Chapter 7
Frege and Benacerraf’s Problem

Crispin Wright

It is a particular pleasure to be here on this very pleasant and auspicious occasion, to honour Bill Demopoulos’ contributions to our subject.¹ I tend to think of Bill as something of a Companion in Arms: we disagree about lots of things, but nevertheless . . . (Demopoulos, from the floor: “We agree on what’s important!”) . . . well, we agree on the great importance of Frege’s philosophy of mathematics. Michael Dummett, of all people, in the introduction to Frege: Philosophy of Language (now more than 35 years ago) said of Frege’s philosophy of language how it was of the utmost contemporary relevance and importance and so on, but of his philosophy of mathematics, Dummett said that it was “indisputably archaic”, in a way in which, he claimed, the philosophies of mathematics of Frege’s contemporaries, Brouwer, Hilbert and Dedekind, are not. So one lonely crusade—perhaps less lonely recently—going back to the early nineteen-eighties was to combat an intellectual milieu in which Dummett’s view was typical. I think it fair to say that up until about the new millennium, almost nobody thought that Frege’s philosophy of mathematics was important: they thought it was of historical interest, worth study for its depth, and technical innovations, and because of the intellectual courage of Frege’s project, and its ultimate tragedy—a bit like Mallory and Irvine’s doomed attempt on Everest—but nobody thought of Frege’s legacy as including much to teach us about the great epistemological and metaphysical problems presented by classical mathematics. Bill and I both strongly demur, and it was, and still is, encouraging and reassuring to have an intellectual alliance in that regard with somebody so distinguished, scholarly and insightful.

This is not a talk I have given before. I don’t have a script, and it may be that we won’t get as far as I plan—namely, provision of the solution to the Benacerraf

¹ This is a lightly edited transcript of a recording of the talk actually given at Analysis and Interpretation in the Exact Sciences, a conference held in honour of William Demopoulos at the University of Western Ontario, May 2–4, 2008.

C. Wright (✉)
Department of Philosophy, New York University, New York, NY, USA
Northern Institute of Philosophy, University of Aberdeen, Aberdeen, UK
e-mail: cw43@nyu.edu
problem! If time cuts us short, you will have to complete the solution as best you can, based on what I do get time to say. The problem, as Benacerraf described it in his classic paper (Benacerraf 1973) was that of reconciling a face-value construal of the ontology of fundamental mathematical theories: numbers, sets, points, lines, functions and so on, with what, writing in 1970, appeared to be the most likely general shape of a satisfactory theory of knowledge. That likely general shape appeared to be something that would involve essential play with causal interaction between the knowing subject and the known subject matter. Since mathematical entities don’t participate in the “causal swim”, they don’t have any causal interactions as normally understood; so there is an immediate crisis, an immediate issue about how any subject matter of this kind could possibly be knowable by natural thinkers, by human beings.

Now, the problem, as Hartry Field (1982) has remarked, is actually more general than that and can be dissociated from broadly causal conceptions of knowledge. The generalised problem is that of explaining how we could justifiably presume ourselves to be reliable in our opinions about a subject matter conceived as we standardly conceive the subject matter of basic classical mathematics, as essentially involving abstract entities. That’s the problem I want to address. It has long seemed to me that Frege had an insight that presents us with real hope of a satisfactory treatment. The promise of this general Fregean direction is really what I want to persuade you of. (Perhaps some of you need no persuasion.) Whether the detail I will go on to outline is a preferable way of implementing the general direction is a further issue.

The general direction is encapsulated in Frege’s Context Principle, on one natural reading of it. In essence, the Benacerraf problem, or the more generalised Field-problem, seem like show-stoppers because we tend to think of knowledge in a way that privileges encounter before thought: you have, we think, to have some kind of dealings with, or interaction with a subject matter before you can go on to so much as generate thoughts about it, let alone know they are true. You have first to form concepts of the material you are interacting with and then maybe you can think about it, and maybe you can know things about it. The Fregean insight, put at its most modest, is that this way of conceiving the matter is at least not compulsory. More ambitiously, perhaps, it is that the idea is definitely mistaken, not just about mathematics, but anywhere. The right way of conceiving the matter, on the interpretation of the Context Principle I am recommending, is rather to take thought first: the first question to ask is, how do we so much as come to attach content to the thoughts that we do have about numbers, points, sets, and so on? How do we grasp these thoughts? If we can give an account of our access to the thoughts, then—the suggestion is—that may be expected to carry with it some account how we might accomplish knowledge of their truth. So, what is being proposed on this interpretation of the Context principle is a reorientation, a kind of propositional turn: rather than starting from the idées fixes about acquaintance or interaction that set up Benacerraf’s bind, put the proposition first. Address the question: how do we so much as grasp propositions with this sort of content in the first instance. What model of that can we provide?

