PETER MILNE

12 Frege's folly: bearerless names and Basic Law V

1 FREGE ON TRUTH

Frege tells us surprisingly little about truth. And some of what little he does say, he repeats:

One can, indeed, say: 'The thought that 5 is a prime number is true'. But closer examination shows that nothing more has been said than in the simple sentence '5 is a prime number'. The truth claim arises in each case from the assertoric sentence, and when the latter lacks the usual force, e.g., in the mouth of an actor upon the stage, even the sentence 'The thought that 5 is a prime number is true' contains only a thought, and indeed the same thought as the simple '5 is a prime number'.[1]

Whereas in the much later 'Thoughts' from 1918 it is sameness of thought/content that is emphasized:

It is also worth noticing that the sentence 'I smell the scent of violets' has just the same content as the sentence 'It is true that I smell the scent of violets'.[2]

in the 1897(?) 'Logic' we find the emphasis is on assertion:

If I assert that the sum of 2 and 3 is 5, then I thereby assert that it is true that 2 and 3 make 5.[3]

Strictly, the claim about assertion is distinct from that about content (sameness of thought). It could be that it is a fact about assertion, e.g., that it aims at truth, as people say, that nothing different is accomplished in asserting that the sum of 2 and 3

[1] 'On sense and reference', *CP*, p. 164.
[2] 'Thoughts', *CP*, p. 354.
[3] 'Logic', *PW*, p. 129.

is 5 and that it is true that the sum of 2 and 3 is 5, rather than a fact about the contents asserted. But this is not Frege's view. In both 'On sense and reference' and 'Thoughts' Frege points to a certain redundancy where truth is concerned. For any assertoric sentence *p*, *p* and 'It's true that *p*' have 'just the same content' ('Thoughts'), they 'contain the same thought' ('On sense and reference'). Nevertheless, in contemporary terms, Frege is no deflationist. He asks rhetorically, 'And yet is it not a great result when the scientist after much hesitation and laborious researches can finally say, "My conjecture is true"?'

A true thought refers to the True; properly speaking, truth is not a property of true thoughts ('On sense and reference', p. 164; 'Thoughts', pp. 354–5). Frege is, to modern eyes, surprisingly lax regarding the difference between a truth-operator and a truth-predicate. While truth may not be a property of true thoughts – they refer to the True rather than falling under some concept – it is hard to see a truth-predicate as doing anything other than picking out some property common to all true sentences: to be sure, a language-relative property, and one possessed derivative upon the sentences expressing (in the context of use) a thought that refers to the True.

As is well known, Frege rejects the correspondence theory of truth. The content of the word 'true' is, he says, *sui generis* and indefinable; the meaning of the word 'true' seems to be altogether *sui generis* ('Thoughts', pp. 353, 354). Here's the best gloss I can put in his remarks about truth.[4]

There should be no difference at all between asking, for suitable *p*, whether *p* and whether it's true that *p* (e.g., there is no difference between asking whether Jena is a city on the Saale and asking whether it is true that Jena is a city on the Saale). This gets us to Frege's fundamental thought: that knowledge of what it takes for a particular sentence to be true cannot be something added on after an understanding of the sentence itself. An understanding of what it is to be true ('in general') cannot come after an understanding of the rest of the language, as would be possible if it were possible to offer a proper *definition* of truth.

[4] I owe some of this to Luis Fernández Moreno, 'Die Undefinierbarkeit der Wahrheit bei Frege', *Dialectica*, 50 (1996), pp. 25–35.

Frege says: 'If I do not know that a picture is meant to represent Cologne Cathedral then I do not know what to compare the picture with in order to decide on its truth.' In the case of a declarative sentence there is *never* an equivalent problem, *if one understands the sentence*, i.e. *if one knows the thought it expresses*. If one understands a sentence then one knows how the world must be for (the thought expressed by) the sentence to be true.

A thought is something for which the question of truth can properly be raised. The *only* thing for which the question of truth can properly be raised is the sense expressed by an assertoric or interrogative sentence, so this is a thought. 'Truth does not consist in correspondence of the sense with something else' ('Thoughts', p. 353). Although it is not at first obvious, this connects with something Frege says much later on in 'Thoughts': 'What is a fact? A fact is a thought that is true' ('Thoughts', p. 368). This is, at first glance, a very odd conception of facts. Thought through carefully, it may not be compelling but it does tie in with how we should think about Fregean thoughts.

Think of declarative sentences in a language (or perhaps some of the stuff inside your head) as 'representations'. One uses them to represent 'ways the world might be', to speak very loosely. In asserting a declarative sentence one is claiming things *are* that way. What is the *content* expressed by a sentence, the content of what does the representing? It is what is represented: a way things might be. When is the representation correct? When is the sentence true? – When the way it represents things as being is the way (or, perhaps better, a way) things are. Put another way, a Fregean thought is *not* a 'picture' or representation of a way things might be; rather, it is the way they are represented as being.

