemma can be finessed by observing that, once we have grasped the new strong senses derived in accordance with this account of (ET), i.e. the strong senses of arithmetical statements as we ordinarily understand them, we will be in a position to see that they share their truth-conditions with some of the old strong senses, i.e. the strong senses of statements of equinumerousity on concepts. However, to give this response the Fregean must explain how we are able to tell that different strong senses share the same truth-condition. As far as I can tell, the only possible explanation for this ability (in the present context) is based upon the assertion that we can unmediatedly grasp weak senses. If that assertion is true, we can tell that two strong senses are different carvings up of the same weak sense. This entails that they have the same truth-conditions. However, this brings us back to our ability to grasp weak senses which, I have argued, is inconsistent with the Fregean construal of (CPR) and (RT).

Let me clarify exactly what my position is here. We have found that (1) cannot express identity of strong sense on pain of inconsistency. This was the effect of Dummett's argument. It was therefore necessary to circumscribe a new notion of sense which did not fall prey to the same *reductio*. This notion was Hale's weak sense. We concluded that instances of (1) express identity of weak sense of their left and right hand sides. However, we then found that the notion of weak sense is not obviously different from the metaphysical notion of state of affairs and that, since the explanation of (ET) requires that we grasp weak senses, it follows that Fregeanism is inconsistent. Assuming that (CPR) and (RT) are true, it follows that (ET) is false. Furthermore, I have argued that even if weak senses do not collapse into states of affairs, we still have good reason to regard them as "ungraspable" in the light of (CPR) and (RT). Hence, again, (ET) is false. In addition, the only other account of (ET) that seems possible separates the job of determining the sense of arithmetical sentences from *explaining* their truth-conditions (as opposed to just fixing them). Consequently, the Fregean was faced with a choice between rejecting (1) or rejecting (ET). Both choices result in abandoning the Argument I. The only way to finesse this dilemma seems to be to appeal to our ability to grasp weak senses. As before, this entails the falsity of (ET).

I am not denying that (1) can express identity of weak sense. On the contrary, I think Hale's notion is a coherent one and that it could be true to assert that statements of equinumerousity on concepts and statements of numerical identity share weak sense in his, or something close to his, notion (whether this be a metaphysical notion or not). What I am denying is that we are capable of the sort of conceptual surgery required to validate (ET), on either of the accounts I have considered, given the assumptions of the Argument I. In light of the Fregean's own assumptions, (ET) is false. What this shows is that, given her own assumptions, the Fregean cannot explain how we acquire a grasp of arithmetical concepts in the way she hopes to. To do so was part of her first aim, to show that our grasp of arithmetical sentences is not epistemologically problematic in the way the Radical Nominalist suggests. Having so concluded, it is time to move on to the next point of contention identified in the Argument I: whether or not the move to L* should be reductively construed. To answer that question, I adopt the methodological assumption that (ET) is true.

2.3 Reductionism and Reference

Let us call any response to the Argument I which denies (OP) a reductionist response. The reductionist asserts that L* is reducible, without ontological residue, to L. This means that (1) is not, in the reductionist's opinion, capable of revealing any ontological commitment to numbers implicit in L. This for the simple reason that there is no such commitment. There are a number of ways to justify a reductionist view of (1). Firstly, one could deny (SS). Since the Fregean infers (OP) from (SS) and (IT), this denial would be sufficient to the denial of (OP). On this view, the sentences introduced by (1), i.e. instances of its LHS, are not sentences at all. They are merely semantically unstructured labels for the corresponding RHS's. Call this view "simple reductionism". Secondly, one could assert (SS) whilst denying (IT). Again, this denial blocks the Fregean's inference to (OP). To hold such a position one must allow that our grasp of instances of the LHS of (1) is derived from a grasp of its sub-sentential expressions and its compositional structure. (IT) is denied by asserting that '=' as it appears in (1) does not express identity. An example of such a view is that under which '=' as it appears in instances (1) is understood to express equinumerousity and under which the referents of the N-terms are taken to be predicates of L. Call such a view "complex reductionism". Finally, one might accept (SS) *and* acknowledge that '=' expresses identity in all the sentences of L* but deny that the same is true of '∃'. Again, this is a denial of (IT). However, it must be motivated by a fairly sophisticated argument for, given a sentence introduced by (1), "Nx:x is a fig = Ny:y is a mango", say, this view accepts that it expresses a genuine identity statement but denies that the inference to "∃x(x = Ny:y is a mango)" is indicative of ontological commitment. Dummett (1973) adopts this strategy by urging that the reference fixed for N-terms by (1) is not the right sort of reference to carry ontological commitment. Obviously, if any of these reductionist views can be defensibly asserted then, even prescinding (as we are) from the falsity of (ET), the Fregean will not be able to achieve her second goal, i.e. to show that the nominalist speaker of L is committed to the existence of the natural numbers. I shall consider each view in turn.[37]

[37] Terminological note: to avoid laborious circumlocutions, "the symbol '=' as it appears in (1)" is abbreviated to '$=_{L*}$'.

Simple reductionism is easily disposed of by our assumption that (ET) is true. However, the truth of (ET) depended upon the truth of (SS) and (IT) and it is one of these that the simple reductionist view challenges: (SS). However, (SS) was described above to be a *stipulation*. To read (1) it is *required* that one conceive of instances of its LHS as semantically complex. I can see nothing wrong with this stipulation. Therefore I do not believe that simple reductionism is tenable.

If this way with simple reductionism is thought too brief, it may be some consolation to learn that no similar quick response is available to refute complex reductionism, the view which denies that '$=_{L*}$' expresses identity. This is not to say that no Fregean has been prepared to give such a response an airing. Wright asserts that:

"...somebody may complain that we seem to be defining a special meaning for '=' for a particular context, whereas its meaning ought to be invariant and general. Frege replies [with Wright's approval]...the definition [(1)], on the contrary, presupposes the standard interpretation of '='; without that presupposition, it will not be possible for the definition to contribute towards explaining the sense of the new kind of singular term." (1983, p107-8)

However, Wright also believes it is incumbent upon the Fregean to check that the symbol for identity retains all the logical properties it has in L after the transition to L* has been effected (1983. p105). I find it hard to see why he should think so. If it is *stipulated* that '=' means in L* exactly what it meant in L, how could it *fail* to retain all its logical properties? It seems to me that it could not. So which word of Wright's should we accept? Rosen (1993) is of the opinion that we should accept the latter. In his opinion, one cannot stipulate that '=' retains its interpretation as identity in the move to L*. Rather, he believes that one must check that '$=_{L*}$' fulfills Wright's logical properties condition and that this will be enough to *show* that the move to L* does not affect the interpretation of '='. More generally, he thinks that it is incumbent upon the Fregean to provide an inferential role semantics for the logical constants of L and to show that the constants mean the same in L* in virtue of validating the same inferences. For Rosen, such a demonstration is enough to justify (IT) (1993, p160-63). I agree with his assessment. Wright is wrong to assert that we can simply stipulate that '=' means identity in *all* sentences of L*. This is because it is enough to found (ET) that we entertain the possibility that '$=_{L*}$' expresses identity. *Supposing* '=' to express identity in all sentences of L* is sufficient to ground the idea of "recarving" the content of an instance of the RHS of (1). Whether '=' *is* interpreted as such in all sentences of L* is something we ought to check.

So what inferential criteria must a sign fulfill if we are to be forced to interpret it as identity?[38] Wright has this to say on the matter:

> "...what is formally required of identity is that it functions as an equivalence relation congruent with respect to every property expressible by an open sentence [of L*] free of epistemic and modal vocabulary." (1983, p105)

A relation, R, is an equivalence relation if it is transitive, reflexive and symmetric. Such a relation is a congruence relation with respect to some predicate, 'F', if, given singular terms 'a' and 'b', 'aRb' and 'Fa' entail 'Fb'.

Now, as matters stand so far in L*, we have no predicates that apply to the N-terms. It is however, easy enough to introduce some. In addition to (1) let the following semantic stipulation hold:

(2) $\forall i=1,....,n$, $\ulcorner \phi_i(Nx:Gx) \urcorner$ means that $\ulcorner F_i(G) \urcorner$, where the φ_i's are congruences for the relation "equinumerousity".[39]

In virtue of this stipulation, we now have a range of predicates that meaningfully bind the N-terms of L*. Furthermore, if "$\exists R(F\ 1\text{-}1_R\ G)$" is true for some F and G and if "$\varphi_i(F)$" is true, for some φ_i, it follows that "Nx:Fx = Ny:Gy" is true, by (1), that "$\phi_i(Nx:Fx)$" is true, by (2) and that "$\varphi_i(G)$" is true, since φ_i is a congruence with respect to equinumerousity. Since "$\varphi_i(G)$" is true it follows that "$\phi_i(Nx:Gx)$" is true. Hence, for any predicate introduced by (2), "Nx:Fx = Ny:Gy" and "$\phi_i(Nx:Fx)$" entail that "$\phi_i(Nx:Gx)$". So the relation expressed by '$=_{L*}$' is a congruence with respect to all predicates that can meaningfully be applied to N-terms.[40]

Unfortunately, this is not going to convince our complex reductionist. She is likely to assert the predicates introduced for N-terms are congruences for the relation expressed by '$=_{L*}$' by construction. Additionally, she is likely to regard the semantic

[38] Note that we really must be *forced* to so interpret "=" everywhere in L* if the Fregean is to defeat this mode of reductionism.

[39] "means that" in (2) should be construed along the same lines as those described for (1) in Section 2.2.

stipulation (2) with as much (reductive) suspicion as (1). Hence, she suggests that the following interpretation should be adopted for L*:

(a) ∀F("Nx:Fx" refers to F"),

(b) $\forall \phi_i(\phi_i$ means the same as $\varphi_i)$,

(c) '=' expresses identity in all sentences of L* other than those in which it is flanked by N-terms, in which case it expresses equinumerousity.

This interpretation is consistent with the stipulations and it is consistent with (SS). It denies (IT) and so blocks the inference to (OP). The problem for the Fregean is that, so far, nothing she has said will refute the claim that this interpretation of L* is correct. How, then, is it to be refuted?

What the Fregean needs to show is that there are predicates of L* which are not introduced by stipulation and which meaningfully bind the N-terms and that all such predicates are congruences with respect to whatever is expressed by '=' in all sentences of L* (in particular, with respect to whatever is expressed by '$=_{L*}$'). This will show that '=' satisfies the logical properties of identity and that this feature cannot be explained by an interpretation like the one given above. So what are these predicates to be? Well, the only predicates that satisfy the description given are those that involve ineliminable occurrences of the cardinality operator. These are predicates like "x is Nx: (Ny:y is a nun)" and so forth (see section 2.1). To show that these are meaningful, the Fregean must, effectively, solve the Caesar Problem, i.e. she must show that all sentences of the form ⌜Nx:Fx = q⌝, q any singular term of L*, F any predicate of L*, have determinate truth-values. Then she will be at liberty to argue that '$=_{L*}$' is a congruence with respect to these predicates, whence, that sign must be taken to express identity. The Caesar Problem is discussed in Section 2.5. Hence, my conclusion with respect to complex reductionism is no more than that, if the Caesar Problem can be solved, then it is untenable.

Assume, for the moment, that the Caesar Problem can be solved. There still remains one form of reductionism to consider. This is the view that blocks the Fregean's inference to (OP) by denying (IT). However, it is distinct from complex

---

[40] Both Rosen and Wright adopt precisely this strategy. See Rosen (1993, p160-1) and Wright (1983,

reductionism in that it denies (IT) by asserting that '∃' does not express ontological commitment when it occurs in positions that bind variables introduced by generalising on N-terms. This is a denial of (IT) since it was assumed in Section 2.1 that '∃' is indicative of ontological commitment wherever it occurs in L. Hence, it could only fail to indicate ontological commitments in the contexts mentioned if its meaning in those contexts is different to its meaning in L. This is Dummett's (1973) view and, as was stated above, it involves some sophisticated considerations from the theory of reference.[41]

Put simply, Dummett's complaint with the introduction of singular terms by sentences like (1), i.e. by the contextual explanation of the senses of a new range of singular terms, understood to refer to a new range of objects, is that the mode of their introduction precludes explaining their reference on the Name-Bearer prototype. For Dummett, this sort of reference is understood to be a substantive relation obtaining between a (use of) a singular term and some objectual constituent of an external reality. In his opinion, only this sort of reference is "robust" enough to support realism concerning the referents of singular terms. Robust reference is contrasted with reference as semantic role. A term has reference in this sense if it makes a systematic contribution to the truth-value of sentences in which it occurs. Now, it seems reasonable to assume that if a term has robust reference, then it has reference as semantic role. Dummett's point is that the converse does not always hold. In particular, it does not hold when the terms in question have been contextually explained in an attempt to introduce a new range of objects. Dummett's view of (1) is, accordingly, that the stipulation of (1) is sufficient to fix a determinate semantic role for the singular terms so introduced, i.e. it is sufficient to fix their reference as semantic role[42], but that it is not sufficient to fix robust reference based upon the Name-Bearer prototype. Since realism with respect to a range of objects can only be supported by

p29-30)

[41] More recently, Dummett (1991) has proposed another view of stipulations like (1) which bears certain affinities to his earlier view. However, the later view denies (SS). It is therefore a version of "simple reductionism" (though his argument in favour of it is by no means simple). Hence, it is unacceptable for the reason described above. I should point out, though, that Dummett's (1991) view of contextual explanations is extremely interesting. However, it is based upon his own interpretation of (CP) which is to be contrasted with Wright's. The reason I do not discuss Dummett's later view is that I have assumed the correctness of (CPR), (CPS) and (RT) as philosophical theses.