One specific kind of model that promises to be fruitful may be elicited by focusing on Abstraction Principles. These are principles which introduce a unary
term-forming operator, $\Sigma$, on some familiar kind of expressions, $a_1 \ldots a_k \ldots$,—these may be singular terms, or first-order or even higher-order predicates or relational expressions—by fixing the truth-conditions of identity statements of the form, $\Sigma(a_k) = \Sigma a$, by reference to the obtaining of some equivalence relation, $\approx$, among the items denoted by $a_1 \ldots a_k \ldots$, etc. Thus an abstraction principle (I’ll often just say, “an abstraction”) takes the form:

$$(\forall a_i)(\forall a_j) (\Sigma(a_i) = \Sigma(a_j) \leftrightarrow a_i \approx a_j)$$

An example is the principle Frege canvasses at Grundlagen §64 for the identity of directions of lines:

The direction of line $a$ is identical to the direction of line $b$ if and only if $a$ and $b$ are parallel, and another, of course, is Hume’s Principle:

$$#F = #G \leftrightarrow (\exists R)(F1-1_R G)^2$$

—for any concepts $F$ and $G$, the number of $F$’s is the same as the number of $G$’s if and only if there is a relation $R$ that is a bijection (a one-one correspondence) from $F$ to $G$.

Suppose it is philosophically admissible to view Hume’s Principle, not just as a truth (which I am sure most will agree it is) about cardinal number, but as a way of giving meaning to contexts of the kind typified by instances of its left-hand side, contexts of identity of cardinal number, in terms of the logical relationship between $F$ and $G$ mentioned on the right-hand side, which we presuppose as previously understood. If the principle can work in that way, then by explicating identity conditions for them in previously understood terms it will allow us to form a notion of cardinal numbers as a kind of object. And that explanation of their identity conditions will establish an association between numbers and properties; in effect, numbers will be given from the start as measures of one-to-one correspondent properties, a notion we can define at second-order. If this can be a legitimate approach,—if Hume’s Principle can serve as an explanation in this way,—then straight away there is the prospect of a very simple answer to the Benacerraf question. How can we know about numbers, how can we at least identify them and distinguish them among themselves, so to speak? Well, by mastering the notion of one-one correspondence and its applications and then making the transition right-to-left across instances of Hume’s Principle. We give content in the first place to propositions of identity of cardinal number by, as it were, laying down Hume’s Principle as a schematic explanatory equivalence.

This “neo-Fregean” proposal will be familiar to many of you who know the contemporary literature in the philosophy of mathematics. But I have come to think

---

2 It was, I think, George Boolos, who first used “$#$” to denote the function: the cardinal number of . . . , giving the sign its official name of “octothorpe” (to somewhat whimsical effect) whenever he uttered it. In the USA, “$#$” is also standardly termed “the pound sign”, and “hash”, of course. I don’t know why George didn’t call it by one of those names, or why he used it in the first place instead of the more usual notation, “$N_x: Fx$”.
that the general idea is at its most forceful if it is given a slightly unconventional setting, a twist which has not featured (much) in the literature. It dawned on me only relatively recently that it would be much better, from the point of view of attaining the right epistemological perspective on what is being proposed, to take Hume’s Principle not as a second-order axiom, or as a schema, but as a pair of rules of inference, conceived as for the purpose of a system of, say, second-order natural deduction. So we would simply have an introduction rule,

\[
\#I \\
\Gamma \implies (\exists R)(F1\Gamma R G) \\
\Gamma \implies \#F = \#G
\]

establishing that when you have derived from a set of premises, \(\Gamma\), that \(F\) and \(G\) are bijectable, you may infer from those self-same premises that their respective cardinal numbers are identical; and an elimination rule,

\[
\#E \\
\Gamma \implies \#F = \#G \\
\Gamma \implies (\exists R)(F1\Gamma R G)
\]

establishing exactly the reverse transition—a beautifully harmonious pair of rules! Then, in effect, we would be characterising the content of the numerical operator by reference to its inferential role. And the neo-logicist thesis could be happily expressed as that when so characterised, octothorpe denotes what is in effect a logical operation, in just the way that conjunction, quantification or negation are viewed as logical operations when taken to be characterised by their distinctive inferential rules. Of course, there is much to say about the inferentialist conception of the meanings of the logical constants in general, which has recently come in for some scepticism.\(^3\) Here I will say merely that it seems to me to provide a happier stage-setting if we are to get a sense of the kind of idea about the epistemology of number-theory that abstractionism proposes.

“An epistemology of number-theory?” Yes indeed. The implications of Hume’s Principle go far, far beyond establishing the truth-conditions of statements of numerical identity. There is an important mathematical fact here which I had better mention just in case someone does not know it: from Hume’s Principle in classical impredicative, second-order logic—and indeed in more modest higher-order logics than that—we can derive each of the five Dedekind-Peano axioms for arithmetic.\(^4\) So Hume’s Principle is in a sense, at least from a mathematical perspective, the master arithmetical thesis. I think it is an amazing fact that you can derive the basis of classical number-theory from a principle whose role is simply to describe, as it were, the conditions for the identity of cardinal numbers and the essence of their applications. Hume’s Principle just says that numbers are measures of one-one correspondence: that their relations of sameness and distinctness map the equivalences and inequivalences among properties effected by that relation. That’s what cardinal numbers are; that’s what they do. And the whole basis for their traditional

\(^3\) I am thinking in particular of Timothy Williamson’s criticisms in his (2007).