Now, if one ascribes truth to sentences, then, as said above, a sentence is true if the way it represents things as being is a way things are. But if, like Frege, one ascribes truth primarily to thoughts, then it is what is represented, not the representation, that is true (or false). And the various ways things are are just some among the way things might be. They're the true ones, the facts. There is no separation between truths, true thoughts, and facts. Facts and truths are the same things, namely – staying with the way I have been speaking – ways things are. Thus Frege ends up with what by his own lights is a perfectly good correspondence theory: thoughts are true if they are identical with, i.e. correspond perfectly to, facts. Frege is an identity theorist of sorts.[5]

This argument turns on Frege (i) taking, if we may speak loosely, truth to be a property of what is represented, not of what does the representing, and (ii) treating what I have called ways things might be fully on a par with what I have called ways things are. Now, one may reasonably object to an identity theory of truth of Frege's sort, a theory that identifies facts with true propositions/thoughts, that the role of the world in determining which propositions are true has been lost. It's all very well identifying facts and true propositions but what is it that fixes the facts? We may resolve that matter if the world is, to borrow a phrase, the totality of facts, but only by granting an explanatory priority to facts that is not Frege's. Truth is, for Frege, unanalysable. Jena falls under the concept *city on the Saale* because the thought *That Jena is a city on the Saale* refers to the True. That is the order of explanation, not, if we take Frege at his word, the reverse. We may limn the laws of truth – the laws of logic – but we cannot give a general account of what it takes for a thought to be true.[6]

What I have just elaborated is the picture that emerges most clearly from 'Thoughts', a late post-paradox piece that contains Frege's only extended discussion of truth – compressed as it may be – in an article that covers an awful lot of ground. How much or how little can be projected back to the 1890s is far from clear. On the one hand, there is the complete absence of any mention of concepts in 'Thoughts'. On the other, in the close to contemporaneous 'Notes

[5] Cf. Julian Dodd and Jennifer Hornsby, 'The identity theory of truth: Reply to Baldwin', *Mind*, 101 (1992), pp. 319–22.

[6] Russell objected to coherence theories of truth that there could be more than one maximal consistent set of beliefs. Likewise, there can be more than one maximal consistent set of Fregean thoughts and it would seem that the world must somehow play a role in determining the set which is the set of true thoughts (facts). Perhaps because he was concerned as much with mathematical truths as empirical ones, Frege says little regarding contingency and modality generally.

Crispin Wright has objected persuasively that coherence theories have a hard job in accommodating contingency say in the sense of genuine chance events ('Truth: A traditional debate reviewed', in S. Blackburn and K. Simmons (eds.), *Truth* (Oxford: Oxford University Press, 1999), pp. 203–38. at pp. 221–2). That is not the objection here. The objection here to a Fregean identity theory is more fundamental: it is that we are given no clue as to what fixes the set of truths. The coherence-theorist has a story to tell about that; Wright's complaint is that that story is defective.

for Ludwig Darmstaedter' (as the editors of the *Nachlass* call it), we find doctrines of long standing reiterated and others portrayed as being of long standing.

What is distinctive about my conception of logic is that I begin by giving pride of place to the content of the word 'true', and then immediately go on to introduce a thought as that to which the question 'Is it true?' is in principle applicable. So I do not begin with concepts and put them together to form a thought or judgement; I come by the parts of a thought by analysing the thought.

… The first thing that strikes us here is that a thought is made up out of parts that are not themselves thoughts. The simplest case of this kind is where one of the two parts is in need of supplementation and is completed by the other part, which is saturated: that is to say, it is not in need of supplementation. The former part then corresponds to a concept, the latter to an object (subsumption of an object under a concept).[7]

What is clear is that Frege's *practice* in *Die Grundgesetze der Arithmetik* is quite different. In §31 he expounds what purports to be a proof that every proper name (including sentences) and first-level function name formed according to the procedures he lays down in §30 has a reference. To carry out the proof he proceeds 'recursively', starting with the primitive signs, using the explanations previously given of them, and moving on to more complex expressions. This 'proof' immediately precedes a section with the title 'Every proposition of *Begriffsschrift* expresses a thought', the first sentence of which summarizes what has, supposedly, been accomplished in §31: 'In this way it is shown that our eight primitive names have denotation, and thereby that the same holds good for all names correctly compounded out of them.' As the system of *Grundgesetze* is inconsistent, the proof must fail. The nub is the

[7] 'Notes for Ludwig Darmstaedter', *PW*, pp. 253–4. Missing from the Darmstaedter notes is any mention of extensions of concepts. The notes end with brief remarks on numerical quantifiers – expressions of the form 'There are *n* …' – which on Frege's reckoning are second-level concepts. The closing lines are:

But still we do not have in them the numbers of arithmetic; we do not yet have objects, but concepts. How can we get from these concepts to the numbers of arithmetic in a way that cannot be faulted? Or are there simply no numbers in arithmetic? Could the numerals help us to form signs for these second level concepts and yet not be signs in their own right? (*ibid*., p. 257) use of second-order quantifiers in forming predicates in which only first-order variables are free.[8]

Little as he says about truth, Frege says far less about the nature of falsity. A false thought is one that refers to the False. Introducing his symbol for the negation function in *Grundgesetze*, he tells us:

> We need no special sign to declare a truth-value to be the False, so long as we possess a sign by which either truth-value is changed into the other.[9]

The thought here is that a sentence (or the thought it expresses) is false if, and only if, its negation is true.[10]

2 ABOUTNESS

Frege's claim that sentences containing non-referring singular terms are neither true nor false follows, in his eyes, from what I think of as Frege's thesis about *aboutness*:

> If words are used in the ordinary way, what one intends to speak of is their reference. ('On sense and reference', p. 159)[11]

There is an unambiguous statement of this thesis in the notes Carnap took at Frege's lectures:

> A proper name has
>
> 1) a meaning [reference]: the thing about which something is said;
> 2) a sense that is part of the thought.[12]

[8] See Michael Dummett, *Frege: Philosophy of Mathematics* (London: Duckworth, 1991), pp. 214–22, for a detailed analysis.

[9] *Gg*, vol. I, §6.

[10] In §5 Frege has introduced the judgement stroke which indicates assertion, the acknowledgement of a thought as true. To say that no special sign is needed in the case of falsity is to say that acknowledgement of a thought as false can be effected by acknowledging its negation as true (and likewise the work done by denial can be accomplished by assertion of the negation).