[42] So he disagrees with my conclusion in Section 2.2 that (ET) is false.

the attribution of robust reference to the singular terms used to denote them, it follows that (1) is not a sound basis for realism concerning the natural numbers.

This is Dummett's view in brief. The question we must answer is whether the Fregean can refute it. I will here sketch a response which it seems to me may provide grounds for that refutation. Unfortunately, there is not space to do more.

Let us examine, in a little more detail, the content of Dummett's distinction between robust reference and reference as semantic role. It is particularly important to be clear about why contextually explained ranges of singular terms cannot lay claim to robust reference. Dummett asks:

> "If the sense of a name is not to be given, or not always to be given, in the form of a criterion for identifying an object as the bearer of the name, how, then, is it to be given? ... what becomes of the realism embodied in the use of the name/bearer relation as the prototype of reference and the principle that the referents of our words are what we talk about? In what sense are we entitled to suppose that abstract objects are constituents of an external reality, when possession of reference by their names has been interpreted as a matter wholly internal to the language?" (1973, p65-7)

Dummett's point is that abstract objects cannot be construed to be independently existing entities. This is because the contextual explanation of singular terms, the only mode of explanation that seems to be available for the explanation of singular terms for abstract objects, does not provide the thinker who receives the explanation with a criterion for identifying the referent of an abstract singular term when it appears in another guise. She has no conception of what it would be to encounter that object again. But this is precisely the sort of conception that is embodied in reference construed on the Name-Bearer prototype. Therefore, the fact that we speak of contextually explained singular terms having reference should not be taken to mean anything more than that we have specified a determinate semantic role for them, than that we have specified how they contribute to the truth-value of sentences in which they appear via the equivalence of those sentences to other sentences in which they do not appear. That is why their interpretation is a matter "wholly internal to language" and does not support realism concerning their referents.

On behalf of the Fregean, we can observe that a solution to the Caesar Problem might refute this view. This is because if we are able, on the basis of (1), to evaluate mixed identity statements, it could be argued that we do acquire, on the basis of (1), a criterion for identifying the referents of contextually defined terms. Of course, Dummett might respond that actually this shows nothing more than that (1) fixes a determinate semantic role for N-terms in all the sentences of L* and that this can still be construed of as semantic role without robust reference. However, that response is likely to cause all sorts of problems when it comes to giving a semantic theory for L*. The semantics of "Nx:x is a planet = Ny:y is a star", "Jezebel = Nx:x is a harlot" and "Jezebel = Delilah" will all have to receive different explanations. As Benacerraf has famously urged (1973), such complexity in semantic theory is to be avoided. Nonetheless, Dummett may be able to accept a solution to the Caesar Problem and yet still adhere to his view of (1) (that view is therefore more subtle than the complex reductionism considered earlier).

The question of whether the Fregean can rebut Dummett's view of (1) is thus the same as the question of whether she has a response that does not depend upon her solution of the Caesar Problem. I think she may. The Fregean can claim that it is simply not true that the "possession of reference by [the names of abstract objects] has been interpreted as a matter wholly internal to the language [they occur in]." We saw in section 2.2 just how the (strong) sense of sentences containing arithmetic vocabulary is to be explained on the basis of (1). This specification is effected by our understanding that the left and right hand sides of an instance of (1) are identical in weak sense. We are supposed to grasp that weak sense and impose upon it the logical form $\ulcorner$aRb$\urcorner$ where the schematic letter 'R' is supposed to express identity and the places indicated by the schematic letters 'a' and 'b' are occupied by the N-terms *construed as robustly referring terms*. This is what, together with the supposition that 'R' expresses identity, is supposed to yield our grasp of the strong sense of instances of the LHS of (1). It is therefore built into the Fregean reading of (1) that the N-terms be construed as robustly referring terms. This conception of those terms partially determines the identity of the senses of the arithmetical sentences introduced by (1). Under the methodological assumption that (ET) is true, and that the picture sketched by the Fregean as to why it is true is correct and defensible, it follows that there is no possible gap between fixing the meaning of the N-terms and endowing them with robust reference (provided the RHS's of instances of (1) are sometimes true). So Dummett's view cannot be maintained.

To repeat: this is only a sketch of the response with which I think the Fregean should meet Dummett's challenge. However, I do believe that it might be made to work. As a result, I conclude that neither of the forms of reductionism considered, nor Dummett's view, are viable accounts of the Fregean's explanation of the move from L to L*. Under the assumptions that (ET) is true and that the Caesar Problem is solved (which we had to assume to refute complex reductionism), the Fregean's account of (1) seems irresistible. Under those assumptions, (OP) must be accepted. having so concluded, it is now time to turn to the second of those assumptions, to question whether the Fregean can solve the Caesar Problem.

2.4 Give (in) unto Caesar...

Recall that the Caesar Problem is a challenge leveled at any proposed definition or explanation of number. In the case we are considering, the contextual explanation of the cardinality operator, the challenge will have been met if the explanation fixes the concept of number such that every sentence of the form ⌜Nx:Fx=q⌝, where q is any singular term in L*, has a determinate truth-value. This is because (as was stated in Section 1.2) the Fregean aims to fix a sortal concept of number. Only if the concept of number fixed is sortal will it be correct to infer that numbers, if they exist, are objects. Since a sortal is a concept with which is associated a criterion of identity and distinctness, and since this criterion is to distinguish between instances of that concept and between instances of that concept and other concepts, it follows that the condition just mentioned must be imposed upon the Fregean's explanation of the concept of number.[43]

According to Wright, "a Caesar Problem" arises for any concept for which the criterion of identity is given by an equivalence relation defined over a domain which includes objects which are not instances of the concept. In other words, a Caesar Problem arises for all Sentences when these are construed as contextual explanations of the functional operator that occurs on their LHS (see n12 for the definition of a "Sentence"). This suggests that a solution to the Caesar Problem when relativised to (1) must be legitimised by an independently well motivated and plausible principle which gives the form of response in the general case. Hale (1994) has aptly named this principle the "Principle of Sortal Inclusion", I shall adopt that terminology here (though shortening it to (PSI)). Wright's suggestion for (PSI) is this:

> "Gx is a sortal concept under which instances of Fx fall if and only if there are, or could be, terms, 'a' and 'b', which recognisably purport to denote instances of Gx, such that the sense of the identity statement, 'a=b', can be adequately explained by fixing its truth-conditions to be the same as those of a statement which asserts that the given equivalence relation holds between a pair of objects in terms of which identity and distinctness under the concept Fx is explained." (1983, p114).

This may be rather a mouthful but the intuition behind it is reasonably simple. Suppose F is the concept *shape*. Then the criterion of identity for F's is given by the

[43] In fact, exactly what the Caesar Problem is may be a somewhat more complex matter than I have made it appear. Nonetheless, I do not believe that my account of the problem distorts it in any significant way. See Heck (1997a) for more on this issue.

equivalence relation of geometrical similarity, as defined on figures (which may be taken to be concrete inscriptions of squares, circles, triangles, etc.). So the following Sentence expresses the criterion of identity for shapes:

($S^=$) ∀x∀y((The shape of x = The shape of y)↔(x is geometrically similar to y)).

Note that this Sentence could be used, according to the Fregean, to explain the concept of shape (See n12). Now suppose that G is the concept *person* and let "Jenny" and "Jim" be singular terms purporting to refer to persons. Wright's idea is that, since the sense of "Jenny=Jim" could never be explained by appeal to facts about geometrical (dis-)similarity on figures, Jenny and Jim cannot be shapes. But since "Jenny" and "Jim" were arbitrary terms, which had to meet only the condition that they purportedly referred to persons, it follows that no people are shapes. Hence every statement of the form ⌜The shape of x=q⌝ is false when q is a person-denoting term. Similarly with all sortals whose criteria for identity and distinctness are distinct from those for the concept *shape*. Therefore all sentences of the form ⌜The shape of x=q⌝ are false unless q purports to refer to a shape. Thus, the Caesar Problem for shapes is solved. By analogy, the Caesar Problem for numbers is solved too. Setting 'Fx' in (PSI) identical to number and invoking (HP) to specify the equivalence relation to be used yields Wright's principle of sortal inclusion for numbers:

"N^d: Gx is a sortal concept under which numbers fall (if? and) only if there are, or could be, singular terms '**a**' and '**b**' purporting to denote instances of Gx such that the [sense] of '**a=b**' could adequately be explained by some statement to the effect that a 1-1 correlation obtains between a pair of concepts."[44] (1983, p117).

Since this principle is available apriori, it follows, assuming that the argument given above is sound, that a nominalist speaker of L is in a position, having moved to L*, to solve the Caesar Problem for number.

The question we must answer, then, is whether the argument described above is sound. I do not think that it is. Firstly, it is not at all clear what to make of Wright's

[44] The bracketing of and querying of "if and" in this quotation reflects the fact that we may not want to prejudice questions concerning whether the numbers can be identified with any other mathematical/abstract objects, e.g. sets. I have added the square bracketed "sense". This is because it is clear that Wright regards N^d as an instance of (PSI). It is only such if the word sense is included as I have included it.

idea of terms "purporting to refer" to instances of a particular sortal. Dummett, in particular, takes issue with this idea, saying that the notion must be taken to mean that the way in which an object is canonically referred to is a property of that object (1991, p159-166). He complains that Frege rejected just this suggestion, saying that he refused:

"...to take the easy way out by simply stipulating that no term formed by means of the [shape-operator] shall be taken as standing for an object denoted by a term of any other kind." (p160)

However, I am not going to discuss Dummett's objection, even though I am, to a degree, in sympathy with his view. It is not clear to me that any terms "purport to refer" only to a certain sort of object. Nonetheless, let us suppose that we can make sense of the notion and that terms do have this property. Even under those assumptions, there is a telling objection to Wright's proposal, suggested by Rosen (1996).

Rosen's point is that even if the sense of a statement of personal identity cannot be fixed by reference to a statement concerning equinumerousity on concepts, that is not enough to show that no referent of a person-denoting term is also the referent of a number-denoting term. The most that Wright's principle N^d can demonstrate is that singular terms that purport to refer to people never have the same sense as singular terms that purport to refer to numbers. However, the *foundation* of the sense-reference distinction is that different senses may determine the same reference. Therefore, just because the sense of a statement asserting the identity of persons cannot be "adequately explained" by fixing its truth-condition to be the same as that of some statement concerning equinumerousity on concepts, it does *not* follow that no people are numbers. That would follow only if one had already demonstrated that the sense of a singular term purporting to refer to a person could never determine the same referent as a singular term purporting to refer to a number, i.e. it would follow if one had already shown that the concepts of number and person are incompatible. But that demonstration, assuming it would generalise to other sortals, would be a solution of the Caesar Problem for numbers. Hence, Rosen concludes that Wright's solution presupposes what it sets out to ascertain.[45]

In response to this objection, the Fregean will probably demand an explanation of how the senses of singular terms for people could refer to numbers and yet not be explicable on the basis of equinumerousity amongst concepts. Certainly, it does seem implausible to say that the truth-conditions of "Caesar = Augustine" and "0 = 1" could be the same yet their senses not admit of the same explanation, viz. that which utilises "$\forall R(Nx: \neg(x=x)\ 1\text{-}1_R\ Nx: x=0 \vee \neg(x=x))$". But this appearance is based upon a confusion. A class of sentences all of which share their truth-condition need not be a class of sentences each of whose senses can be explained on the basis of the sense of another, where an admissible explanation must cohere with the basic assumptions of Fregeanism. It may be the case that the truth condition of "0 = 1" is "$\forall R(Nx: \neg(x=x)\ 1\text{-}1_R\ Nx: x=0 \vee \neg(x=x))$". Even so, we concluded (see Section 2.2) that there was no route available to the Fregean to extract the sense of the former from the sense of the latter. The routes suggested were both found to be inconsistent with (CPR) and (RT).