\(^4\) A helpful outline of the definitions and proofs that go into this result is Zalta (2010).
mathematical treatment turns out to flow just from that. That’s an extraordinary finding. But it’s quite another issue, of course, whether it really can be presented as in someway epistemologically foundational. (This may be one point on which Bill and I are not fully in concert.) One way of trying to make out that it is epistemologically foundational is to think of things in the kind of way I just illustrated: take Hume’s Principle not as an axiom but as a principle of inference whose content may be captured by a pair of rules constraining the use of the numerical operator, octothorpe, in the way illustrated. And now regard those rules of inference, like any sound rules of inference, as ways of extending and generating knowledge. Reason in accordance with them and you can arrive at the Dedekind-Peano axioms, which may then be regarded as known in just the way in which any propositions establishable on no undischarged premises by sound logical reasoning may be known. So, no need to think of Hume’s Principle as a first truth, any more than one need to think of, say, Modus Ponens and Conditional Proof as truths. Rather they are, or so we would like to think, sound but primitive principles of logical inference. This is how Hume’s Principle may accrue epistemological significance when taken as a pair of principles of inference. I don’t think this is just a cosmetic difference. I think the epistemological perspective alters significantly. In particular, we should so far as possible think of issues to do with the justification of Hume’s Principle under the aegis of whatever should be said about the justification of basic logical rules. I do not here assert that it is unproblematic to sustain the analogy. But I do think that it sets the stage for the structure of a philosophically fruitful discussion: a discussion that focuses on what one should say about the epistemology of basic logic in general, and—whatever that may be—what problems there may be about viewing Hume’s Principle in the light of that comparison, as encoding a pair of rules of inference of primitive logical standing.

Well, those of you who know something of the debates will know there will be a whole host of at least apparent problems. We know, to begin with, that some principles of this broad character are not acceptable. Frege’s Axiom V, or at least the specialisation of it to extensions of concepts:

\[(\forall F)(\forall G)(\{\text{ext} : Fx\} = \{\text{ext} : Gx\}) \iff (\forall x)(Fx \leftrightarrow Gx)\]

looks like an abstraction principle. At any rate, it likewise configures an identity between abstract objects on the left and a second-order equivalence relation—co-extensiveness—on the right. But this principle leads to Russell’s paradox. Nor is the worry just a matter of there being inconsistent relatives of Hume’s Principle. There are also individually satisfiable but pairwise incompatible principles of the same general structural kind, each of which might seem in isolation to be acceptable, but which cannot both be acceptable. So which are the acceptable and which are the unacceptable such principles, and what are the criteria we should use to distinguish the acceptable cases from the unacceptable?\(^5\)

---

5 This question was first forcefully urged by George Boolos (1990). The problem was further explored and deepened by Alan Weir’s (2003), and has recently received the attention of a full special number of Synthese, edited by Oystein Linnebo (2009).
That is what I once christened the Bad Company problem. I am going here to assume that it is solved, and will say no more about it today! We can pick it up in discussion if you like. A second major issue is the so-called Julius Caesar problem: that of explaining the distinction between objects introduced by abstraction and items of other kinds. This is not just a metaphysical problem. To found arithmetic on Hume’s Principle will demand that meaning be given to open sentences of the form, “x = #F”, an understanding of which presumably requires understanding their instances, for instance “Julius Caesar = #(x is a planet)”. But how does Hume’s Principle contrive to give sense to such “mixed” identity contexts? Again, I won’t address this matter further here.

Then there is a worry about what happens when you unpack instances of the right-hand side of Hume’s Principle into primitive notation. “F 1-1R G” will come out as something like:

\[(\exists R)[(\forall x)[Fx \rightarrow (\exists y)(Gy & Rxy & (\forall z) Gz & Rxz \rightarrow z = y)]

& (\forall y)[Gy \rightarrow (\exists x)(Fx & Rxy & (\forall z)Fz & Rzy \rightarrow z = x)]\]

in which there are *impredicative* first-order quantifiers—quantifiers whose range will need to be taken to include the referents of the terms introduced on the left-hand side of Hume’s Principle—the octothorpe terms—if Frege’s constructions are to go through (because we will need to work with numbers of properties that are themselves instantiated by numbers: for instance, the number belonging to the property, *predecessor of 2 in the series of natural numbers.*) So on the right hand side we are here quantifying over the very things we are using the principle to introduce. That may raise a concern about circularity. Is it a good concern? I don’t think so—but I am not going to discuss it. 6

Finally, there are worries about the underlying higher-order logic. Full classical impredicative second-order logic is a very powerful, in some ways unsurveyable system. We don’t need its full strength of to derive the Dedekind-Peano axioms from Hume’s Principle, but we do need it when it comes to giving abstractionist foundations for classical real analysis—that is, deriving axioms for a completely ordered field from a suitable abstractionist base. That will need the full power of classical second-order logic. And, even for the more modest purpose of founding arithmetic, there are worries about the unavoidable impredicativity of the higher-order quantifiers that will be required.