[11] I have silently restored 'reference' and its cognates for the translation of '*Bedeutung*' and the like here and in subsequent quotations.

[12] E. Reck and S. Awodey (eds.), *Frege's Lectures on Logic: Carnap's Student Notes, 1910–1914* (Chicago: Open Court, 2004), p. 148.

Given that, there's nothing a sentence containing a bearerless name is about, hence nothing for what it predicates to be true or false of.

Take a sentence containing a Fregean proper name, i.e. either a proper name or a definite description. Replace all occurrences of said name or description by some place-holding marker, say 'ξ'. We have what Frege calls a (first-level) concept-word, an expression that refers to a concept. First-level concepts are functions from objects to truth-values. Going back to the original sentence, the Fregean proper name refers to an object, if it refers at all. If it refers, its reference is an argument of the function. If it refers, the sentence is true or false as the value of the function for that argument is the True or the False. If it does not refer, no argument is selected, hence the function takes no value: no input, no output.[13]

This is what leads Frege to maintain that sentences containing bearerless proper names are neither true nor false. In a discussion of the name 'Odysseus', which he holds may be bearerless, he says,

> Is it possible that a sentence as a whole has only a sense, but no reference? At any rate, one might expect that such sentences occur, just as there are parts of sentences having sense but no reference. And sentences which contain proper names without reference will be of this kind. The sentence 'Odysseus was set ashore at Ithaca while sound asleep' obviously has a sense. But since it is doubtful whether the name 'Odysseus', occurring therein, has reference, it is also doubtful whether the whole sentence has one. Yet it is certain, nevertheless, that anyone who seriously took the sentence to be true or false would ascribe to the name 'Odysseus' a reference, not merely a sense; for it is of the reference of the name that the predicate is affirmed or denied. Whoever does not admit the name has reference can neither apply nor withhold the predicate. ('On sense and reference', p. 162)

In one published article from 1897, and in posthumously published writings dated 1897 and 1914 by the editors (although neither date may be reliable), the connection made between 'aboutness' and lack of truth-value is rendered quite transparently, this time with the name 'Scylla':

> In poetry and legend … there occur sentences which, although they have a sense, have no reference – like, e.g., 'Scylla has six heads'. This sentence

[13] Cf. Susan Haack, *Philosophy of Logics* (Cambridge: Cambridge University Press, 1978), p. 212; and Scott Lehmann, 'Strict Fregean free logic', *Journal of Philosophical Logic*, 23 (1994), pp. 307–36.

is neither true nor false since, for it to be one or the other, it would have to have a reference; but no such reference is available, because the proper name 'Scylla' designates nothing.[14]

The sentence 'Scylla has six heads' is not true, but the sentence 'Scylla does not have six heads' is not true either; for it to be true the proper name 'Scylla' would have to designate something. ('Logic', pp. 129–30)

And when we say 'Scylla has 6 heads', what are we talking about? In this case nothing whatsoever; for the word 'Scylla' designates nothing. Nevertheless we can find a thought expressed by the sentence, and concede a sense to the word 'Scylla'.[15]

Likewise, back with Odysseus, in the 1906 diary entries that form 'Introduction to Logic':

Proper names are meant to designate objects, and we call the object designated by a proper name its reference. On the other hand, a proper name is a constituent of a sentence, which expresses a thought. Now, what has the object got to do with the thought? We have seen from the sentence 'Mont Blanc is over 4000 m high' that it is not part of the thought. Is then the object necessary at all for the sentence to express a thought? People certainly say that Odysseus is not an historical person, and mean by this contradictory expression that the name 'Odysseus' designates nothing, has no reference. But if we accept this, we do not on that account deny a thought-content to all the sentences of the *Odyssey* in which the name 'Odysseus' occurs. Let us just imagine that we have convinced ourselves, contrary to our former opinion, that the name 'Odysseus', as it occurs in the *Odyssey*, does designate a man after all. Would this mean that the sentences containing the name 'Odysseus' expressed different thoughts? I think not. The thoughts would strictly remain the same; they would only be transposed from the realm of fiction to that of truth. So the object designated by a proper name seems quite inessential to the thought-content of a sentence which contains it. To the thought-content! For the rest it goes without saying that it is by no means a matter of indifference to us whether we are operating in the realm of fiction or of truth.[16]

[14] 'On Mr Peano's conceptual notation and my own', *CP*, p. 241.

[15] 'Logic in mathematics', *PW*, p. 225.

[16] 'Introduction to logic', *PW*, p. 191. I cite this lengthy passage in full because it is so very much at odds with the reading given by Gareth Evans, *The Varieties of Reference* (Oxford: Oxford University Press, 1982), pp. 29–30) and by John McDowell ('Truth-value gaps', in McDowell, *Meaning, Knowledge, and Reality*

For Frege the truth-functional logical connectives are literally functions (or, if you think of connectives as linguistic, refer to functions). They are functions whose values are truth-values. For reasons we shall briefly examine below, Frege demands that functions defined for any objects be defined for all, so we cannot say that the connectives are functions whose arguments are truth-values; we can, however, say that it is only for arguments that are truth-values that we need to take note of the values they assign. Negation maps the True to the False and the False to the True; conjunction maps the pair <the True, the True> to the True, the pair <the True, the False> to the False, and so on. As a consequence of this understanding of the connectives, a sentence containing any sentential clause in a direct/non-oblique context that contains a bearerless proper name must lack a truth-value (must fail to refer).