The picture described in Section 2.2, that which was to explain how (ET) was to be understood, can be invoked here to clarify matters. Under the assumption that that picture coheres with the Fregean's other claims, in particular with her assertion of (CP), it looks as if the Fregean has a response to Rosen's objection. Recall that our understanding of (ET) with respect to (1) was supposed to be this: that we pass from a grasp of the strong (compositional) sense of the RHS of an instance of (1) to a grasp of its weak (non-compositional) sense; that we understand (1) to assert that this weak sense is common to the corresponding instance of the LHS and that we "recarve" this weak sense to obtain the strong sense of the LHS . How could this picture be used to overcome Rosen's objection? Clearly, what one must do is to show that the weak sense of a statement of, say, personal identity is never identical to the weak sense of a statement to the effect that two numbers are equal. From this we could conclude that the weak sense of the former could not be carved up to yield the strong sense of the latter, and *vice versa*. Consequently, it would be plausible to suggest that the sense of

[45] Though he nobly remarks that, "There may be a construal of Wright's remarks that I have missed." (1993, p174). I am not convinced there is.

no term referring to a person could refer to a number.[46] Hence, we could conclude that the concepts person and number were incompatible. It is clear that this argument would generalise to the case of other sortals, thus the Caesar problem for numbers would be solved. This is precisely Hale's response to Dummett's objections to the putative solution (Hale, 1994). With Wright's consent (Hale, 1994, p131, n6), he reparses (PSI) in accordance with the picture described in Section 2.2, derives a correlative N^d, and follows it with an assertion which beautifully captures the strategy recently suggested:

> "This general principle, together with $N^=$ [(HP)], gives Wright's principle N^d: that instances of the sortal concept *number* can be among the instances of another sortal concept G only if some statements of identity linking G-terms have the same content [weak sense] as appropriately corresponding statements asserting that a 1-1 correlation obtains between certain concepts. Since the content [weak sense] of an identity-statement linking person-denting terms is never that of a statement of 1-1 correlation among concepts, it follows that numbers aren't people."[47] (1994, p131)

The question is, of course, how do we know that the weak senses of statements of personal identity are never identical to the weak senses of sentences of the form $\ulcorner \exists R(Nx{:}Fx\ 1\text{-}1_R\ Ny{:}Gy) \urcorner$? If we do know this, then perhaps Hale's strategy will work and the Caesar problem will be solved. However, we concluded in Section 2.2 that the suggested account of (ET) was not acceptable because it was inconsistent with (CPR) and (RT). This was because it was found to depend upon our ability to pre-representationaly grasp (unstructured) weak senses. The claim to have this ability is what was found to be inconsistent with (CPR) and (RT). So even if the weak senses of statements of personal identity were never identical to those of sentences of the form $\ulcorner \exists R(Nx{:}Fx\ 1\text{-}1_R\ Ny{:}Gy) \urcorner$, and even if it could be shown that this demonstrated the incompatability of the concepts of person and number, because of (CPR) and (RT), the Fregean is not in a position to make use of these facts. Hence, there is no solution to the problem of whether any people are numbers that could be generalised to the case where number is contrasted with other sortals. I can see no other way to respond to Rosen's criticism. Neither are there any other responses to the Caesar Problem in

[46] I say "plausible" here but I do not know for certain that this argument will work. I am just describing what seems to me to be the only possible response which the Fregean has to hand.

[47] In fact, Hale had not yet developed his notion of weak sense which is to be found in (Hale, 1997). However, as the later paper builds upon the earlier, I assume he would have no objection to my account of his view.

the literature. Therefore, I conclude that the Fregean cannot overcome the Caesar Problem.

2.5 Conclusion

The Argument I had as fundamental assumptions (CPS), (CPR) and (RT). It had three objectives: to dispel the epistemological doubts associated (by Conservative and Radical Nominalists) with our grasp of arithmetical concepts and our recognition of arithmetical truth; to show that Platonism is true by demonstrating the independent, objectual existence of the natural numbers and to show that some form of Logicism can be upheld, at least with respect to number theory. We were interested in the first two of these which the Argument I was to achieve by the stipulation of (1). We were construe (1) as a contextual explanation of the sortal concept number which, at the same time, was to fix epistemologically unproblematic truth-conditions for some sentences containing singular terms for numbers. Unfortunately, it failed to deliver the goods.

This conclusion has been reached by identifying the claims that a Fregean must make with respect to (1), and, more generally, the move from L to L*, in order to support the Argument I. These were (ET), (SS) and (IT). We also identified the condition that was to be satisfied in order for (1) to have satisfactorily explained a sortal concept of number. This was that (1) must furnish us with the means to solve the Caesar Problem.

In Section 2.2, I argued that (ET) is inconsistent with the Fregean's fundamental assumptions (CPR) and (RT). Given the truth of these assumptions, it followed that (ET) was false. The argument depended upon determining exactly what (1) should be taken to assert. This was the identity of the weak sense of the left and right hand sides of an instance. Given this refinement of our understanding of (1), it was possible to suggest two explanations of (ET). However, both of these explanations were found to depend upon an ability that we cannot have if (CPR) and (RT) are true, the ability to grasp weak senses. Hence, the Argument I fails to explain how we can acquire arithmetical concepts. That we cannot was one of the doubts motivating Radical Nominalism. Therefore Section 2.2 shows that the Argument I cannot attain its first objective.

In Section 2.3, I distinguished between three forms of reductionism that someone might hold with respect to L*. On the assumption that (ET) is true, I suggested that, given a solution of the Caesar Problem, all three forms of reductionism could be rebutted. However, in Section 2.4, we found that a satisfactory solution of the Caesar Problem depended upon exactly the same ability which the truth of (ET) depended upon, the ability to grasp weak senses. This was not really surprising since if (ET) was true, there would be no reason to expect the Caesar Problem to present an insurmountable challenge. However, given the conclusion of Section 2.2, we had to recognise in Section 2.4 that Wright's solution to the Caesar Problem is insufficient, that it could not be defended against Rosen's objection. This opens the door for the second form of reductionism considered in section 2.3. Therefore, the Argument I fails to attain its second objective, to show the independent objectual existence of the natural numbers.

Chapter 3

The Argument II

3.1 Introduction

I concluded in the last chapter that the Fregean Argument fails on its first reading. That is to say, if (HP) is taken to be stipulatively true, as a definition or explanation of new vocabulary or concepts, then the Fregean cannot justify her claims. Recall from Chapter 1 that the Argument I was supposed to demonstrate three things: that number-theory is not epistemologically problematic in the ways that nominalists suppose it is (i.e. that we can acquire the concepts necessary to discourse on the numbers, that we can refer to them and that we can know about them); that natural numbers exist and that number-theoretic logicism can be upheld. Recall, also, that we were interested only in the first two claims. The failure of the Argument on its first reading means that the Fregean has so far substantiated neither of those claims. It is now time to consider the second reading of the Argument, the reading which claims that (HP) can be known apriori. Under this reading it is assumed that we grasp all the concepts involved in (HP), no explanation of how we acquired them is given. As was stated in Chapter 1, it follows that the argument on its second reading, and without any augmentation, can at most substantiate the ontological claim, that the natural numbers enjoy mind-independent objectual existence, and the claim that we can know about them. "The Argument II" is less ambitious than "the Argument I", perhaps, as a result, it has more chance of success.

For (HP) to be knowable apriori, it is necessary that apriori knowledge be possible. So much is obvious. It is therefore reasonable to pose the question, "What sort of sentences or contents can be known apriori?". There have been various answers to this question. Perhaps the most well known is this: analytic sentences, those that are true in virtue of meaning, are apriori. Of course, as it stands this answer is hardly satisfactory. It is too vague in content to admit of precise evaluation. Nonetheless, I think it is clear, on the basis of this very imprecise formulation of the thesis, that some theory of apriority that founds apriority in analyticity is most likely to be of service to the Fregean in the present context. This is because the assumption that distinguishes the Argument II from the Argument I is (as stated above) that we grasp the concepts required to form number-theoretic beliefs. These concepts are really all the Fregean has to work with in justifying her claim that (HP) is apriori. So, under the assumption that these concepts are the meanings of number-theoretic vocabulary, it follows that the most plausible route to the conclusion that (HP) is apriori is to argue that it is true in virtue of the meanings of the words it contains, i.e. to argue that it is analytic.

The term "analytic" was first introduced by Kant. For him an analytic judgeable content is one in which what is predicated of the subject is already contained in the subject. Of course, not only analytic judgeable contents are apriori for Kant. He also believes that there exist apriori *synthetic* judgements (where a synthetic judgeable content is simply one that is not analytic). It took the *Critique of Pure Reason* to justify(?) this claim. There is one rather obvious reason why Kant's view of apriority is not going to appeal to the Fregean. (HP) is not a subject-predicate sentence, so the Kantian definition of analytic does not apply.

Perhaps we might use Frege's characterisation of analyticity. He urged (*Grundlagen*, §3) that an analytic content is one whose proof depends only upon general logical laws and definitions and that analytic contents can be known apriori. More recently, Boghossian has developed a theory of apriority which can be seen as an reparsing of Frege's view (Boghossian, 1997). For Boghossian, *Frege-analytic* sentences are apriori. Such sentences are those that are reducible to logical truths by the substitution of synonyms for synonyms. However, both Frege's view and Boghossian's development of it are not appropriate for use in the Argument II. Firstly, (HP) is not a definition of the concept number, as concluded in the previous chapter. Hence, there can be no one-line proof that demonstrates the analyticity of (HP) under the Fregean account. Secondly, it is not at all clear that (HP) could be reduced to a truth of logic by the substitution of synonyms for synonyms. As far as I can tell, such a reduction would have to presuppose (HP) *as a definition*. Thus, there can be no many-lined proof that demonstrates its analyticity under the Fregean account and Boghossian's does not apply. Thirdly, I have presented the Argument II without any commitment to answering the vexed question of whether arithmetic is reducible to logic. Adoption of Boghossian's account prejudices that question in favour of an affirmative answer. Adoption of Frege's looks to so much the same (though it is not as obvious). But this would make the Argument II more ambitious than necessary. Hence, it would have less chance of success.

So is there a theory of apriority based on analyticity that the Fregean can appeal to? I think there is, the account proposed by Peacocke (1992b). His account differs from the three mentioned above. Contra Kant, he holds that all apriori knowledge is analytic and that not all analytic contents need be in subject-predicate form; contra Frege and Boghossian, there is no commitment to the reducibility of apriori knowledge to logical knowledge. Nonetheless, there are similarities. Peacocke agrees that Kant's synthetic apriori judgments are apriori (or, at least, agrees that mathematics is apriori), he just gives a different explanation of this fact. Furthermore, everything classified as apriori by the Frege/Boghossian approach is classified as such by Peacocke's theory. In light of these contrasts and comparisons, it is possible to see that Peacocke's theory is the widest of any that I have mentioned, i.e. it classifies the most contents as apriori. Furthermore, as I will explain in Section 3.2, it is a theory of apriority based on analyticity. Peacocke's theory, then, seems to offer the best chance to the Fregean who is concerned to show that (HP) is apriori.

Unfortunately, adopting a theory of the apriori based on analyticity puts the Fregean in range of certain detractors of such theories. In the general case, Quine has argued against any such theory, in fact, against the possibility of apriori knowledge in general (Quine, 1936, 1951, 1963). More specifically, Boolos has urged that, whatever we think of analytic theories of the apriori, (HP) should not be considered analytic (Boolos, 1990, 1997). There is not space in a study of this length to address the sceptical (Quinean) attack. However, it is incumbent upon us to consider Boolos' objections to the analyticity (and, hence, apriority) of (HP) for these are addressed directly to the Fregean. The remainder of this chapter is therefore structured as follows: in Section 3.2 I outline Peacocke's theory of the apriori, there will be space neither to criticise nor to defend the view; in Section 3.3 I show that (HP) is (very probably) apriori under that account (this might be called the positive phase of my answer to the question of whether (HP) is apriori); in Section 3.4 I explain and try to answer Boolos' objections to the analyticity (hence, apriority) of (HP) (this might be called the negative phase); finally, in Section 3.5, the strands of argument are drawn together to see what can be made of the Argument II and my conclusion is put into the wider context of this study.

3.2 Peacocke's Theory of the Apriori

Peacocke's theory of the apriori is an application of his theory of concepts (Peacocke, 1992). The latter theory is concerned with concepts construed as Fregean senses or modes of presentation. Such things are to be individuated by *possession conditions* (PC)s, non-circular accounts of what must be true of a thinker for him to possess a given concept. These will all be of the form:

The concept F is that concept C to possess which a thinker must meet the condition A(C).[48](1992a, p6)

For logical concepts, the conditions for possession will mention certain inferences which a thinker will find "primitively compelling" in virtue of their form, for other concepts the condition will mention certain belief-forming practices to which a thinker will conform. Assuming that words express concepts, it follows that Peacocke is committed to giving an inferential role semantics for the logical constants and a conceptual role semantics for other words.[49] A (PC) is non-circular if the concept it individuates does not occur in the thinkers propositional attitudes which appear in the condition A(C) (1992a, p9).[50] This will become important later (Section 3.3).

To ensure that contents formed from Peacockian concepts are apt to carry truth, Peacocke invokes an analogue for the Fregean claim that sense determines reference. He asserts that, for each (PC), there exists a *value-fixing relation* that holds between the (PC) and an entity when that entity is the semantic value of the concept individuated by the (PC) (1992a, p16-8). The existence of this relation does not

[48] In fact, these will not all be of this form, but we can manage here without the refinements necessary to Peacocke's theory.

[49] Both an inferential role semantics and a cognitive role semantics explain the meaning of a sub-sentential expression by identifying certain belief-forming practices that are to be regarded as constitutive of its identity. The distinction between the two lies in what is taken as the class of belief-forming practices from which the practices constitutive of meaning are drawn. Put very roughly: for an inferential role semantics, this class is a class of classes of inferences each member of which is the class of inferences that a thinker is disposed to draw from a given sentence containing the sub-sentential expression; for a cognitive role semantics, the class is the aforementioned class *unionised with* the class of sentences containing the sub-sentential expression that she is disposed to believe given certain experiential evidence (e.g. perceptual experience).