So there are many dragons to slay, to run this programme through—not all of them are dead! But today, I want to focus on what I take to be the central *ontological* idea, about which there may seem to be serious issues even before any of the foregoing concerns arise. To fix ideas, it may be helpful to

---

6 It is discussed in my exchange with Michael Dummett in Schirn ed. (Dummett 1998, Wright 1998a, b)
take a first-order case. Consider Frege’s own example of direction—the first level abstraction:

\[(\forall a,b) \text{Direction}(a) = \text{Direction}(b) \iff a//b\]

could be conceived as laying it down as necessary and sufficient for the identity of the directions of a pair of lines that they be parallel. That we may legitimately so conceive the Directions abstraction is the master thought I want to concentrate on. There is

a prima facie commonsensical idea that is offended by that thought. We don’t naturally suppose that there is any inexplicit additional structure on the right-hand sides: the right-hand side contexts are, we think, just as they appear to be, about lines and a relation on lines. But the explicit ontology on the left-hand side is of course different. Here there are not just lines and the relation of parallelism. There are extra things: directions, being referred to. So how can these two types of context be equivalent? It looks as if the left-hand kind of context makes a specific ontological demand that the right-hand kind of context does not. So how can it possibly be admissible for us to lay the principle down? How can we possibly successfully stipulate that the truth of the one kind of context is to be sufficient, without further ado, for the truth of the other?

Well, there are three options:

(i) That indeed we cannot—that the truth-conditions of the left-hand sides are stronger than those of the right-hand sides; or

(ii) That we can—but only at the cost of arguing that the syntactic structure of the left-hand sides masks their logical form, so that the direction-terms aren’t genuine, referential singular terms, though they look like it on the surface; or

(iii) That the two forms of context do indeed have the same truth-conditions but that the left-hand sides make explicit something that’s only implicit on the right.

The third is the way of looking at the matter that’s implicit in Frege’s metaphor of “recarving” of content. When you recarve content, you preserve truth-conditions in the broad sense of the necessary equivalence of the right- and left-hand sides: it’s still necessarily the case that the left-hand side is true iff the right-hand side is. But you re-orchestrate those truth-conditions within a different conceptual repertoire. You reconceptualise the states of affairs thought of as the truth-conferrers of the right-hand sides. By forming the concept of direction via just this pattern of introduction, you thereby put yourself in position to recognise some objects associated with those states of affairs that you overlooked before.

Now of course, we also have to maintain that this new conception of the ontological implications of the right-hand sides is to be consistent with their initially epistemically unproblematic character. It mustn’t be the case that by making the move to institute the new abstracts, we thereby “up the ante” as far as what it takes

---

7 *Grundlagen* §64.
to know or verify the right-hand sides is concerned. So our problem is to address that issue squarely, to explain how the trick can be pulled. My main purpose here is to outline how such an explanation might go.

In order to do that, I want to begin by separating two questions, one metaphysical, the other epistemological. The metaphysical question is:

What does the World have to be like in order for the best examples of abstraction principles, whatever they are—the best examples being those which survive a solution to the Bad Company problem, contain only unproblematic forms of quantification, and so on—what does the World have to be like in order for such abstraction principles to be true?

And associated with that, the epistemological question is

How do we know, what reason do we have to think, that the transition right-to-left across the biconditional in examples of the best kind of abstraction is truth-preserving? What reasons can we give ourselves for thinking that when we move, for instance, from the parallelism of a pair of lines to the identity of their directions, we wind up with a truth? In short: how do we know that the World is as the answer to the metaphysical question requires it to be?

At this point we come to a parting of the ways. There are two ways of thinking about the metaphysical question. One is to think that we had better do some metaphysics! We better try to win through to a perspective from which when the right-hand side of an instance of a good abstraction obtains, the World will co-operate and deliver up objects of the appropriate sort, objects that reflect in their identity-conditions the partitions effected by the equivalence relation on the right. An example of a metaphysical theory that will provide for that is Matti Eklund’s Maximalism (Eklund 2006). That’s the view that, at least when we are concerned with abstract objects, all possible varieties actually exist! There is a plenitude of abstract objects, so merely form a consistent concept of a certain kind of abstract object and there are guaranteed to be abstract objects to answer to your concept. If we are maximalists, then we have a supplementary metaphysical theory that will underwrite the transition from right to left across the instances of a good abstraction principle.

The first approach to the metaphysical question, then, is to try to grease the right-to-left transitions by well-motivated supplementary metaphysics. The truth-conditions of the two halves of an instance of a good abstraction will be argued to coincide courtesy of a metaphysical guarantee that the semantically additional commitments incurred by the left-hand sides are met as a matter of independent metaphysical necessity.

But there is a second possible approach: to show you don’t actually need any such metaphysical greasing. Personally, I much prefer the sound of that! So my project is to show how the transition from right to left across instances of the best abstraction principles doesn’t need additional collateral metaphysical assurances; that it is somehow already guaranteed because the commitments of the two halves are already strictly the same. As you see, I propose to take the recarving metaphor very seriously. What we need, though, is not to rely on that metaphor, but to give acceptable sense to it.