3 THE LOGICAL PROBLEM OF BEARERLESS NAMES

Frege holds that *any* sentence containing a bearerless name in a direct/non-oblique context is neither true nor false. That is the completely general thesis advanced in the quotation from 'On sense and reference'. He terms the thought expressed by such a sentence 'fictitious' and a 'mock thought' ('Logic', p. 130); they are such exactly and only in that they fail to be about actually existing objects. In particular, he says

'Scylla has six heads' is not true

and

'Scylla does not have six heads' is not true.

Lack of a bearer for a singular term spreads lack of truth-value pervasively to logically complex sentences. What holds for negation applies equally to the other familiar connectives. We can set out the Fregean picture in what look like truth-tables for three-valued logic:[17]

(Cambridge, Mass.: Harvard University Press, 1999), pp. 212–13) to Frege's talk of 'mock proper names' and 'mock thoughts' in the 1897 piece 'Logic' as to encourage me in the belief that they have simply misread Frege. Cf. David Bell, 'How "Russellian" was Frege?', *Mind*, 99 (1990), pp. 267–77, §4.

[17] Cf. Timothy Smiley, 'Sense without denotation', *Analysis*, 20 (1960), pp. 125–35.

FREGEAN TRUTH TABLES

A	**¬A**
T	F
–	–
F	T

A&B (rows: **A**; columns: **B**)	T	–	F
T	T	–	F
–	–	–	–
F	F	–	F

A∨B (rows: **A**; columns: **B**)	T	–	F
T	T	–	T
–	–	–	–
F	T	–	F

A→B (rows: **A**; columns: **B**)	T	–	F
T	T	–	F
–	–	–	–
F	T	–	T

A↔B (rows: **A**; columns: **B**)	T	–	F
T	T	–	F
–	–	–	–
F	F	–	T

Beware! The bar is not a third truth-value; it signifies the absence of a truth-value. Where both **A** and **B** have truth-values, the connectives behave classically.

Lack of truth-value bothered Frege, his reason being that it subverts classical logic.[18] Going by the truth-tables above and taking for granted that a valid inference transmits truth from premises to conclusion, Frege was right to be bothered. Some familiar natural deduction rules fail:

> ∨-introduction, →-introduction (conditional proof), *reductio ad absurdum*, *ex falso quodlibet*, the law of excluded middle.

On the other hand, enough of classical logic survives for Frege to be in deep trouble, very deep trouble. Various classical equivalences still hold:

> $P \leftrightarrow Q$ and $\neg P \leftrightarrow \neg Q$; $\neg\neg P$ and P;

and we still have this rule:

> from $P \leftrightarrow Q$ and $Q \leftrightarrow R$ infer $P \leftrightarrow R$.

Albeit that numerous familiar rules fail to be uniformly truth-preserving, there is a simple criterion of validity in this setting:

> the inference from premises Σ to conclusion φ is valid iff (i) it is classically valid and (ii) no proper name occurs in φ that does not occur in at least one member of Σ.

[18] See, e.g., 'Function and concept', *CP*, p. 148; *Gg*, vol. II, §165.

There are two routes to trouble. The first adopts (and adapts) an argument due to Herbert Heidelberger.[19]

The indirect argument

With the valid equivalences and the rule noted above in play, we can do this:

> It's true that *P* if, and only if, *P*
> So, it's not true that *P* if, and only if, not-*P*
> But it's true that not-*P* if, and only if, not-*P*
> And it's true that not-*P* if, and only if, it's false that *P*
> Hence it's not true that *P* (if, and) only if it is false that *P*.

And:

> It's false that *P* if, and only if, it's true that not-*P*
> And it's true that not-*P* if, and only if, not-*P*
> So, it's not false that *P* if, and only if, not-not-*P*
> Thus, it's not false that *P* if, and only if, *P*
> But it's true that *P* if, and only if, *P*
> Hence it's not false that *P* (if, and) only if it is true that *P*.

Putting that all together we get,

> It's not true that *P* and it's not false that *P* only if it's both true that *P* and false that *P*.

In short, everywhere we think there's a truth-value gap, there's also a 'glut'! (And vice versa!)

In reaching this conclusion we have used a little logic and Frege's claim about the sameness of thought expressed by *P* and 'It's true that *P*'. Is the little logic used sound with respect to the Fregean truth-tables? Well, if *P* is neither true nor false the biconditionals above are all neither true nor false. But that's not really germane. What matters is that in asserting that a sentence containing a bearerless name is neither true nor false Frege surely intends to say

[19] Herbert Heidelberger, 'The indispensability of truth', *American Philosophical Quarterly*, 5 (1968), pp. 212–17.

something true: he asserts it, so, by his own lights has judged it to be true, not truth-value-less. Now, as the transitions licensed by the biconditionals above are truth-preserving (even if the biconditionals themselves are neither true nor false), we can indeed claim that there is a truth-preserving inference from the supposedly true

It's not true that *P* and it's not false that *P*

to the contradictory

It is both true that *P* and false that *P*.

The latter is certainly contradictory for it expresses the same thought as

P and not-*P*.

Heidelberger's argument is perhaps not as well known as it should be. It's not a knock-down argument that any theory that acknowledges truth-value gaps must acknowledge all instances of gaps as being simultaneously instances of truth-value gluts. One needs to know what logical principles are in play.[20] In Frege's case enough is in play to use at least a variant of the argument: the claim that a sentence containing a bearerless name is neither true nor false is contradictory, provably so even in a logic that allows for gaps as profligate as those of the Fregean truth-tables.