[50] The non-circularity condition, under the assumption that the meanings of words are Peacockian concepts corresponds to Dummett's claim that a theory of meaning should be "full-blooded" rather than "modest", i.e. it should explain in what an understanding of language consists rather than just describe that understanding. See Dummett (1975).

depend upon the existence of any thinkers. What does so depend is the existence of *determination theories* (DT)s. These are correct specifications of the value-fixing relation for any given (PC). As Peacocke has it:

"A determination theory ... states how a possession condition, commonly together with the way the world is, determines semantic value for the concept individuated by the possession condition."(1992b, p177)

It is a constraint upon the legitimacy of any given (PC) that it is possible to supplement it by giving a true (DT) for the concept it supposedly individuates. The crucial thesis linking (PC)s to their associated (DT)s is that, given a legitimate (PC), the associated (DT) licenses the assumption that any inferential or belief-forming practices mentioned in the (PC) are correct. So if Σ(C) are the conditions mentioned in the (PC) for the concept red, say, some of which mention belief-formation, then the (DT) will assign to that concept a semantic value which, if a thinker's use of the concept conforms to Σ(C), results in the formation of true beliefs of the kind mentioned. There is a great deal more to say about Peacocke's theory of concepts in explanation, criticism and defense. As we are concerned with his theory of apriority, I shall not pursue that discussion here. However, it should be noted that the condition upon the legitimacy of a (PC), that it be possible to give a *true* associated (DT), in fact yields the analogue of a strong relation between sense and reference. Under this conception of the relation, a sense that fails to determine a reference (a (PC) for which there is no true (DT)) is not a sense at all (is not a legitimate (PC) so fails to individuate a concept). A well known champion of this sort of relationship between sense and reference is Evans (1982). But note that there have also been champions of a weaker relationship, notably Dummett (1971), under which a sense is construed as a way of fixing reference which may fail, i.e. which may not actually determine a reference but is an acceptable sense for all that. If this were Peacocke's conception, then it would not have been important for (DT)s to be true.

Having sketched the background required from Peacocke's theory of concepts, it is possible to describe his account of the apriori. I shall first describe it in outline. Peacocke's approach to apriority depends upon distinguishing between apriori truth and apriori knowledge. For him, the apriority of a given truth is independent of any facts about thinkers who assent to that truth. Accordingly, Peacocke's explanation of apriori truth is a non-epistemic account under which apriority is a property of (some) truths which they possess (or don't possess) independently of whether and why they are believed. In contrast, the apriority of knowledge is relative to thinkers and is an overtly epistemic notion. Whether a thinker knows a content apriori is a fact about her which will obtain (or not) depending upon facts about her justification for that content.

The aim of Peacocke's theory is to give an account of apriori truth, where a truth is understood to be a content formed from concepts, which explains how apriori knowledge is possible. To that effect he first proposes his account of apriori truth, the *metasemantic account* (MA). This theory is supposed to give the fundamental, philosophical explanation of the apriori status of any given class of truths. In order to explain how (MA) is related to theories of apriori knowledge he states what he calls the intuitive criterion of apriority which I call (IC). This is the thesis that:

(IC) A content is known apriori if it is known to be true in the actual world independently of any particular experience. (1992b, p174)

Peacocke elucidates (IC) by asserting that it classifies a thinkers knowledge of a content as apriori if their justification for that content does not mention any particular experience. The second component of Peacocke's programme is to show that those contents classified apriori by (MA) are those that are knowable apriori in accordance with (IC). The effect of such a correspondence between (IC) and (MA) will be to show the possibility of apriori knowledge. There are, I think, two main reasons to hope for such an achievement. Firstly, as will become clear below, whether a given truth is apriori under (MA) depends upon facts about the (PC)s and associated (DT)s for the contents constituent concepts. So, for us to know that a given truth is apriori under (MA), we must be able to specify those (PC)s and (DT)s. Peacocke himself admits that, for many concepts, this is no easy task. However, if the putative correspondence holds, there is no need to complete it for, under this condition, if one is justified, in a sense strong enough to support knowledge, in one's belief in a given content and if (IC) classifies one's justification as supportive of apriori knowledge, then that justification is sufficient to account for one's apriori knowledge of the content. One does not need to show that the content is apriori under (MA). Secondly, though this point is only implicit in Peacocke's work, a satisfactory theory of apriori knowledge should not explain apriority in such a way as to render the acquisition of apriori knowledge impossible for anyone save a philosopher. If it were the case that to claim apriori knowledge of a content it was necessary to show that the content was apriori under (MA), precisely that situation would obtain. However, if the putative correspondence between apriori truth and apriori knowledge holds, the difficulty is avoided. One need not show that a truth is apriori under (MA) in order to claim apriori knowledge of it in accordance with (IC). The final component of Peacoke's programme, which I shall not describe, is to refine (IC) into a precise and defensible philosophical thesis.

(MA) is the following thesis, where K is a class of contents all of which are supposed to be apriori:

"The a priori status of a content in the class, K, is fundamentally explained by the fact that the possession conditions for the concepts from which those contents are formed, together with the corresponding determination theories for those concepts, jointly guarantee that the content is true in the actual world." (1992b, p179)

In effect, the claim amounts to this: if it is possible to derive (i.e. prove) the truth of a content from the truth of the (PC)s and associated (DT)s of the concepts involved in the content, then that content is apriori. However, as Peacocke rightly observes, no one is going to accept this account of apriori truth unless (PC)s and (DT)s are themselves apriori truths. Peacocke argues that they are and that this can be shown by applying (MA). So, to show that some given (DT) is apriori, one specifies (PC)s and (DT)s for all the concepts it involves and proves its truth from those conditions. It should be clear now why Peacocke's theory is a theory based on analyticity. Assuming that Peacocke's concepts are the meanings of words, (MA) explains the apriori truth of sentences by reference to what its constituent expressions mean.

In order to show that the contents classified apriori by (MA) are the contents that can be known apriori under (IC), it is necessary to clarify the relationship between the metasemantic proof of the truth of an apriori content and a justification invoked by a thinker to found a knowledge claim with respect to that content. Peacocke does not explicitly state the relationship between the theories of apriori truth and apriori knowledge. However, he does draw a distinction between types of justification and suggests a link between (PC)s and justifications both of which are relevant to an understanding of how he conceives that link (the link between (MA) and (IC)).

According to Peacocke, there are two sorts of justification, *ground-level justifications*, which do not mention facts concerning the (PC)s and (DT)s associated with those concepts which constitute the content which it justifies, and *metasemantic justifications*, which do (1992b, p179). Intuitively, the former are the sorts of justifications that a non-philosopher might appeal to in order to justify a content, the latter are those that philosophers might invoke. To explain the link between justifications and (PC)s it is necessary to get a little clearer about what justifications are (or, at least, can be taken to be). Peacocke takes a justification to be a set of mental states and propositional contents (1992b, p190). To facilitate the discussion, he limits his attention to direct sufficient justifications. These are sets of mental states and propositional contents which do not require the addition of further contents or states in order to justify the content in question (they are direct) and which are conclusive, i.e. strong enough to support knowledge claims (they are sufficient). With respect to the link between (PC)s and justifications, Peacocke states that:

> "When a given set S of contents and states provide a direct, sufficient justification for accepting a given content, p, the following is a consequence of the possession conditions for the concepts in the contents involved: accepting the contents in S, together with enjoyment of the states in S, is sufficient to lead a thinker to judge that p, if the question arises." (1992b, p190)

So S is a justification for p precisely because possession of the concepts involved in S entails that, if enjoys that states in S and believes in the contents it contains, one will form the belief that p. Furthermore, it is possible to explain why, given this link between (PC)s and justifications, the disposition to form the belief in p should be explained on the basis that S is a justification for p rather than for some less epistemically respectable reason. Recall that a legitimate (PC) must be associated with a true (DT). Recall, also, that this entailed that the belief-forming practices mentioned in a legitimate (PC) are truth-preserving. It follows that if all the contents in S are true and believed by a thinking subject who is in the sates in S, then p will be true. So the thinker is disposed to form true beliefs on the basis of S. This is a very attractive result which explains why S is a justification.

The suggested link between (PC)s and justifications has two consequences, the second dependent of the first. To begin with, it suggests a picture of how any theory of justification might go about explaining the legitimacy of ground-level justifications for beliefs. The picture links the justification for a belief with the way that belief is formed. Thinkers are put into certain mental states by their experience and, as a result, are disposed to form certain beliefs, provided they do not believe anything else with which those beliefs do not cohere. In order to give a ground-level justification of that belief they can merely list the mental states they were in from which they derived their new belief. In light of the link between justifications and (PC)s, we may say that to philosophically explain the legitimacy of a ground-level justification, i.e. to explain why the derivation of some content p from a set of mental states and contents S is a good way to acquire beliefs, we need to give a metasemantic justification of the content which the ground-level justification is supposed to justify. This will be a derivation of the content from the states and contents in S together with (PC)s and (DT)s for the concepts they involve. Derivatively, (MA) can be seen as the theory which philosophically explains why a belief formed without recourse to particular experience can be a justified belief. This is because no experience is involved in such a beliefs formation, only reflection upon the concepts it involves. So a ground-level justification for that belief will not mention experience, only facts about concepts. This entails that the metasemantic justification for that belief, the philosophical explanation of why forming beliefs in that way results in justified beliefs, will derive the content of that belief only from facts about (PC)s and (DT)s. But this is the proof of the contents truth from the relevant (PC)s and (DT)s which classifies it as an apriori truth under (MA). Since such a belief is known apriori, according to (IC), it follows that what is knowable apriori in accordance with (IC) is what is apriori true under (MA). This completes my description of the first two thirds of Peacocke's account of the apriori. As stated above, I am not concerned here with the third part, to refine (IC) into a precise and philosophically defensible thesis.

3.3 Could (HP) be Apriori?[51]

So much for an outline of Peacocke's account of apriority. We turn now to the positive phase of our answer to the question of whether (HP) is apriori. For the purposes of this section, we assume that (MA) is perfectly acceptable. Of course, that is no trivial matter. I have only sketched the theory and have done nothing to persuade the reader that it should be accepted. Much ought to be said given enough space to say it. Nonetheless, my strategy is to prescind from any debate concerning (MA) and to make the case for classifying (HP) as apriori. This is because I think that if Peacocke's theory can be defended against the objections brought against it, then there is good reason to think that (HP) is apriori. That is to say, it looks as if (HP) *is* apriori on Peacocke's account.

What is required to show that this is so is a proof of the truth of (HP) from the (PC)s and (DT)s of its constituent concepts. Without such a proof, ground-level justifications for (HP) will not be vindicated, so non-philosophers will not be able to claim apriori knowledge of (HP). Also, without such a proof, the sceptic is hardly going to accept that (HP) is apriori. Thus, it is necessary to specify those (PC)s and (DT)s. With respect to the logical vocabulary of (HP), this specification amounts to the description of an inferential role semantics for the logical constants (as was noted in the previous section). I am not going to give such a semantics because I don't have one. Therefore, I assume that we have such a semantics to hand. More precisely, I assume that we have an inferential role semantics that explains the meanings of the logical connectives in this set: $\{\neg, \wedge, \vee, \rightarrow, \leftrightarrow, \exists, \forall, =\}$. To turn such a semantics into (PC)s for the concepts expressed by the constants, it is sufficient to slot the inferences constitutive of the meaning of each constant into the condition A(C) of the associated (PC), this will state that the inferences are found "primitively compelling" by a thinker who grasps the concept.[52] I shall call the possession condition for each constant, \$, "PC(\$)". An analogous convention is adopted for the associated

[51] Terminological note: in this and the following sections it is necessary to distinguish between Fregean and Peacockian concepts. I denote the former with the word "Concept", the latter with the word "concept".

[52] In fact, this is not quite right, matters are not this simple. Peacocke has argued (1987) that the inferences mentioned in possession conditions for the logical constants need not be found primitively compelling, but that they must be found primitively compelling or compelling upon a certain amount of reflection. Exactly what this means and just why he said so are niceties that need not detain us here.

determination theories, each of which will state that the semantic value of logical concept $ is that logical object which makes the inferences mentioned in (PC$) truth preserving. That such a theory, (DT$), exists and is true for each of the constants in the set mentioned is assumed.

With (PC)s and (DT)s for the logical concepts assumed, it remains to specify a (PC) and (DT) for the concept number. Call these "(PCN)" and "(DTN)". To formulate a plausible candidate for (PCN), it is necessary to identify what conditions are necessarily fulfilled by any (PC) in order that it be a candidate. Hence, it is necessary to identify what belief forming practices can reasonably be considered *constitutive* of a thinker's possession of that concept. Note that there is no need to indicate why these practices individuate the concept number and not some other concept. In the context of the Argument II, the fact that we possess and deploy a sortal concept number is assumed. All we have to do is specify what belief forming practices distinguish between this concept and others. Furthermore, we do not have to delimit *sufficient* conditions for a given (PC) to be a (PCN). There may be different ways of formulating the condition, but if two ways both satisfy the necessary conditions, then they are best regarded as superficially different ways of saying the same thing.

So what might reasonably be considered to be the set of conditions satisfaction of each of which is necessary by candidate (PC)s for (PCN)? One constant of this study has been that the concept number is a sortal concept. Hence, there is a criterion of identity associated with it. This fact must be reflected in the belief forming practices in the (PC) for the concept number. Indeed, this is a similar condition to the one that that lead Frege to introduce the Caesar Problem in *Grundlagen*. He was trying to fix the sense of number terms and, derivatively, the sense of "number" and he was not to be satisfied with a definition unless it could be shown that it established our grasp of a criterion of identity for numbers. In giving the (PC) for number, however, we do not need to solve the Caesar Problem. This is because we do not need to *define* the concept number (as Frege sought to) but just to explain what it is to grasp or possess that concept. Part of the grasp of the concept number is grasping a criterion of identity for numbers. Whence, in specifying the (PC) for the concept number, we are free to take advantage of our intuitive responses to mixed identity statements (which, of course, Frege admits we have) without explaining on what basis these responses are made.