Here is a simple-minded answer to the metaphysical question. What does the world have to be like in order for the best examples of abstraction principles to be
true? Answer: their Ramsey sentences have to be true. The Ramsey sentence for an abstraction principle is just the result of existentially generalising into the place occupied by its abstraction operator; thus, for the case of Hume’s Principle, what we have is

$$(\exists f)(fF = fG \leftrightarrow (\exists R)(F 1-1_R G))$$

There is a function whose values for a pair of properties are the same just if those properties are one-one correspondent. There may of course be more than one such function; there are philosophical issues raised by the plethora of functions that may play that role. I am not going to discuss that particular kind of concern; it may come up in discussion. I am going to make the assumption that, provided there is at least one, that will be good enough. So our question is, what guarantees the existence of such a function? What guarantees that, whenever we have an equivalence relation associated with a good abstraction,—and whatever a good abstraction is,—there is going to be a function to deliver the truth of the corresponding Ramsey sentence? That is the heart of the issue.

Let me quote a passage from George Boolos worrying about just this point. George writes

...what guarantee have we that there is such a function from concepts to objects as [Hume’s Principle] and its existential quantification [Ramsey sentence] take there to be?

I want to suggest that [Hume’s Principle] is to be likened to “the present king of France is a royal”

—What he has in mind is a sentence that is, as it were, analytic modulo a presupposition of existence. It is guaranteed that the present king of France— if there is one—is a royal, but

...we have no analytic guarantee that for every value of “F”, there is an object that the open definite description, “The number belonging to F” denotes...

Notice incidentally George’s assimilation of singular terms introduced by abstraction to open definite descriptions. I actually think this is an important mistake. It’s a natural mistake because of our practice of informally paraphrasing terms introduced by abstraction as, for instance, “The number of Fs”, “The direction of a”, and the like; and those sound like definite descriptions. But they are not—I will explain why in a minute. Boolos continues

Our present difficulty is this: just how do we know, what kind of guarantee do we have, why should we believe, that there is a function that maps concepts to objects in the way that the denotation of octothorpe does if [Hume’s Principle] is true? If there is such a function then it is quite reasonable to think that whichever function octothorpe denotes, it maps non-equinumerous concepts to different objects and equinumerous ones to the same object, and this moreover because of the meaning of octothorpe, the number-of-sign, or the phrase “the number of.”

—So he is granting, in effect, that we have fixed the sense, up to a point, of the octothorpe functor by the proposal of e.g. Hume’s Principle as an implicit definition.
The issue concerns whether the sense so fixed is such that we are assured of a reference—

But do we have any analytic guarantee that there is a function which works in the appropriate manner?
Which function octothorpe denotes and what the resolution is of the mystery how octothorpe gets to denote some one particular definite function that works as described are questions we would never dream of trying to answer. (Boolos 1997, p. 306)

Actually that last remark is changing the question slightly—it raises the question of the determinacy of reference of e.g., “#(x is a past US President)”, and that’s another issue. Our concern is with whether or not there is any reference at all.

Let’s consider the question at its most general. If somebody asks you, how do you know that there is a function of a certain kind, how might you answer? It is natural to think that what you do is to consult your repertoire of functions,—your favourite set theory, for example,—and see what it can deliver to meet the purpose at hand. But that is not the right way to take the question in the context we are in. We might have no background set theory. If abstraction is a reasonable procedure, it should be reasonable independently of any previous ontological commitments save those implicated on the right-hand sides of the principle in question. So, even for someone hitherto totally innocent of the notion of a function in general, and without any entrenched repertoire of functions or sets, it should, once he is introduced to the idea of a function, be a reasonable question: what grounds are there for thinking that there is any function fit to serve the purpose of a particular abstraction?

Now, it cannot always be the case that the way to answer an ontological challenge is to, as it were, produce some item(s) that fit the relevant bill. Suppose you ask, “How do we know there is any object that meets a certain condition, C?” And I say, “Well, here is one! This will do it.” That answer, “Here is one. This will do it”, involves my bringing the thing, whatever it is, under some other already available concept. In giving that answer appropriately, I must already have another concept of the thing in question that I can then marshal to serve the purpose in hand. But that cannot be the only way we can in general answer ontological questions, since it is obviously regressive. The question we are asking—and it won’t make it any easier to put it like this, but it will give it the right focus—the question we are asking is: how in fundamental cases should we assure ourselves of existence? The fundamental cases are the cases where, exactly, there is no presupposition that the objects we are looking for, if such there be, fit any other anterior concepts that we have.

It is easy to overlook the force of this, because with ordinary middle-sized dry goods there is always the resource of demonstrative concepts. If somebody says, “How do we know there is any object of such-and-such a sort?” I may be able to reply, “Well look, there is one,” and point it out. But with abstract entities that cannot be available. So if we are taking the possibility of an abstract ontology seriously, we need to take on the question: what would it be to satisfy ourselves that a class of fundamental terms—terms that if they refer at all, refer in this fundamental way to things of which we may have no other concept, yet—what would it be to have evidence that such a class of terms refer?
Mull that question!

We are, to be sure, not yet talking about objects; we are in the first instance talking about functions, especially the putative referent of octothorpe. But I propose to understand the question in such a way that if we satisfy ourselves that the function exists, that will carry in train the existence of the referents of the complex terms formed by using it. If the function exists, it will have a value for each of the appropriate range of arguments, so the values will exist too.