The direct argument

A step taken in the course of the indirect argument suffices to establish the incoherence of Frege's claims about sentences containing bearerless names and the thoughts they express. It is a step that

[20] There are non-standard logics in which biconditionals do not contrapose (see, e.g., Richard Holton, 'Minimalism and truth-value gaps', *Philosophical Studies*, 97 (2000), pp. 137–68, at pp. 154–5, for an application to present subject matter) and logics in which the negations of logically equivalent formulas need not be logically equivalent (such as Nelson's Logic of Constructive Falsity and Priest's Logic of Paradox).

The observation regarding biconditionals is something of a red herring. When we reconstruct the argument in terms of inferences, what matters is the rule of proof inversion (a weak form of *reductio ad absurdum*): if *A* entails *B* then not-*B* entails not-*A*. Now, true enough, this rule will not hold in general when bearerless names give rise to sentences that are neither true nor false, but the instances that we need do preserve truth-preservation.

Frege ought to have considered, for it turns on the answer to the simple question, what is the difference between it's not being true that *P* and not-*P*'s being true? For the Fregean there can be none.

(1) By the truth-equivalence, *P* and 'It's true that *P*' express the same thought.
(2) By the functional understanding of negation, not-*P* and 'It's not true that *P*' must therefore express the same thought.
(3) By the truth-equivalence, not-*P* and 'It's true that not-*P*' express the same thought.
(4) Therefore, 'It's not true that *P*' and 'It's true that not-*P*' express the same thought.

Crispin Wright says, '[T]he equivalence schema entails, given only the most basic assumptions about its scope and about the logic of negation, that truth and negation commute as prefixes'.[21] More narrowly, we have used only claims about sameness of meaning (thought expressed) to obtain the same conclusion. Frege wants 'It's not true that *P*' to be TRUE when *P* contains a bearerless name; and, at the same time, he wants 'It's true that *P*' to say the same as *P* (and 'It's true that not-*P*' to say the same as not-*P*) even though, in virtue of containing a bearerless name, *P* is, he wants to say, NOT TRUE and NOT FALSE. Now, the very fact that Frege tells us so little about falsity, and what he does tell us is exactly that judgement of a thought as false is accommodated by judging its negation to be true, shows us that he takes 'It is false that *P*' and 'It's true that not-*P*' as ways of expressing the same thought. But if he is right about this then he cannot coherently maintain of any thought that it is neither true nor false, for 'It's true that not-*P*' is entailed by, indeed *says the same as*, 'It's not true that *P*'.

The following constitute an inconsistent triad (which we may call 'Frege's trilemma'):

(i) The truth-equivalence
(ii) The functional reading of negation
(iii) The truth-value gap thesis concerning the thoughts expressed by sentences containing bearerless names.

[21] Wright, 'Truth', p. 213.

In the *Notes Dictated to G.E. Moore in Norway, April 1914* Wittgenstein states as a *definition*:

p is false = ~(*p* is true) Def.[22]

In a Tarskian, recursive definition of truth we standardly have the clause

~*p* is true if, and only if, *p* is not true.

Such stipulations threaten not just the functional understanding of negation but bring pressure to bear on the very foundation of Frege's function/argument analysis of propositions. For suppose *P* has the form *F*(*a*) where the name *a* does not refer. Then *P*'s negation has the form ¬*F*(*a*) and is true, since *P* is not true. But the function denoted by ¬*F*(ξ) does not, on its own, name either the True or the False, and yet *a* supplies no argument for it. To avoid this consequence, it would seem that one must give up the functional understanding of the logical constants.

Otherwise, one must give up either the truth-equivalence or deny the existence of truth-value gaps. Michael Dummett gives up the former.

A Dummettian interlude

We have seen that (i), (ii) and (iii) are inconsistent. Dummett has argued that (i) and (iii) are inconsistent. The argument is given originally in his article 'Truth'. It has been endorsed by many. Simon Blackburn and Keith Simmons, in the Introduction to their collection, *Truth*, rehearse it and wield it fiercely without further ado. Richard Holton has said of it that it is as damaging as it is simple. It's certainly simple. Here's the argument:

Suppose that *P* contains a singular term which has a sense but no reference: then, according to Frege, *P* expresses a proposition which has no truth-value. This proposition is therefore not true, and hence the statement

[22] Ludwig Wittgenstein, *Notebooks 1914–1916*, ed. G. H. von Wright and G. E. M. Anscombe, trans. G. E. M. Anscombe, 2nd edn (Oxford: Basil Blackwell, 1979), p. 116.

'It is true that *P*' will be false. *P* will therefore not have the same sense as 'It is true that *P*', since the latter is false while the former is not.[23]

As it stands this argument is hardly compelling. It is an argument in the logician's sense: it has premises; it has a conclusion. What connects them is a premise that Dummett has endorsed time and again: that 'It's true that …' is an *oratio obliqua* context,[24] an oblique, opaque or indirect context. 'It's true that *P*' is to be read as predicating truth of the thought that is the reference of 'That *P*'. There being no failures of reference in indirect contexts, 'It's true that *P*' cannot be neither true nor false.

Dummett admits that the context governed by 'It's true that …', unlike, say, propositional attitude contexts, fails the standard substitution test for opacity. That test, however, he takes as being only a sufficient criterion. What is at stake here is the way we should read 'It's true that *P*'. As I have said, Dummett reads it as predicating truth of the thought referred to by the name 'That *P*'. This is to be contrasted with how we read a sentence of the form 'It is not the case that *P*'. Here 'It's not the case that …' attaches, as an operator to the sentence *P*.[25] Why can we not read 'It's true that *P*' analogously (for surely 'It is the case that …' should be like *both* 'It is true that …' *and* like 'It is not the case that …')? Before we come to Dummett's response to that question let's ask another. Does Frege concur with Dummett's reading?