The second necessary condition that a putative (PCN) must satisfy is founded on the apparent equivalence of adjectival and substantival attributions of number. For example, the (PC) must reflect the fact that a belief in the content of "There are three cows" is much the same as a belief in the content of "The number of the cows is three". Frege argued that the substantival form was more basic than the adjectival, that the latter was to be explained on the basis of the former. However, I do not see any good reason to accept this claim in the present context, given that we are not trying to define the concept number but to describe what it is to possess it. There seems no obvious reason why the two forms, and their equivalence at the level of belief, couldn't just reflect some property of our grasp of the concept number. In what follows, I assume that this is the case.

The third and final necessary condition upon a putative (PCN) must reflect the belief that if anything is a number, it must be an abstract entity. It may be objected to this condition that it is not at all obvious that numbers must be thought of as abstract and that, since it is not, there is no need for that belief to be reflected in (PCN). This intuitive response might be strengthened by the observation that the distinction between abstract and concrete entities, if it exists, is notoriously difficult to draw.[53] Both points are well made. However, in the present context I believe they are somewhat misguided. Both the Fregean, who proposes the Argument II, and the Conservative Nominalist, who rejects abstract entities, take numbers to be abstract.[54] What we have to do is to outline a (PC) for the concept number which both of them would accept and see what can be made of it with respect to the apriority of (HP) under (MA). Whether or not the (PC) is appropriate in a wider debate need not concern us.

My suggestions for a possession condition and associated determination theory for the concept number are these:

[53] Though see Hale (1987) for an interesting attempt.

[54] The Conservative Nominalist denies the existence of numbers precisely because they are abstract and she believes that this leads to insurmountable difficulties in their epistemology. See Section 1.1.

(PCN) The concept number is that concept C to possess which a thinker must conform to the following belief forming practices whenever she believes that F and G are Concepts;

(1) If she believes that the C of F's equals the C of G's, then she must be disposed, if the need arises, to form the belief that F and G are equinumerous.
(2) If she believes that F and G are equinumerous, then she must be disposed, if the need arises, to form the belief that the C of F's equals the C of G's.
(3) If she believes that the C of F's is n, and if she believes that n is a numeral, then she must be disposed, if the need arises, to form the belief that there are n F's.
(4) If she believes that there are n F's, and she believes that n is a numeral, then she must be disposed, if the need arises, to form the belief that the C of F's is n.
(5) She is disposed to believe that the C of F's is not equal to q, whenever she believes that q refers to an object falling under some concrete sortal.
(6) If she believes that q is a C, for any singular term q, she must be disposed, if the need arises to form the belief that q is an abstract entity.

(DTN) The semantic value (referent) of number is that one-to-one function which maps distinct equivalence classes of Concepts onto distinct objects.

A word of explanation may be appropriate as to my specification of (PCN). It should be clear that conditions (1), (2) and (5) jointly guarantee that our thinker grasps a criterion of identity for numbers and so grasps a sortal concept number. (1) and (2) are to establish that possession of the concept number is possession of a concept whose referent is a sortal. (5) is to distinguish it from other concepts which have a sortal referent (i.e. to answer the Caesar Problem).[55] (3) and (4) reflect the fact that a thinker who grasps the concept number will be sensitive to the equivalence at the level of belief of adjectival and substantival attributions of number.

[55] It may be objected that (5) does not distinguish the sortal expresses by "C" from other abstract sortals. That is quite true. However, I do not think it needs to do so since that would precludes the identification of numbers with other abstract objects. The Argument II is supposed to show that numbers exist. It need not tell us everything there might be to know about their character.

The most obvious point of objection with respect to my specification of (PCN) is the use made in clauses (3), (4) and (5) of our thinkers abilities to reason with numerals. Why, for example, could we not have put the following in place of (3): If she believes that the C of F's is n, and if she believes that n is a singular term for a number, then she must be disposed, if the need arises, to form the belief that there are n F's?[56]

There are two reasons why this is unacceptable. Firstly, the appearance within our thinkers propositional attitudes of the concept number violates Peacocke's non-circularity constraint for (PC)s (see Section 3.2). Secondly, not all singular terms for numbers give rise to well-formed sentences when manipulated in this way. So the putative modification of (3) is unacceptable.

Nonetheless, there must be something about the thinker's beliefs which makes her disposition to infer "There are n F's" from "The number of F's is n" (and *vice versa*) partly *constitutive* of her grasp of the concept number and not just an idiosyncratic tendency. It is in response to this demand that reasoning with numerals is introduced. The alert objector who saw their last suggestion rebutted is now likely to complain that one cannot explain the notion of a numeral without recourse to the concept number, whence the explicit circularity of her suggestion has been replaced by an implicit circularity in mine. Although I do not agree that this is the case, the objection is well made. Since our thinker must believe, concerning given expressions, that they are or are not numerals, it is incumbent upon us to give a characterisation of numerals understanding of which does not depend upon the concept number. Otherwise, Peacocke's circularity constraint for (PC)s is violated.

To specify criteria for numerality, we need to consider in what an understanding of the concept numeral consists. Intuitively speaking, it is clear that an arbitrary singular term 'q' could not be a numeral unless it were a member of a potentially infinite sequence of singular terms, with some base, which is ordered by the concept is-next-in-the-sequence. Our criteria for numerality must reflect this fact. In other words, it must distinguish between arbitrary sets of numerals and arbitrary sets of

[56] In fact, the appearance of numerals need not be too surprising. It is based around the idea that the possession of certain concepts might be best explained by an account of how one reasons with notations which somehow characterise those concepts. It might be argued that such concepts are sign-dependent, such that one cannot possess them unless one uses the relevant notation. I remain neutral on this matter but see Peacocke (1997) for an interesting discussion.

arbitrary singular terms. I claim that a non-circular explanation of numerals, i.e. one which does not require grasp of the concept number, can be provided by a non-circular explanation of two of their uses. Firstly, numerals are used as singular terms. Our criteria for numerality should reflect that fact. Secondly, they are used to index the existential quantifier in order to formulate an ordered sequence of numerically definite quantifiers. Our criteria should reflect that fact. To avoid circularity, it will be necessary to explain the notion of a sequence of numerically definite quantifiers without recourse to the concept number. Therefore, the criteria for numerality I propose are as follows:

A potentially infinite set of sub-sentential expressions, Γ, is a set of numerals if:

(Numerality 1) For all τ in Γ, τ can be used as a singular term,

(Numerality 2) Γ is ordered by the relation is-next-in-the-sequence,

(Numerality 3) For all τ in Γ, τ can be indexed to the existential quantifier to form a well formed complex quantifier. The proper expansion of such a a quantifier is defined, for open sentence Fx and quantifier $\exists_\tau$, as

(i) $\forall x \neg Fx$, if nothing occurs before τ in the set as ordered sequence,

(ii) $\exists x_\rho \ldots \exists x_\gamma \exists y(x_\rho \ldots x_\gamma, y$ are distinct F's $\wedge \forall z(Fz \rightarrow z$ is x_ρ or...or $y)$, where $\rho \ldots \gamma$ precede τ in the set Γ as ordered sequence.

Given these criteria for when a set of expressions is a set of numerals, being a numeral is explained simply as being a member of such a set. Now, can a thinker believe that some set of terms satisfies these criteria without believing anything about numbers or the concept number? Is it the case that a thinker could understand these criteria without violating Peacocke's circularity constrain for (PC)s? The concept involved in these criteria grasp of which seems most objectionable in this respect is the concept potential infinity. Dummett, for one thinks that our concept actual infinity cannot be separated from our concept number (1998). It may be felt that his claim should be extended to the concept potential infinity. However, I tentatively submit that there is a way which a thinker can grasp this concept without a prior grasp of the concept number. This is because I think it is plausible to maintain that a thinkers formation of beliefs in accordance with various subsets of the clauses in (PCN) is sufficient to ground her possession of certain "proto-concepts" of number.[57] In particular, I think that if a thinker forms beliefs in accordance with clauses (1) and (2), then she possesses a concept, call it '\$', for which can be framed a set of axioms analogous to the Dedekind/Peano axioms for number. Furthermore, since the cardinality operator is sufficient to the analytic reduction of the Dedekind/Peano axioms (see Section 2.1), *these* axioms can be expressed in a second-order logic just by the addition to its lexicon of a non-logical functional operator which expresses \$. In addition, our thinker is also in a position to acquire a belief analogous to (HP), i.e. a belief in the content of "The \$ of F's equals the \$ of G's iff the F's and G's can be put into one-to-one correspondence", merely by reflecting upon how she forms beliefs. Since it is possible to prove Frege's Theorem with (HP) (see Section 2.1), it follows that there exists a proof, analogous to Frege's Theorem, from the sentence "The \$ of F's equals the \$ of G's iff the F's and G's can be put into one-to-one correspondence" of the Dedekind/Peano axioms for \$. That she can construct this proof is sufficient ground for the claim that she possesses (at least) a concept of potential infinity. Hence, it seems plausible to suggest that there is no obstacle to her understanding the criteria for numerality without having a prior grasp of the concept number. Therefore, (PCN), as stated above, satisfies Peacocke's non-circularity constraint. Since it also satisfies the constraints placed upon a (PC) in order for it to be a candidate for (PCN), I suggest that it should be adopted.

Returning to the question of whether (HP) is apriori under (MA), it is necessary now to give a (PC) and (DT) for the concept equinumerous. My suggestions for the (PC) and (DT) for equinumerous, (PCE) and (DTE), respectively, are as follows:

(PCE) The concept equinumerous is that concept C to possess which a thinker must conform to the following belief forming practices if she believes that F and G are Concepts:

(1) If she believes that C(F,G) then she must be disposed, if the need arises, to form the belief that $\exists R(F\ 1\text{-}1_R\ G)$.

(2) If she believes that $\exists R(F\ 1\text{-}1_R\ G)$, then she must be disposed, if the need arises,

[57] By this I do not mean notions which approximate to the concept number but which do not qualify

to form the belief that C(F,G).[58]

(DTE) The semantic value (referent) of the concept equinumerous is that equivalence relation on Concepts which relates Concepts whose extensions have the same cardinality.

As with the (DT)'s for the logical concepts, (DTN) and (DTE) are assumed to be both true and known. Now, as Peacocke states, the truth of a (DT) for a given (PC) ensures that the inferential or belief forming practices in it are truth-preserving. So, given that (PC)s and (DT)s for the logical concepts have been assumed, (PCN) + (DTN) and (PCE) + (DTE) combine to ensure that the following are true:

(i) $\forall F \forall G$("Nx:Fx=Nx:Gx" is true iff "F and G are equinumerous" is true),

(ii) $\forall F \forall G$("F and G are equinumerous" is true iff "$\exists R(F\ 1\text{-}1_R\ G)$" is true).

Note that these sentences are consequences of (PCN) + (DTN) and (PCE) + (DTE), only if (PC$\forall$) and (DT$\forall$) license their derivation. Now, given a sound formalisation of arithmetic, A, and a satisfactory theory of truth, T, for that language, it is possible to construct the required derivations. By transitivity of the biconditional, we have, from the two truths just mentioned, that:

(iii) $\forall F \forall G$("Nx:Fx=Nx:Gx" is true iff "$\exists R(F1\text{-}1G)$" is true).

(Again this inference is subject to validation by (PC$\forall$) and (DT$\forall$)). Since T is satisfactory we may assume that it satisfies Tarski's Condition-T, i.e. it derives all the T-sentences for A (Tarski, 1944), therefore we have that:

(2) "Nx:Fx=Nx:Gx" is true iff Nx:Fx=Nx:Gx,

(3) "$\exists R(F1\text{-}1G)$" is true iff $\exists R(F\ 1\text{-}1_R\ G)$,

(4) "$\forall F \forall G$ Nx:Fx=Nx:Gx $\leftrightarrow \exists R(F\ 1\text{-}1_R\ G)$" is true iff

---

for concepthood but rather notions which *are* concepts which somehow underpin our concept number.

[58] It may be objected that no-one is going to find these inferences obvious. However, they do not need to be obvious without any reflection, see n54.

$\forall F \forall G(Nx{:}Fx=Nx{:}Gx \leftrightarrow \exists R(F\ 1\text{-}1_R\ G)$.

Adding (1), which is justified by the (PC)'s and (DT)'s relevant to (HP), to the theory T as an additional axiom makes possible this derivation:

(D1) $\forall F \forall G$("$Nx{:}Fx=Nx{:}Gx$" is true iff "$\exists R(F\ 1\text{-}1_R\ G)$" is true),	by (1),
(D2) "$Nx{:}Fx=Nx{:}Gx$" is true iff "$\exists R(F\ 1\text{-}1_R\ G)$" is true,	by (PC$\forall$) and (DT$\forall$),
(D3) $Nx{:}Fx=Nx{:}Gx$ iff "$\exists R(F\ 1\text{-}1_R\ G)$" is true,	MP on (D1) and (2),
(D4) $Nx{:}Fx=Nx{:}Gx$ iff $\exists R(F\ 1\text{-}1_R\ G)$,	MP on (D2) and (3),
(D5) $\forall F \forall G(Nx{:}Fx=Nx{:}Gx$ iff $\exists R(F\ 1\text{-}1_R\ G))$	by (PC$\forall$) and (DT$\forall$),
(D6) "$\forall F \forall G(Nx{:}Fx=Nx{:}Gx$ iff $\exists R(F\ 1\text{-}1_R\ G))$" is true,	MP on (D5) and (4).