So, how can we address the question, what in general should fundamentally satisfy us that a function exists, where the stress on “fundamentally” reminds us that we are not just going to cop out and say, e.g., “This set can serve as the relevant function.” Forget about sets. (After all, if there are any, how are we fundamentally assured of their existence?) How can we be assured that octothorpe denotes?

Let’s ask a related but an importantly different question. What should satisfy us that a property exists? Functors are one kind of Fregean incomplete expression, so in the hope of illumination, let’s consider predicates: the basic and canonical genre of incomplete expression. What should it take to satisfy us that a predicate has a referent, so that a corresponding property (or if you prefer, Fregean concept) exists?

There are two broadly different ways of approaching the question. On one conception we are asking a question about the nature of divisions “out there”, in the real world. When we ask if a property exists, we are asking whether the satisfaction condition associated with the predicate takes us to—whether to satisfy that satisfaction condition is to have—a fundamentally real property, a “natural joint”. It is in this spirit that someone who thought that colours, for example, if they exist at all, would have to be natural kinds, might say there is no such thing as the property of being red. It has turned out there is an immense physical diversity of conditions which result in a preponderance of red light being emitted by a surface. So there is no real essence associated with redness, there is no natural kind of redness. One conclusion in response to that discovery is to say that the predicate “red”, although associated with a coherent satisfaction-condition, presents no real property.

Such a view of properties is what has come to be known as a sparse view—Sparsism. It’s the view that properties are metaphysically sparse, that there are many more significant distinctions that we can draw using predicates with well-conceived conditions of satisfaction than there are actual properties that correspond to them, in the real world. But contrasted to that is an Abundance metaphysics of properties. Abundance says, “No, a property is just a way things can be; and when you have got a determinate satisfaction-condition associated with a predicate, there is of course a way things can be whereby they satisfy the predicate, namely: satisfying that condition.” For abundance, there is no deep issue about predicate reference. Sense—having a satisfaction condition—suffices for reference in the case of predicates, more or less. There will, to be sure, be some predicate expressions, like “... is Wright’s favourite colour”, which, although meaningful, will still fail of reference because Wright has no favourite colour. But in the general run of cases, from the point of view of Abundance, it suffices for a predicate to present a real property that it have a well-explained sense, that is, that it be associated with a coherent satisfaction-condition.
So there are two quite different ways of thinking about the issue of property existence. Now here is the matter I want to press. Is there a way of so conceiving of functors that reference for them too is abundant? Is there an analogue for them of the liberal conception just articulated for the case of predicates? If there is, then merely conferring such a sense on a function expression that it is enabled to make a determinate contribution to the truth-conditions of statements containing it, will suffice for the existence of a denoted function.

But there is, you will likely say, an obvious problem. We cannot treat the matter that simply. You may succeed in conferring a sense on a function expression and still fall foul of constraints of uniqueness and existence of value, required if there to be a function that it stands for. You may confer a sense on “#G” in such a way that it turns out that there is more than one thing that can count as the number of Gs; or you may confer a sense in such a way that, alas, there is no such thing. Those are possibilities that the mere conferral of sense cannot pre-empt.

So we should revise our question: how can we explain—an abundance conception, as it were, of functions which pays proper heed to these additional constraints of uniqueness and existence, unmatched in the case of predicates, but still make the question of the reference of octothorpe and its ilk relatively easy and non-metaphysical, as on an abundance conception of predicate-reference and properties?

Let us think about uniqueness first. I claim there is no coherent uniqueness worry in the cases that concern us. The reason why not is the same as the reason why it is a mistake to think of the singular terms introduced by abstraction principles as definite descriptions. If they were definite descriptions, they would be of this form:

\[ \iota x : xR F \]

There would be a relational condition on F:

\[ \ldots RF \]

with a gap for a first-level argument, which one would then bind with the description operator. That would be the semantic structure. But for that to work, there has to be an embedded relation. So we would need a notion in general of what it is for an object to “number” a concept, without any presupposition that this relation is functional. But I claim that when sense is given to “#F” and its ilk via Hume’s Principle, there is no such half-way house. There is no intermediary grasp of such a numbering relation, F is numbered by x, which we then, in forming octothorpe-terms, take to be many-one; there is simply no appeal to any such concept.

I am not of course denying that numerical terms, introduced via Hume’s Principle, are semantically complex. The point is rather than they do not have—are not, by the means of their explanation, endowed with—the kind of semantic complexity possessed by a definite description, wherein an operation is executed upon a predicate, or a relation with one unfilled argument place, requiring that it is satisfied uniquely. And in the case that concerns us, there could be a legitimate worry about uniqueness only if there were such a relation and a possible doubt about
its functionality. Then there might be more than one \(x\), such that \(F\) is numbered by \(x\), so the legitimacy of the term, “The unique \(x\), such that \(x\) numbers \(F\)”, would be hostage. But unless there is such an embedded relation, there is no space for a coherent doubt about uniqueness. Since no such relation is either presupposed by or explained in the course of fixing the sense of a range of singular terms, “\(\Sigma(a_k)\)”, by abstraction, a lucid doubt about the uniqueness of reference of abstract terms is pre-empted. There are some kinds of semantically complex singular terms that nevertheless aren’t definite descriptions. And as I have argued, that’s important.