Nothing Frege says encourages the thought that he does. There are substantial reasons to think that he does not. Here's one. The passages quoted from 'On sense and reference' that give us theses (i) and (iii) occur before any mention of oblique contexts (more properly, of the customary/indirect distinction for sense and reference). It would be disingenuous in the extreme, not to say outright dishonest, of Frege to use a locution that requires that distinction for its

[23] Michael Dummett, 'Truth', in his *Truth and Other Enigmas* (London: Duckworth, 1978), pp. 1–24. at p. 4.

[24] E.g. *ibid*., Michael Dummett, *The Interpretation of Frege's Philosophy* (London: Duckworth, 1981), ch. 6, and 'Of what kind of a thing is truth a property?', in S. Blackburn and K. Simmons (eds.), *Truth* (Oxford: Oxford University Press, 1999), pp. 264–81.

[25] Cf. A. N. Prior, 'Oratio Obliqua', *Aristotelian Society Supplementary Volume*, 37 (1963), pp. 115–26, at p. 116, and 'Is the concept of referential opacity really necessary?', *Acta Philosophica Fennica*, 16 (1963), pp. 189–99, at pp. 193–4.

proper interpretation prior to advancing it, the more so as *nowhere* does he ever so much as mention it as an example giving rise to an indirect context. Here's another reason for thinking Frege didn't adopt Dummett's reading. Take a sentence such as 'Jena is a city on the Saale'. The reference of the name 'Jena' is a particular German city. If the sentence is true then so too is the sentence 'Jena exists'. Jena's existence is an existential commitment of that sentence's being true. On Dummett's reading of 'It's true that Jena is a city on the Saale', this second sentence has no such existential commitments, or at least has none such directly, because in this sentence the name 'Jena' now refers to the customary sense expressed by the name 'Jena', the sense expressed in the original sentence, its indirect reference. There is no obvious explanation why the truth of 'It's true that Jena is a city on the Saale' has the existential commitments of 'Jena is a city on the Saale'. Now, surely, Frege, had he intended Dummett's reading, would have realized this and, having realized it, balked at the sameness of thought claim.

Frege does not, I contend, concur with Dummett's reading. Should we? As I read him Dummett presents only one argument that is intended to clinch the claim that we should.[26] It is this:

[I]f there are meaningful sentences which say nothing which is true or false, then there must be *a* use of the word 'true' which applies to propositions; for

[26] In 'Of what kind of a thing is truth a property?', he offers another which he describes as providing a strong reason in favour of the *oratio obliqua* thesis. The argument form

> *X* believes that *P*
> It's true that *P*
> Therefore, *X* has a true belief

is, as Dummett puts it, unquestionably valid. It is also, as he says, unproblematic if we read 'It's true that *P*' as predicating truth of the proposition 'That *P*', i.e., the *oratio obliqua* reading. It may, Dummett concedes, be objected that the form

> *X* believes that *P*
> *P*
> Therefore, *X* has a true belief

is equally valid, but problematic on the view Dummett maintains. The difficulty can be 'localized' and validity explained by allowing inference of 'It's true that *P*' from *P*. But if this fact 'provides strong reason for construing the phrase "it is true that" as inducing an opaque context' (p. 271), it does so at the cost of rendering the validity of the latter inference wholly unexplained: to allow the inference is not to explain it.

if we say 'It is neither true nor false that *P*' the clause 'that *P*' must here be in *oratio oblique*, otherwise the whole sentence would lack a truth-value.[27]

What reason could there be to believe this conclusion? I suspect that Dummett is making an assumption that we have already seen to be false on pain of contradiction *when the truth-equivalence is accepted*: that if a sentence or thought *P* is neither true nor false so too is its negation. The conclusion of the direct argument shows us that, if *P* is neither true nor false, then, since in particular it is not true, its negation *is* true. Without an assumption to the contrary in play, I cannot see how Dummett's conclusion follows. To show that it does not we must elaborate a coherent position that admits that 'It is neither true nor false that *P*' is true for some sentences *P* consistently with the truth equivalence.

Before that, notice that, if Dummett is to avoid the same morass Frege got himself into, he must deny one of the following (what we might call 'Dummett's trilemma'):

(i) That *A* entails 'It's true that *A*'
(ii) That not-*B* entails not-*A* when *A* entails *B*
(iii) That 'It's true that not-*A*' entails 'It's false that *A*'.

Here I take entailment to be necessary truth-preservation. (i) and (ii) suffice to get from 'It's not true that *P*' to 'It's true than not-*P*'.

A semantic conception of falsity

The Fregean wants to say that an assertoric sentence may be neither true nor false. For this to be possible while endorsing the truth-equivalence, he must *not*, as we have seen, equate being false and not being true. If a sentence is not true that is either because it is false or because it is neither true nor false. Frege's account of falsity, such as it is, fails to allow for this second possibility.

Dummett has argued that the truth-value gap thesis is incompatible with Frege's claim that *P* and 'It's true that *P*' express the same thought/content. I have suggested that Dummett's grounds for this are not compelling. I do, nonetheless, hold that the truth-equivalence and the truth-value gap thesis are incompatible with Frege's conception of how sentences come to be false, i.e., how they come to name

[27] Dummett, 'Truth', p. 5.

the False. What we have been led to this far, via the direct argument, is an account of negation that assigns to it this truth-table:

A	**¬A**
T	F
–	T
F	T

This is a three-valued truth-table in a purely formal sense: as before, '–' stands for neither true nor false, not for some distinct, third *value*.