This is a proof that demonstrates the apriority of (HP) under the metasemantic account.[59]

I admit that I have done no more than sketch a proof that (HP) is Peacocke-apriori (and hence apriori *tout court*, if (MA) is acceptable). However, it looks to me as if something like this could be made to work. Certainly, it seems that if anything can legitimately be called apriori under (MA), then there is a good case to be made for saying that (HP) can. However, the negative phase remains. Boolos' objections against asserting that (HP) is analytic (hence, apriori) remain to be considered.

[59] I say "a proof" not "the proof" for two reasons. Firstly, it has been left open that there may be other (PCN)s and (DTN)s that satisfy the conditions mentioned. These would allow the construction of other proofs of (HP). Secondly, Peacocke allows that different proofs of the same content might be forthcoming from the same (PC)s and (DT)s, see Peacocke (1987).

3.4 Could (HP) *really* be Apriori?

Boolos presents an impressive battery of objections against taking (HP) to be an analytic truth or a truth of logic (Boolos, 1990, 1997). I explained in Section 3.1 why I think that the Fregean is most likely to substantiate her claim that (HP) is apriori by appeal to a theory of apriority which is based upon analyticity. We saw in section 3.2 that Peacocke's account is just such a theory. For the Fregean, therefore, Boolos' objections to the analyticity of (HP) are objections to its apriority also. The objections are made from three perspectives, worries concerning the role of (HP) in mathematics, worries concerning the ontological commitments of (HP) and worries concerning its consistency. I shall only deal with the objections from the second two perspectives. This is because they strike me as the strongest objections Boolos suggests.[60]

There are two objections from ontological commitment. Firstly, Boolos points out that (HP) can only be true in interpretations with infinite domains (1997, p250). To see why this is so, observe that Frege's Theorem can only be proved if the functional expression "Nx:Fx" is construed as a total function from equivalence classes of Concepts (under equinumerousity) to objects. If (HP) is true, it follows that the set of Concepts over which its second-order variables range is partitioned into equivalence classes by the relation of equinumerousity. Suppose there is only a finite number, n, of elements in the domain of some interpretation of second-order logic extended by (HP). It follows that there is an equivalence class containing all Concepts under which no objects fall, an equivalence class containing all Concepts under which one object falls,.........and an equivalence class containing all Concepts under which n objects fall: n+1 equivalence classes in sum. Hence, since "Nx:Fx" is a total function, (HP) must be false in this interpretation since there are not enough objects in the domain for the function to be well-defined. There are not enough elements of the domain to go

[60] The objections from mathematics begin with the claim that second-order arithmetic is an extremely powerful theory. Consequently, the fact that it can be interpreted in second-order logic extended by the addition of (HP) as a further axiom strikes Boolos as a good reason to doubt that sentences apriority. Furthermore, he *proves* that arithmetic can be interpreted in second-order logic extended by the statements that nothing precedes zero and that precedes is a one-one relation (Boolos, 1996). (HP) itself is shown not to follow from the conjunction of these two statements, even though they are amongst its consequences. It is therefore stronger than some of its consequences. In the face of these facts, Boolos wonders why we might even want to claim that (HP) is apriori. My response is just that what is and what is not mathematically possible is strictly irrelevant to the question of whether (HP) is apriori. That should be decided by appeal to a philosophically sound and well-motivated theory of the apriori.

round, so to speak. Therefore, (HP) is true only in interpretations with infinite domains. Now, Boolos claims that analytic sentences, if they exist at all, have two very distinctive properties:

"first, they are true; secondly and roughly speaking, they lack content: that is, they make no significant or substantive claims or commitments about the way the world is; in particular, they do not entail the existence of particular objects or of more than one object ." (1997, p248)

As a result, the fact that (HP) is true only in interpretations with infinite domains is an objection to holding it analytic. It entails the existence of infinitely many objects and this is something that no truly analytic sentence could do.

Secondly, (HP) entails the existence of a total function mapping concepts to objects, the function expressed by "Nx:Fx". Since (HP) can only be true if such a function exists and since it can only be analytic if it is true, it follows that the claim that (HP) is analytic entails the claim that such a function exists. But what guarantee do we have for its existence? Why should the sentence "There exists a third-level total function from concepts to objects." be regarded as apriori? Boolos claims, of course, that we have no such guarantee and that this is sufficient reason to assert that (HP) is not analytic (1997, p251).

There are also two objections from consistency. The discovery that second-order arithmetic is inconsistent is, for Boolos, a live possibility:

"Do we really know that some hotshot Russell of the 23rd Century won't do for us what Russell did for Frege?" (1998, p259)

he asks. To see his point, note that extending second-order logic by the addition of (HP) as a further axiom yields a theory in which second-order arithmetic can be interpreted. Therefore the question of its consistency is directly related to the truth of (HP): if it is inconsistent, (HP) is false. But, according to the two distinctive properties of analytic statements which Boolos alludes to (see quotation above), false statements cannot be analytic.

The second objection from consistency is the final objection I shall consider. It has become known as the "Bad Company Argument". Since it is more subtle than the previous objections, I shall first explain it in outline before fleshing it out with details. Recall from Section 1.2 (n12) that (HP) is an example of what I called a "B-Sentence", i.e. a second-order sentence of the form $\ulcorner$#F=# G $\leftrightarrow$ E(F, G)$\urcorner$ where '#' denotes a third-level function from Concepts to objects and 'E(F, G)' is an equivalence relation defined on Concepts. We know that some B-Sentences cannot be analytic. Frege's Axiom V is a B-Sentence and it leads to Russell's paradox. The principle that the "relation number" of one relation is equal to the "relation number" of another relation if and only if the two relations are isomorphic is also a B-Sentence.[61] It leads to the Burali-Forti paradox when adjoined to second-order logic.[62] Hence, these Sentences are false and so cannot be analytic. Therefore, we need some method of distinguishing between those B-Sentences which are analytic and those which are not. Now, it can be proved that (HP) is equiconsistent with analysis.[63] However, Boolos argues that it will not do to take the relative consistency of (HP) as enough to guarantee its analytic status. Firstly, there is the worry identified above that we do not really *know* that second-order arithmetic is consistent, we only think that it is. Secondly, and this is the foundation of the Bad Company Argument, it is possible to construct B-Sentences which, although they yield consistent extensions of second-order logic, *are not consistent with (HP)*. It follows that not all (individually) consistent B-Sentences can be analytic truths. So some criterion is required to distinguish between the (individually) consistent B-Sentences which are analytic from those that are not. Boolos is sceptical that such a criterion can be provided. He concludes that, if it cannot, no B-Sentences should be thought analytic.

So much for the outline, what about the detail? These details consist in showing that a B-Sentence exists which is consistent with second-order logic, but is inconsistent with (HP).[64] To complete the first task, let Concepts F and G differ evenly if and only if the number of objects that fall under F but not G, or that fall under G but not F, is even. Define 'de(F,G)' as that relation upon concepts which relates two concepts if and only if they differ evenly. Then 'de(F,G)' is an equivalence

[61] Observe that isomorphism is an equivalence relation on relations and that "The relation number of..." is a functional operator mapping predicates to singular terms to see that the principle satisfies the characterisation of a B-Sentence given in Section 1.3.
[62] This result is proved in Hodes (1994). Note that not everyone blames the inconsistency of second-order logic extended by Axiom V etc. solely on the inconsistency of those Sentences. Dummett (1991) argues that the real problem is impredicativity combined with a objectual interpretation of second-order quantification. Boolos responds to this claim in Boolos & Clark (1993), Dummett replies in his (1994).
[63] This is independently noted by Burgess (1984), Hazen (1985) and Hodes (1984), proved by Boolos (1990).

relation on Concepts. From this it follows that there is a B-Sentence which has "$\forall F \forall G de(F,G)$" as its RHS. Let the functional operator on its LHS be "Px:Fx", this should be read "The Parity of F". Then the B-Sentence in question is the "Parity Principle":

(PP) $\forall F \forall G((Px{:}Fx=Px{:}Gx) \leftrightarrow \forall F \forall G de(F,G))$

This reads "The Parity of F equals the Parity of G if and only if F and G differ evenly.". Now, to show that (PP) is consistent (with standard second-order logic) it is sufficient to show that the extension of second-order logic that results from the addition of (PP) as sole non-logical axiom has a model. So let X be any finite set which contains 0 and 1. For all subsets, Y, of X let Px:Yx equal 0 if the cardinality of Y is even and let Px:Yx be equal to 1 otherwise (i.e. if the cardinality of Y is either finite and odd or infinite). Under these stipulations (PP) is true. Hence it has a model and so is consistent.

Now it has to be shown that (PP) and (HP) are inconsistent. Recall that (HP) is only true in interpretations that have *infinite* domains. The way to show that (PP) is inconsistent with (HP) is to show that it is only true in interpretations with *finite* domains. Let X be an infinite set and let Y and Z be subsets of X. Then the following (in)equations hold:

1. $|\{Y| de(Y,Z)\}|$
2. $= |\{<A,B>| A, B \subseteq X, A \subseteq Y, B \subseteq Z, A \cap Z = \varphi, B \cap Y = \varphi,$
   $|A \cup B|$ is even and $Y=(ZA)B\}|$
3. $\leq |\{<A,B>| A, B \subseteq X$ and $|A \cup B|$ is even$\}|$
4. $= |X|$.

Now let $f{:}P(X) \rightarrow X$ be a function such that for all Y and Z, subsets of X, $f(Y)=f(Z)$ only if de(Y,Z). Note that this presupposes that there exists such a function, which

[64] What follows in this and the next paragraph is taken from Boolos (1990, p250-52). I have changed the terminology somewhat and added a little explanation where I felt that those unfamiliar with certain results from mathematical logic might get lost.

entails that (PP) is true. Since $|\{Y de(Y,Z)\}| = |\{Y f(Y)=x\}|$ for any x a member of X, the assumptions entail that:

5. $\forall x \in X |\{Y f(Y)=x\}| \leq |X|$
6. $\Rightarrow |P(X)| \leq |X| \times |X|$, since f is an equivalence relation on P(X) and such a relation partitions its domain into equivalence classes (of which there are at most as many as the cardinality of the domain),
7. $\Rightarrow |P(X)| = |X|$, since X is infinite,
8. Contradiction, since the cardinality of the power set of Ψ is greater than that of Ψ, for any set Ψ.

Therefore, (PP) cannot be true in interpretations which have infinite domains. Thus, (PP) and (HP) are mutually inconsistent. But now we must ask the question: Which one is an analytic truth? Boolos' conclusion is that until we have a principled way to decide such questions, we must deny the analyticity of both. More generally, until we have a principled way to decide the question given any individually consistent but mutually inconsistent B-Sentences, we must withhold analyticity from them all.

These are Boolos' objections. Before considering how the Fregean is to respond, it is worth discussing them a little further. Note, first, that none of them draw upon the foundation of any particular theory of analyticity. In fact, the only theoretical claims that Boolos commits himself to are the assertions that analytic truths cannot carry existential commitments and that they must be true though neither claim is argued for. This is an important point for, in effect, the fact that Boolos' finds the ontological commitments of (HP) to be objectionable is not to be explained in virtue of any facts that pertain, in particular, to (HP). These are not objections to the analyticity of (HP) *qua* (HP). It is only the objections from consistency that satisfy that description.

Observe, also, that there is a significant relationship between the existential commitments entailed by (HP). The fact that (HP) entails the existence of infinitely many objects and the fact that it entails the existence of a total function mapping distinct equivalence classes onto distinct objects may well be thought of as two sides of the same coin, at least on one plausible explanation of when it is correct to assert the existence of functions. According to this plausible explanation, the existence of functions is parasitic on the existence both of the entities which are supposed to be arguments to the function and of the entities that are supposed to be values of the function. In other words, functions only exist when domains and ranges exist over which the functions can be (well-) defined. Thus, on this view of the metaphysics of functions, that (HP) entails the existence of infinitely many objects and that it entails the existence of the aforementioned function comes to much the same thing. This is not to say that the objections from content are *exactly* the same, just that there is more in common between them than the fact that they are motivated by the same doubt, that analytic truths could not entail ontological commitments. At the least, it would seem to follow that our knowledge of the two commitments is inextricably linked.[65]

So how is the Fregean to respond to Boolos' attack? The first, and most obvious, response for the Fregean is to reject Boolos' characterisation of analyticity. Indeed, it looks as if accepting that characterisation is tantamount to throwing in the towel. Assume, for the moment, that Boolos' characterisation of analyticity is correct. The Fregean must show that that his objections are misplaced. She must show, in particular, that the existence claims which are supposed to follow from the truth of (HP) do not, in fact, follow from the truth of (HP). This is precisely Wright's approach, at least in answer to the assertion that (HP) entails the existence of infinitely many objects (Wright, 1997). In his opinion, this fact is to be explained, not solely on the basis of the truth of (HP), but also on the basis of the truth of the right hand sides of a certain (infinite) class of its instances. So, for example, the sentence "The number of dogs is the number of cats" will be true, if it is, because there exists a bijective map from the Concept *dog* to the Concept *cat*, not just because "The number of dogs is the number of cats if and only if there exists a bijective map from the Concept *dog* to the Concept *cat*" is true. The class of instances Wright is particularly interested in are those the right hand sides of which can be known apriori. Start with the Concept $\neg(x=x)$. This Concept is bijectively mapped to itself by the identity map, hence its number equals its number, so to speak. Let this number be defined as "0". Next consider the Concept $x=0$. Again, the identity map bijectively maps this Concept to

[65] Even if the metaphysical picture is not accepted, it is clear that the proof of the claim that (HP) is true only in infinite domains could not be given without the supposition that the function mentioned exists.

itself, so its number again equals its number. Let this number be defined as "1". If n numbers have been defined, then the $(n + 1)^{th}$ will be given by considering the action of the identity map on the Concept $x=0 \ldots \vee x=n$. Hence, the natural numbers are recursively defined and their existence is shown to follow from the truth of the right hand sides of certain instances of (HP) together with the truth of (HP). Thus, one cannot object to the analyticity of (HP) on the grounds that we know on its basis alone that infinitely many objects exist. This is an attractive argument. Unfortunately, in the present context, it is inconsistent.