What about existence? Well, for the same reason, you cannot either entertain an analogous worry about existence. If the worry was: maybe in some cases, \(F\), there is no object that numbers \(F\)—maybe some properties are not numbered by anything,—that’s a doubt that is going to make sense only if we have a numbering relation to work with, in terms of which to formulate the doubt. Of course, there are nonsortal properties that, plausibly, are commonly admitted to have no numbers—terms for stuffs and kinds, like \(water\), \(uranium\), and \(treacle\), and vague attributives like \(yellow\) and \(warm\). But these are outside the range of the cardinal number operator in any case. The doubt about existence should be a doubt about whether the function associated with octothorpe is total within its proper range, that of, as the matter is usually expressed, sortal predicates. If there is no appropriately embedded relation, \(x\) numbers \(F\), of which an understanding of octothorpe requires grasp, there is no such coherent doubt.

But let me hasten to say, as hackles rise, that I don’t expect that these considerations are likely to be received as completely disposing of any worry about existence. I can imagine someone saying, “Very well, then I can’t articulate my worry in that particular way. But surely we can still intelligibly raise the question: what if there just aren’t any such things as cardinal numbers? The left-hand sides of instances of Hume’s Principle collectively call for an ontology. What if there just aren’t any such objects? What have you to say to address that?”

And this looks like a sticking point for any attempt to go abundant on functions, to try to analogue the case to that of predicates and properties on an abundant construal. We can grant that the fixing of sense for an abstraction operator all but gets us a function, so to say, and that the function, if any, that it gets us is manifest in the sense thereby given to the operator, as an abundant property is manifest in the sense of its associated predicate. But the point has not gone away that there will actually be such a function only if there is an appropriate range of values. So, you may say, there is no way to finesse the question whether that is so. Boolos’s worry—How do we know that there is any such function as the referent of octothorpe?—still remains to be addressed. We will be entitled to take it that there is such a function only if we can assure ourselves of the existence of a suitable range of values. And whether that is so cannot be resolved just by bestowing appropriate sense on octothorpe, even if we are sympathetic to the precedent of an abundant view of properties, and sympathetic to the attempt to enlist it to help out here.

But whether this is a sticking point depends on whether the residual existential doubt really is intelligible. Here is where it is important to remember that we are talking about potentially fundamental terms, that is, terms such that, if they refer, we may have no other apparatus in terms of which to pick out their reference. So, in
its most general form, the doubt that is being pressed is a doubt about fundamental terms: how do we know that any such terms refer? Since they are fundamental, we cannot assure ourselves by cross-identifying their referents with those of other expressions whose reference is not in doubt. So what assurance can we obtain, not just with numbers but with any objects for which putative means of reference is being introduced for the first time?

It is at this point that the Context Principle comes into its own. For so far as I can see, there is only one possible shape for an answer to take. To verify that a fundamental term, “a”, refers, we have to verify that some context, “Fa”, that configures it in a reference-demanding way is true. It cannot be a matter of verifying that “a = q” where q’s reference is not in doubt. If the matter could be addressed like that, then “a” would not be fundamental. But in order to verify a relevant proposition, “Fa”, you have first to understand that proposition and to have some accepted conception of what counts as good evidence for it. So you must have already established the content of thoughts of that kind, and you must have established those contents in such a way that allows that we have some conception of what we count as good evidence for their truth. Well, that is exactly what an abstraction principle is proposing to give you for contexts of identity of its proper abstracts. Hume’s Principle assures us that the best, canonical evidence for the truth of thoughts concerning numerical identity consists in finding out facts about one-one correspondence of properties. That’s the essence of the abstractionist proposal in the first place. But the more general point is that doubt whether some potentially fundamental class of terms refer, has to be handled, if it is to be handled at all, in broadly the way that the Context Principle schematises, by fixing the content of claims embedding the terms concerned and then, guided by that fixed content, by seeking out appropriate evidence.

The point is absolutely general. Imagine a scenario in which the only terms that we have for referring to middle-sized particulars are sortal demonstratives, like “This pen”, “That notepad”, and so on; that’s all we have got. Our conception of the content of claims containing such terms in reference-demanding ways would still be that, in the most basic case, the relevant kind of evidence for their truth is sense-experience: if you want to verify a claim of that kind, you need to attend to the object concerned and check it out. But now suppose that someone said, “Yes but, you know, even given that kind of evidence, mightn’t it be the case that terms of this kind—sortal demonstratives for perceptible middle-sized particulars—just don’t refer; that there are actually no such things as pens, notepads, etc?” Well, whatever the right thing to say in response to that, it’s a very familiar kind of position: it is material world scepticism! That we are inundated with experience is conceded but now there is supposed to be a doubt about whether certain kinds of claims requiring a certain ontology are true for which, if anything is good evidence, it is experience; there is simply nothing better than that. If the doubt is whether the best possible evidence is good enough, then that’s just the familiar shape of a classical sceptical doubt.