This is an unorthodox truth-table, one that follows from our Fregean theses (together with the commonplace that the negation of a truth is false), but one which there can be little doubt Frege failed to consider. Why so? Because *P*'s negation has a truth-value even when *P* doesn't. In Fregean terms, even when the assertoric sentence *P* fails to refer either to the True or the False, because containing a non-referring singular term, its negation, which contains exactly the same non-referring term or terms, succeeds in referring to the True. But the negation operator, for Frege, stands for a function, a function that maps the True to the False and everything else to the True (see, e.g., 'Function and concept', pp. 149–50, *Gg*, vol. I, §6). It must, as all functions must according to Frege, be defined for all objects, but the whole point of the passage about Odysseus is that the sentence containing a non-referring term fails to refer, hence can supply no argument for the function for which negation stands to act upon. There is, as we have seen, a fundamental incoherence in Frege's *use* of negation in 'On sense and reference'.

It is important to appreciate exactly which of Frege's semantic theses poses the problem. It is *not* the thought that sentences are proper names referring to the True and the False if to anything, nor the thought, however problematic elsewhere, that the True and the False are objects on a par with tables, chairs and extensions of concepts. Nor is it the thought that logically unstructured predicates refer to concepts and the latter are functions from objects to truth-values. Nor yet is it a mere commitment to compositionality, at least as that is understood in contemporary terms, for there is no failure of compositionality in having a negation satisfying the truth-table above. Where the problem lies is in the supposition that *any* sentence containing a non-referring singular term must itself express a thought that fails to have a reference, i.e. fails to be either true or false. That thesis, entirely plausible taken on its own, no doubt, is incompatible with the conjunction of the truth equivalence and the truth-value gap thesis. Frege holds to this thesis because he has a narrow reading of compositionality in functional terms, from which follows the principle 'no input, no output'.[28]

Holding to the truth-equivalence, we must give up the claim that there are sentences expressing thoughts that are neither true nor false, or the thesis that *any* assertoric sentence containing a non-referring singular term must express a thought that fails to be either true or false. The latter, of course entails the former, so that we cannot give the former up without rejecting the latter. The entailment does not reverse. What *is* the case is that if we give up on the specifically Fregean thesis that any sentence containing a non-referring singular term must itself express a thought that fails to be either true or false, it may seem that we have little reason to continue endorsing the claim that there are sentences expressing thoughts that are neither true nor false. Why should the Fregean continue to suppose some sentences containing non-referring terms are neither true nor false when having to give up on the claim that all are? The best answer, it seems to me, is that anyone with Fregean sympathies would have to give up on the stronger thesis anyway, irrespective of any problems occasioned by his/her treatment of negation.[29]

[28] Evans makes the weaker point that Frege had no means to rule out a 'wide scope' negation with our truth-table. But Evans makes a mistake when he goes on to say:

> I said this was essentially the same point, because it rests upon the incomprehensibility of the idea that the thought that *p* and the thought that *it is not true that p* can both fail to be true. Surely the thought that it is not true that *p* is true just when the thought that *p* is not true. So resistance to the idea that both thoughts may fail to be true is, once again, resistance to the idea of a gap between a determinate thought's failing to have the value True and its having the value False. (Evans, *The Varieties of Reference*, p. 25)

We agree that the thought that it is not true that *p* is true just when the thought that *p* is not true. But that, as we shall show, does not preclude the possibility of *p*'s being neither true nor false.

[29] We should note that the weaker claim is compatible with what Frege lays down in *Grundgesetze* in a section headed 'When does a name denote something?' There he says,

Our Fregean truth-tables for conjunction and disjunction have a quite extravagantly damaging effect on what Christine Tappolet calls truisms about truth.[30] To be fair, her one example – that a conjunction is true if, and only if, its conjuncts are true – survives; but the parallel truism for (inclusive) disjunction, that a disjunction is true if, and only if, at least one of its disjuncts is true, fails, as does the truism that a conjunction is false if, and only if, at least one of its conjuncts is false. Tappolet also proposes it as a truism that truth is what is conserved in valid inference. If so, with these truth-tables the natural deduction rule of ∨-introduction fails, as we noted above. If one supposes that, likewise, that it is a truism that if the conclusion of a valid inference is false so too is at least one premise – and is this any less of a truism? Perhaps it is – then the natural deduction rule of &-elimination (simplification) also fails.

Now, of course, Frege did think that in order to save logic, by which he meant classical logic, a logically perfect language must satisfy the requirement that there be no non-referring singular terms, and that ordinary language is sadly deficient in this respect. Anyone less sanguine than Frege about a wholesale revision of everyday conceptions in favour of the logically perfect, will, I suspect, feel moved to hold on to Tappolet's truisms and so reject the Fregean truth-tables for the sentences of everyday language. Rather, she will endorse these truth-tables:

> A proper name has a denotation if the proper name that results from that proper name's filling the argument places of a denoting name of a first-level function of one argument *always* has a denotation, and if the name of a first-level function of one argument that results from the proper name in question's filling the ξ-argument places of a denoting name of a first-level function of two arguments *always* has a denotation, and if the same holds also for the ζ-argument-places. (*Gg*, vol. I, §29, emphasis added)

He has preceded this by saying:

> A name of a first-level function of one argument has a denotation (denotes something, succeeds in denoting) if the proper name that results from this function-name by its argument-places being filled by a proper name always has a denotation if the name substituted names something.

[30] Christine Tappolet, 'Truth pluralism and many-valued logics: A reply to Beall', *Philosophical Quarterly*, 50 (2000), pp. 382–5.