The problem for the Fregean is that, under the assumption that Boolos' characterisation of the analytic is correct, and under the assumption that this characterisation extends to apriori truths (which it does in the context of the Argument II as I have presented it), one cannot claim to know apriori the right hand sides of those instances of (HP) described above. These are all of the form "$\exists R(F\ 1\text{-}1_R\ F)$" so to know them apriori, we would have to know apriori of the existence of a particular entity. Now, under Boolos' characterisation of the analytic this is just the sort of thing that cannot be known analytic. So, if what is apriori just is what is analytic, as it is for the Fregean in the present context, then these sentences cannot be apriori. But the argument given above requires that they are. To put the point as simply as it may be, no consistent argument can result from accepting that one cannot know apriori of the existence of infinitely many or particular objects and then claiming that we know, a priori, of the existence of particular objects.

Of course, I do not mean to attribute the inconsistent argument to Wright. My discussion of it is supposed to show only that in the present context the Fregean cannot both accept Boolos' characterisation of the analytic and give Wright's response to the ontological commitment objections. Nonetheless, Wright's response to these objections is somewhat enigmatic. He cannot accept Boolos' theoretical claim concerning analyticity and then give the argument outlined above. He does give the argument outlined above. Hence, one presumes that he rejects Boolos' characterisation of analyticity. But if this is the case, then there is no particularly good reason for him to identify some source other than the truth of (HP) as the origin of the ontological commitments.

I now turn to those objections of Boolos' which play upon distinctive features of (HP). These, really, are the interesting objections, since the others turn upon what can be taken as an acceptable conception of analyticity. Pending a final answer to that question, we would like to know if there are any convincing objections to the apriority of (HP) over and above the purely general ones that are based around Boolos' conception of analyticity. If there are not, we may conclude that there is no special reason for denying the apriority of (HP), whatever we actually think about the analytic.

The first objection from consistency is simply overcome. The suggested objection is that because we do not know that (HP) is not inconsistent, we do not know that it is true. At least this is the position in Boolos (1997). The position in Boolos (1990) is slightly different. There he notes that (HP) is equiconsistent with analysis and concludes that it would be unreasonable (irrational?) to reject (HP) on the basis that it may be inconsistent (this for the reason that it would be unreasonable to reject analysis on the same basis). Now, it seems fairer to take Boolos at his latest word, though it would have been nice to know why he changed his mind. However, there seems nothing stopping us offering the retort suggested by his earlier self. This is the response given by Wright (1997) and I think it is perfectly acceptable.

The Bad Company argument is not as quickly dealt with. The particular claim (which is subsequently generalised) is that there are certain B-Sentences which are, individually, as consistent as (HP) but which are mutually inconsistent with (HP). Since they can lay as much claim to being analytic as (HP), Boolos states that, unless a principled choice can be made between them, we should hold that neither they nor (HP) are analytic.[66]

To begin with observe that we cannot say that (HP) is analytic, as opposed to (PP), say, because we possess and deploy the concept number but not the concept parity. This response would entail that the only relevant difference between the concepts is that we have evolved to use one, not the other, and that the one that should be thought apriori is the one we have developed to use. But that cannot be correct, what contents are analytic should not be dependent upon what concepts we possess and deploy, as a matter of contingent fact.

[66] Of course, this principle should also tell us why the demonstrably inconsistent B-Sentences are not in the market for analyticity.

However, Wright (1997) claims that there is a relevant difference between (PP) and (HP). The difference is supposed to show that (HP) is an acceptable B-Sentence for the Fregean Argument and to show that (PP) is not. Generalising from this case, Wright concludes that there is a principle by which we can partition the class of B-Sentences into those that are, and those that are not, apt for adoption by the Fregean. Of course, by this he means that there is, in his opinion, a principle of individuation by which we can pick out those B-Sentences which are capable of introducing, or explaining, concepts, i.e. a principle which will pick out those B-Sentences apt to found an argument like the Argument I. I have rejected the Fregean Argument on that reading, so it will be important to determine whether Wright's response can be modified to suit the Argument II before considering whether, if such a modification is possible, it yields a cogent response.

Wright asserts that because (PP) is true only in finite domains, it has implications for the cardinality of extensions of Concepts other than the Concept *parity*. This includes concrete sortals like fish, house and person. Hence, if (PP) is analytic, it follows that we know apriori that there are only a finite number of fishes, houses and people, under the assumption, that is, that logical consequences of apriori contents are, themselves, apriori. In contrast, (HP) only has implications for the cardinality of the extension of the concept number; this must be infinite. It does not entail any statements about entities other than numbers. Wright takes this to indicate that (HP), and not (PP), is a B-Sentence capable of introducing a concept and that it follows that (HP), and not (PP), can be thought of as analytic. More generally, he asserts that only those B-Sentences which do not entail any consequences for the cardinality of concepts other than those with which they are specifically concerned can be thought of as legitimate means of concept-formation. Only those B-Sentences can be thought of as analytic. As Wright goes on to state:

"What is at stake in this disanalogy is, in effect, *conservativeness* in (something close to) the sense of that notion deployed in Hartry Field's exposition of his nominalism." (1997, p231, Wright's italics)

But what exactly is the notion of conservativeness to which Wright appeals?[67] Let A be any B-Sentence, with functional operator '#'. Introduce 'Px', a predicate true of

[67] In what follows, I paraphrase Wright's own explanation (1997, n49), this is to retain consonance with the notation used throughout this study.

only the referents of the #-terms. We now define the #-restriction of any sentence, S, to be S with its first-level quantifiers restricted to only those objects of which 'Px' is not true (so the first-level variables in the #-restriction of S range over all objects other than possible referents of #-terms). Given a theory, T, with which A is consistent, we will say that A is conservative with respect to T if and only if for any S expressible in the language of T, T + A entails the #-restriction of S only if T entails S. As Wright points out, one difference between his notion of conservativeness and Field's is that for him a B-Sentence cannot be conservative over a theory with which it is inconsistent, whereas, for Field, a mathematical theory can be conservative over a nominalistic theory with which it is inconsistent. Another difference is that Field's notion is defined in terms of nominallistically acceptable consequences of theories whereas Wright's is concerned with consequences that do not involve quantification over possible #-term referents.

Now, can this response be adopted by the Fregean who is defending the Argument II against the Bad Company argument? To begin with, observe that, since the Caesar Problem is not a problem in the context of the Argument II, the introduction of 'Px' should not be thought contentious. In the context of the Argument I, this would not have been so. Now, Wright's response is convincing (if it is) because he is concerned to distinguish between those B-Sentences that can introduce concepts and those which can't. Hence, his notion of conservativeness, which basically says that a B-Sentence is conservative if it does not entail any consequences about objects that do not fall under the concept it supposedly introduces, might be expected to do the job. The principle at work here is that an acceptable way of determining one concept should not prejudice questions about other concepts, in particular whether their referents are finitely of infinitely instantiated. Maybe this is true. What we must decide is whether there is some other principle which validates the conservativeness response for the defender of the Argument II and, if there is, whether it is a plausible principle to adopt.

To that effect, observe that what really gives Wright's response its force is that, included in the questions which are prejudiced by the acceptance of (PP) and other similar principles, are questions about the cardinality of the concrete world. It is these questions which we presume are not decidable by analytic sentences. So the adoption of conservativeness as the discriminatory principle with which to distinguish between analytic and non-analytic B-Sentences is justified by the principle that it is not possible to know existence claims about concreta apriori. This principle seems to me to be perfectly acceptable. Furthermore, it may be the case that classical mathematics is conservative. Field thinks it is and if it is conservative in Field's sense, it is conservative in Wright's sense (Field, 1980). Given that each extension of second-order logic by the adjunction of a B-Sentence as extra-logical axiom is a mathematical theory, it may be true that only those extended by conservative B-Sentences are acceptable. So maybe only those B-Sentences are in the market for analyticity. If it be objected that this argument does not explain why conservativeness should be taken as *evidence* for the analyticity of a B-Sentence, then the objector is directed to the principle that it is not possible to know "analytically" about the concrete world.

There are three problems with this response to the Bad Company argument. The first is that it looks too *ad hoc* to be acceptable. The conservativeness constraint effectively picks out as acceptable all and only the consistent B-Sentences which are mutually consistent with (HP) (and each other). In answer to this the Fregean can of course claim that mathematics *is* conservative and so whether adopting the conservativeness constraint looks *ad hoc* or not doesn't really matter. I have some sympathy with this response though it must be said that whether mathematics is conservative or not, and, if not, in what sense not, is still an open debate.[68] In addition though, it can perhaps be argued that the response is not as *ad hoc* as it looks. The objector who, like Boolos, doesn't think that any ontological claims can be analytic must surely accept that one cannot know apriori, on the basis of certain analytic sentences, about the cardinality of concrete sortals. But this is the principle that founds the conservativeness response, so why should they reject that?

This brings us to the second problem. The "evidence" mentioned above is not evidence *for* the analyticity of some such sentences, it is evidence *against* the analyticity of certain others. It may be sufficient to show that (PP) and its associates are not analytic to observe that they are not conservative with respect to concrete sortals. But this does not show that (HP) and other B-Sentences which are conservative with respect to the concrete world *are* analytic. That only follows if it is possible to know, on the basis of analytic sentences, about the abstract realm, e.g. if it

[68] See Shapiro (1983) and Field (1985) for discussion of this issue.

is possible to know that there are infinitely many numbers on the basis of (HP). But that is precisely one of the things the Argument II is supposed to show! The Fregean, no doubt, will exclaim that these ontological commitments do not flow solely from the truth of (HP) but from (HP) and the truth of instances of its RHS (as was argued above). But now she is in trouble. If these ontological commitments do not flow from the truth of (HP) alone, why should we believe that the ontological commitments associated with, say, (PP) follow from its truth alone. Clearly we should not. But if we have no need to believe that, we have no need to believe that (PP) *alone* is not conservative with respect to the concrete world. So conservativeness could not be the principle with which to discriminate between apriori and aposteriori B-Sentences.

The Fregean is now in a dilemma neither horn of which she can grasp. Either she concedes that (HP) entails the ontological commitments Boolos says it does or she does not. In the first case, her appeal to conservativeness which is supposed to defeat the Bad Company Argument seems to rely on the claim that we can know apriori of the existence of abstract objects. But the Conservative Nominalist, who doesn't believe that we can know about abstracta in *any* sense, is hardly going to accept that. Even if this were not the case, some new response to the ontological commitment objections would have to be mooted, a different one from that suggested above. Either way, the Argument II does not succeed. In the second case, the objections from ontological commitment may be defeated, but some other response to the Bad Company Argument is called for, i.e. the Fregean is left without a principle by which to individuate the good and the bad B-Sentences. But without such a principle, the Bad Company Argument is surely effective. So in this case, too, the Argument II fails.

3.5 Conclusion

So is (HP) knowable apriori? I have no definite answer to this question. We have discovered (Section 3.3) that (HP) is (probably) apriori under (MA). So if (MA) is accepted then we have good reason to accept the apriority of (HP) and the conclusion of the Argument II. However, we have also found that the Fregean response to Boolos' objections to the analyticity (hence, apriority) of (HP) is insufficient (Section 3.4). However, this was under the assumption (at least with respect to the objections from ontological commitment) that Boolos' characterisation of the analytic was correct. So, if it can be shown that (MA) is to be preferred to Boolos' characterisation of the analytic, and if it can, at the same time, be made plausible to the Conservative Nominalist, then the Argument II will succeed.

To see why, recall that (DT)s are apriori truths under (MA) and that apriori truths are those contents that can be known apriori on Peacocke's account. Recall, also, that (DT)s commit one to the existence of the semantic values of the concepts they deal with. So Peacocke's account of the possibility of apriori knowledge is committed to the possibility of knowing some ontologically loaded contents apriori. In particular, note that apriori knowledge of (DTN) entails apriori knowledge of the existence of precisely the function which Boolos doubts we could know of analytically. I have urged that our knowledge of the existence of this function is not separable from our knowledge of the existence of numbers (Section 3.4). Therefore, a comparative study of Peacocke's account of the apriori and Boolos' (minimal) account of the analytic would defeat the ontological commitment objections if, as a result of that study, the former was adopted in favour of the latter. Furthermore, this would leave open the possibility of invoking the conservativeness response to the Bad Company Argument, since that would no longer have to provide evidence *for* the analyticity of (HP) (that work being done by (MA)) but only evidence *against* the analyticity of (PP) (and its friends). This task it can perhaps fulfill. So if (MA) is to be preferred, by the Conservative Nominalist, to Boolos characterisation of the analytic, the dilemma identified in the previous section would be finessed and the problem with the conservativeness response overcome. (HP) would be apriori and the Argument II would succeed. It is beyond the remit of this study to determine the outcome of this debate.