What I am saying now brackets all other issues about the satisfactoriness of the explanations offered by Hume’s Principle. The target is someone who agrees that, yes, an abstraction can successfully fix meaning at least to the extent that Boolos
agreed—it can fix truth-conditions conditionally upon the existential presuppositions made by its left-hand sides—but then claims that a doubt remains unaddressed whether there simply are any such things as the objects thereby presupposed. This doubt, I claim, is essentially of the same shape as the sceptical doubt of one who grants that if there are material objects at all, then one’s sensory experience provides excellent evidence for how things are with them, but denies that it provides sufficient evidence that there are indeed such objects.

I do not, however, suggest this parallel in a spirit of dismissal. I do say that anyone who regards material world scepticism as absurd owes an explanation now of why scepticism about abstracts, when the contents of statements about them are explained by means of abstractions, is not also absurd. But my own sense is that material world scepticism traffics in an intelligible doubt, so I am not content to leave the matter there.

I want to persist with the prospects for an abundant ontology of functions, on the closest analogy we can sustain with an abundant ontology of properties, thought of as given by the satisfaction-conditions of predicates. What should that analogy be, exactly? There are two polar views here that we need to steer between. On the one hand, there is the view illustrated by these remarks of Peter Sullivan and Michael Potter writing about Hume’s Principle:

What is the image they have in mind? It is a kind of Fisherman’s view. It represents us, in laying down Hume’s Principle, as both introducing a conception of a distinctive kind of object, a kind whose instances are to behave in certain ways, and in doing so, as casting our net at the world and hoping to enmesh objects that behave in just the ways that the abstraction requires. So then it is just down to what’s “out there” whether we catch any fish or not. In parallel, we have a certain conception of what gold is, we point it at the world, and it’s just down to what is out there whether anything suitable is delivered, whether there is any worldly substance that appropriately underwrites our conception.

That’s one extreme. The Fisherman’s view embraces the analogue, for objects, of a sparse conception of properties. The exact analogue, correspondingly, of an abundant conception of properties will be Meinongianism: that just as every predicate associated with a well-explained satisfaction-condition determines an (abundant) property, so every significant singular term has some kind of referent. It might be a non-existent referent, but it is a referent nonetheless.

We—Fregeans—want something between those two views. We want the reference of terms introduced by good abstraction principles to be real, just as abundant properties are real. For the distinction between the abundant properties and the sparse properties is not one in point of reality; it’s to do with the contrasted natures of the two types of property. There are some properties that are interesting as far
as dividing the world up into its fundamental kinds is concerned and some that are not, that don’t do that, but simply answer to the distinctions that can be drawn by significant predication. It is very easy to get assurance of the existence of the second kind of property. It’s enough to have a well-behaved predicate. It’s not so easy to get assurances of the first. But we—Fregeans—don’t want to need the deeper kind of assurance that goes with the Sullivan and Potter picture, the Fisherman’s picture. That picture is what sets up the problem. We want the assurance of reference to be easy, but we also want it to be world-driven, an assurance of reality, not of Meinongian subsistence. We don’t want it to be the case that abstracts reduce to creatures of language, or our thought.

So then, let us propose a conception of a certain kind of object which is abundant—which as closely as we can run the parallel, stands to the significant use of singular terms as abundant properties stand to the significant use of predicates—but is also appropriately disciplined by the world. We get an exact parallel with the kind of discipline required if we recall Aristotle’s views on properties in general. For Aristotle, properties, in order to exist, need real instances. There are no empty properties. Combine that conception of a property with the abundant view: the result is a conception of properties such that to every satisfied predicate—every predicate with determinate satisfaction conditions which are actually satisfied—but only to such predicates, corresponds a property. So the sense of the predicate all but takes you to the existence of a property. Properties are transparent in the senses of the predicates that express them. But a predicate’s possession of sense, on this hybrid conception, no longer ensures that it presents any property. It is necessary in addition that the world steps in and actually delivers something that satisfies the predicate before you get to a property at all: the world must ensure that the predicate applies to something. The truth of an atomic predication of it will thus suffice.

That’s as close as I know how to guide you to the way I believe we should understand the general conception of an abstract—the general notion of the sort of object it is that one obtains by successful abstraction. In contrast with any Meinongian view, we need the truth of the right-hand side kind of context before we can claim existence. It is not enough that the abstract terms have a sense. Appropriate (atomic) statements containing them have to be true. But those truths can be objective. And the truth of the left-hand sides of instances of abstraction principles will be an objective matter just if that of their right-hand side counterparts is, because that is given as a necessary and sufficient condition. Thus where it is objectively so that a pair of properties are one-one correspondent, it will correspondingly be objectively so that some one number is the number of them both. But there will be no metaphysical hostage, no “fishing”, in drawing this conclusion about their number. The reason is that numbers, like all abstracts, are to be compared to abundant Aristotelian properties: entities knowledge of which is fully grounded in knowledge of the truth of atomic predications and identity statements, respectively, and embodies no further conjecture about the nature of the World.

It is thus very easy on this conception of what an abstract object is to know (some things) about them, at least in the best case. The possibility of knowledge about abstracta just falls out of the way in which the content of thought about them
is fixed in the first place, just as was promised by the “propositional turn”, just as it should do. That’s the outline of the (neo-) Fregean solution to Benacerraf’s Problem. No doubt it could use some further filling in.

References