So where does all this leave Fregean Platonism? Recall that Fregean Platonism was supposed to convince both Radical and Conservative Nominalist's that their nominalism should be recanted, at least with respect to the natural numbers, that it was supposed to show that the natural numbers exist and that it was to found logicism with respect to number-theory. We were interested in the first two of these objectives. To attain these objectives, the Fregean proposed an argument founded upon Wright's interpretation of Frege's Context Principle, which was found to contain two thesis (CPS) and (CPR), and the thesis (RT). I distinguished two readings of this argument. The first reading I called "the Argument I". The reader will recall (Section 2.1) that it was based on the following four claims:

"(i) that the vocabulary of higher-order logic plus the cardinality operator, 'Nx:...x...', provides a sufficient definitional basis for a statement of the basic laws of arithmetic [the Dedekind/Peano axioms];
(ii) that when they are so stated, $N^{=}$ [(HP)] provides for a derivation of those laws within higher-order logic
(iii) that someone who understood a higher-order language to which the cardinality operator was added would learn, on being told that $N^{=}$ is analytic of that operator, all that it is necessary to know in order to construe any of the new statements that would then be formulable.
(iv) Finally and crucially, that $N^{=}$ may be laid down without *significant epistemological obligation*: that it may simply be stipulated as an explanation of the meaning of statements of numerical identity, and that - beyond the issue of the satisfaction of the truth-conditions it thereby lays down for such statements - no competent demand arises for an independent assurance that there *are* objects whose conditions of identity are as it stipulates." (Wright, 1998b, p389)

In effect, the suggestions contained in these four claims are that it is possible to reduce number-theory to second-order logic extended only by (HP), that (HP) can be taken as stipulatively true, as a contextual explanation of the cardinality operator, and that Frege's Theorem can be proved. I argued in Chapter 2 that (HP) cannot contextually explain the concept number, therefore the Argument I failed to convince the Radical Nominalist of the error of her ways. Recall, from Section 1.1, that one of the doubts motivating her nominalism was the belief that we cannot form the concepts necessary to forming beliefs about abstract entities. The Argument I did not convince the Radical Nominalist that this belief is mistaken when the abstract entities involved are the natural numbers.

The second reading of the Fregean's argument that I considered, the Argument II, was a weakening of the Argument I. This is because it did not seek to explain how we acquire the concepts required to understand number-theoretic discourse. Hence, it could not hope to refute Radical Nominalism. Its aim was rather to show that the natural numbers exist and that we can acquire knowledge of them. Thus, the Argument II was aimed at the Conservative Nominalist whose nominalism was motivated by the worry that, even if abstract entities did exist, none of the beliefs we might form about them could acquire the certainty of knowledge (see Section 1.1). The Argument II depended upon taking (HP) as an apriori truth (in effect it asserts claims (i) and (ii) above but replaces claims (iii) and (iv) with the thesis that (HP) is a number-theoretic statement which we can grasp and know, apriori, to be true). As explained above, I have not reached a definite conclusion with respect to the Argument II. Rather, I have concluded that it is no worse off than Peacocke's theory of apriority even in light of Boolos' objections to the apriority of (HP) *qua* (HP) (the objections from consistency). If that can be upheld and made plausible to the Conservative Nominalist, then the Argument II will succeed.

References

Benacerraf, P.
- (1965) 'What numbers could not be', *Philosophical Review*, 74, 47-73. Reprinted in Benacerraf & Putnam (1983), 272-94.
- (1973) 'Mathematical Truth', *Journal of Philosophy*, 70, 661-80. Reprinted in Benacerraf & Putnam (1983), 403-20.

Benacarraf, P. & Putnam, H.
- (1983) *Philosophy of Mathematics, Selected Readings*, ed. Benacerraf & Putnam. Cambridge: Cambridge University Press.

Boghossian, P.
- (1993) 'Does an Inferential Role Semantics Rest Upon a Mistake?', *Mind & Language*, Vol. 8, 1, 27-40.
- (1997) 'Analyticity', Hale & Wright (1997).

Boolos, G.
- (1984) 'To Be is To Be the Value of a Variable (or To Be Some Values of Some Variables', *Journal of Philosophy*, 81, 430-9.
- (1985) 'Nominalist Platonism', *Philosophical Review*, 94, 327-44.
- (1987) 'The Consistency of Frege's *Foundations of Arithemetic*', Thomson (1987), 3-20. Reprinted in Demopoulos (1995), 211-33.
- (1990) 'The Standard of Equality of Numbers', Demopoulos (1995), 234-54.
- (1996) 'On the Proof of Frege's Theorem', Morton & Stich (1996).
- (1997) 'Is Hume's Principle Analytic?', Heck (1997b), 245-61.

Boolos, G. & Clark, P.
- (1993) 'Whence the Contradiction?', *Proceedings of the Aristotelian Society*, suppl. 67, 213-33.

Brown, J.
- (1990) 'π in the Sky', Irvine (1990), 95-120.

Burgess, J.
- (1984) 'Review of Frege's Conception of Numbers as Objects', *Philosophical Review*, 93, 638-40.

Demopoulos, W.
- (1995) *Frege's Philosophy of Mathematics*, ed. Demopoulos. Cambridge, Mass.: Harvard University Press.

Dummett, M.
- (1973) *Frege: Philosophy of Language*. London: Duckworth.
- (1975) 'What is a Theory of Meaning? (I)', Guttenplan (1975). Reprinted in Dummett (1993), 1-33.
- (1976) 'What is a Theory of Meaning? (II)', Evans & McDowell (1976). Reprinted in Dummett (1993), 34-93.
- (1991) *Frege: Philosophy of Mathematics*. London: Duckworth.
- (1993) *The Seas of Language*. Oxford: Clarendon Press.
- (1994) 'Chairman's Address: Basic Law V', *Proceedings of the Aristotelian Society*, 243-251.
- (1998) 'Neo-Fregeans: in Bad Company?', Schirn (1998), 369-87.

Evans, G.
- (1982) *The Varieties of Reference*, ed. McDowell. Oxford: Clarendon Press.

Evans, G. & McDowell, J.
- (1976) *Truth and Meaning*. Oxford: Clarendon Press.

Feigl, H. & Sellars, W.
- (1949) *Readings in Philosophical Analysis*, ed. Feigl & Sellars. New York: Appleton-Century-Crofts.
Field, H.
- (1980) *Science Without Numbers*. Oxford: Blackwell.
- (1984) 'Critical Notice of Wright's *Frege's Conception of Numbers as Objects*' *Canadian Journal of Philosophy*, 14, 637-62. Reprinted as 'Platonism for Cheap? Crispin Wright on Frege's context principle', Field (1989), 147-70.
- (1985) 'On Conservativeness and Incompleteness', *Journal of Philosophy*, 81, 239-60. Reprinted in Field (1989), 125-46.
- (1989) *Realism, Mathematics and Modality*. Oxford: Oxford University Press.
- (1993) 'The Conceptual Contingency of Mathematical Objects', Mind, Vol. 102, 404, 285-99.
Frege, G.
- (1950) *The Foundations of Arithmetic*, trans. Austin. Oxford: Blackwell.
Gödel, K.
- (1947) 'What is Cantor's Continuum Problem?', *American Mathematical Monthly*, 54, 515-25. Reprinted in Benacerraf & Putnan (1983), 447-69.
Guttenplan, S.
- (1975) *Mind and Language*, ed. Guttenplan. Oxford: Clarendon Press.
Hale, B.
- (1984) 'Frege's Platonism', *Philosophical Quarterly*, 34, 225-241.
- (1987) *Abstract Objects*. Oxford: Blackwell.
- (1990) 'Nominalism', Irvine (1990), 121-44.
- (1994) 'Dummett's Critique of Wright's Attempt to Resuscitate Frege', *Philosophia Mathematica* (3), Vol.2, 122-47.
- (1997) 'Grundlagen §64', *Proceedings of the Aristotelian Society*, 243-61.
Hale, B. & Wright, C.
- (1992) 'Nominalism and the Contingency of Abstract Objects', *Journal of Philosophy*, Vol. 89, 3, 111-35.
- (1994) 'A Reductio Ad Surdum? Field on the Contingency of Mathematical Objects', *Mind*, Vol.103, 410, 169-84.
- (1997) *Blackwell Companion to the Philosophy of Language*, ed. Hale & Wright. Oxford: Blackwell.
Hart, W.
- (1996) *The Philosophy of Mathematics*. Oxford: Oxford University Press.
Hazen, A.
- (1985) 'Review of Crispin Wright's Frege's *Conception of Numbers as Objects*', *Australian Journal of Philosophy*, 63, 251-54.
Heck Jnr., R.
- (1992) 'On the Consistency of Second-Order Contextual Definitions', *NOÛS*, 26, 491-94.
- (1997a) 'The Julius Caesar Objection', Heck (1997a), 273-308.
- (1997b) *Language, Thought, and Logic: essays in honour of Michael Dummett*, ed. Heck. Oxford: Oxford University Press.
Hodes, H.
- (1984) 'Logicism and the Ontological Commitments of Arithemtic', *Journal of Philosophy*, 81, 123-49.

Irvine, A.
- (1990) *Physicalism in Mathematics*, ed. Irvine. Dordrecht: Kluwer Academic Publishers.

Lee, O.
- (1936) *Philosophical Essays for A.N. Whitehead*, ed. Lee. New York: Longmans.

Maddy, P.
- (1980) 'Perception and Mathematical Intuition', *Philosophical Review*, 84, 2, 163-96.

McDowell, J.
- (1997) 'Another Plea for Modesty', Heck (1997b), 105-29.

Millar, A.
- (1994) 'Possessing Concepts', *Mind*, Vol. 103, 409, 73-82.

Morton, A. & Stich, S.
- (1996) *Benacerraf and his Critics*, ed. Morton & Stich. Oxford: Blackwell.

Neale, S.
- (1995) 'The Philosophical Significance of Gödel's Slingshot', Mind, Vol.104, 416, 761-825.
- (1997) 'Slingshots and Boomerangs', Mind, Vol.106, 421, 144-68.

Parsons, C.
- (1979) 'Mathematical Intuition', *Proceedings of the Aristotelian Society*, 80, 145-68
- (1997) 'Wright on Abstraction and Set Theory', Heck (1997b), 263-71.

Peacocke, C.
- (1987) 'Understanding Logical Constants: A Realist's Account', *Proceedings of the British Academy*, LXXIII, 153-200.
- (1992a) *A Study of Concepts*. Cambridge, Mass.: MIT Press.
- (1992b) 'How Are A Priori Truths Possible?', *European Journal of Philosophy*, 175-99.
- (1997) 'Concepts Without Words', Heck (1997b), 1-34.

Quine, W.
- (1936) 'Truth by Convention', Lee (1936). Reprinted in Benacerraf & Putnam (1983), 329-54.
- (1951) 'Two Dogmas of Empiricism', *Philosphical Review*, 60, 20-43.
- (1958) 'Speaking of Objects', *Proceedings and Addresses of the American Philosophical Association*. Reprinted in Quine (1969), 1-25
- (1963) 'Carnap and Logical Truth', Schilpp (1963). Reprinted in Benacerraf & Putnam (1983), 355-376.

Rosen, G.
- (1993) 'The Refutation of Nominalism (?)', *Philosophical Topics*, Vol. 21, 2, 149-86.

Russell, B.
- (1919) *Introduction to Mathematical Philosophy*. London: George Allen & Unwin.

Schilpp, P.
- (1963) *The Philosophy of Rudolph Carnap*, ed, Schilpp. La Salle, Ill.: Open Court.

Schirn, M.
- (1998) *Philosophy of Mathematics Today*, ed. Schirn. Oxford: Oxford University Press.

Shapiro, S.
- (1983) 'Conservativeness and Incompleteness', *Journal of Philosophy*, 80, 9, 521-31.

Tappenden, J.,
- (1993) 'The Liar and sorites Paradoxes: Toward a Unified Treatment', *Journal of Philosophy*, 90, 11, 551-577.

Tarski, A.
- (1944) 'The Semantic Conception of Truth', *Philosophy and Phenomenological Research*, Vol. IV. Reprinted in Feigl & Sellars (1949), 52-84.

Thomson, J.
- (1987) *On being and saying: essays in honour of Richard Cartwright*, ed. Thomson Cambridge, Mass.: MIT Press.

Wittgenstein, L.
(1968) *Philosophical Investigations*, trans. Anscombe. Oxford: Blackwell.

Wright, C.
- (1983) *Frege's Conception of Numbers as Objects.* Aberdeen: Aberdeen University Press.
- (1988) 'Why Numbers Can Believably Be', *Revue Internationale de Philosophie*, 42, 425-75.
- (1989) 'The Verification Principle: Another Puncture -Another Patch', *Mind*, Vol.98, 392, 611-22.
- (1990) 'Field and Fregean Platonism', Irvine (1990), 73-93.
- (1997) 'On the Philosophical Significance of Frege's theorem', Heck (1997b), 201-44.
- (1998a) 'On the Harmless Impredicativity of $N^{=}$', Schirn (1998), 339-68.
- (1998b) 'Response to Dummett', Schirn (1998), 389-405.