NON-FREGEAN TRUISTIC TRUTH TABLES

A	**¬A**
T	F
–	T
F	T

A&B (A \\ B)	T	–	F
T	T	–	F
–	–	–	F
F	F	F	F

A∨B (A \\ B)	T	–	F
T	T	T	T
–	T	–	–
F	T	–	F

A→B (A \\ B)	T	–	F
T	T	–	F
–	T	?	–
F	T	T	T

A↔B (A \\ B)	T	–	F
T	T	–	F
–	–	?	–
F	F	–	T

I put a '?' rather than '–' because one may well maintain that *any* instance of $A \to A$ should be TRUE, but equally clearly not every sentence of the form $A \to B$ in which both A and B contain non-referring proper names should be TRUE. Arguably it's a truism that if A entails B then 'If A then B' is true; and even more arguably, it's a truism that A entails A. Likewise, if A and B say the same thing, 'If A then B' should be true.[31]

At the same time as endorsing these truth-tables (which merely enshrine truisms),[32] we don't have to give up entirely on the original Fregean perception that leads to the thesis that sentences containing non-referring singular terms are neither true nor false.

Think of simple predications. A predicate refers to a concept; some objects fall under the concept, some don't. A predicate therefore just refers to something that maps objects to the semantic values of sentences, which, for Frege, are truth-values (as Dummett has emphasized: e.g., chapter 6 of *The Interpretation of Frege's Philosophy*). And we need feel no obligation to say that a sentence comprising

[31] Regarding truisms, it has to be admitted that in twentieth-century logic nothing is sacred. Quantum logic has been seen as admitting true disjunctions neither of whose disjuncts need be true. Even the rule of &-elimination (simplification) has been denied: a conjunction may not entail its conjuncts (see, e.g., Robert Gahringer, 'Intensional conjunction', *Mind*, 79 (1970), pp. 259–60, and Bruce Thompson, 'Why is conjunctive simplification invalid?', *Notre Dame Journal of Formal Logic*, 32 (1991), pp. 248–54).

[32] In the case of the conditional, conditionally so: if there's any truth in the material implication account of the conditional.

a simple predication and a non-referring singular term is anything other than neither true nor false.

We must divorce falsity and non-truth. We do this for atomic sentences: for an atomic sentence to be false it must be not true *and* all the singular terms it contains refer. Otherwise it is neither true nor false. Rather than an analogue of the simple truth equivalence, we then proceed to give a recursive definition of falsity, guided by the platitudes enshrined in the non-Fregean truth-tables. We must also give an account of the quantifiers. The reader interested in seeing how this goes may consult the Appendix.

Now, it is a fact to be celebrated that, if we take Tappolet's truism that truth is what is preserved in valid inferences to heart, we find that the (formal) truth-tables above for negation, conjunction and disjunction deliver that all and only classically valid propositional logic inference patterns involving those connectives preserve truth. Employing classically valid inference patterns in conjunction with the truth equivalence, we can then *derive* the standard recursive clauses for truth concerning these connectives – accepting that 'if *A* then *B*' is true when *A* entails *B*. What is more, we also have that an atomic sentence is neither true nor false when, and only when, at least one of its terms fails to refer. And, returning to Dummett's discussion of 'It's true that …' contexts, we find that not only may the thought expressed by such a sentence be neither true nor false, it is then true that it is neither true nor false.

4 STATEMENTS OF NON-EXISTENCE

To paraphrase Leonard Linsky, Frege does not address the analysis of negative existential sentences involving proper names, but we can construct a Fregean account of them.[33] Existence is a higher-level concept, under which concepts of lower level fall, or not: to say that there is at least one φ is to say that at least one object falls under the concept that is the reference of 'φ'. To say that Jena exists is to say that at least one object falls under the concept *identical to Jena*; formally, $\exists x(x = a)$. To say that unicorns don't exist is to say that nothing falls under the concept *unicorn*; formally $\neg\exists x\phi x$. Likewise, one

[33] Leonard Linsky, *Names and Descriptions* (Chicago: University of Chicago Press, 1977), p. 5.

might expect, to say that Pegasus does not exist is to say that nothing falls under the concept *identical to Pegasus*. But this suggestion falls foul of Frege's general thesis concerning sentences containing non-referring terms: the thought expressed, that Pegasus does not exist, is neither true nor false on Frege's account.[34]

It has struck some as intolerable that this should be a consequence of Frege's theses on non-referring singular terms. Linsky, for example, says that 'Pegasus does not exist' is true, a truth, indeed, that we must insist upon.[35] Now, while common sense balks at Meinong's 'There are objects of which it is true to say that there are no such objects', it happily countenances assertions such as 'There are lots of things that don't exist: Atlantis, El Dorado, the planet Vulcan, Santa Claus, the Abominable Snowman, the Loch Ness Monster, the Big Grey Man of Ben Macdui …'[36] But be that as it may, it hardly sanctions a reflective insistence on the correctness of saying 'Pegasus does not exist'. Of what is one denying existence? Not of Pegasus, for, as one would like to say, the point is that there is no Pegasus of which existence may be denied. There's a real and pressing sense in which the sentence cannot be about Pegasus (even if, in some more attenuated sense of 'about', it is about Pegasus). The name 'Pegasus' does not refer; it expresses a sense but nothing answers to this mode of presentation.

To say that Pegasus does not exist may, in some indirect way, be a claim about the name 'Pegasus' or the sense it expresses. The former is perhaps Frege's view. The latter is Linsky's contention. In negative existential sentences 'exists' 'induces an oblique context in which the proper name denotes its customary sense'.[37] This is, he says, 'a rather satisfying result since it exploits the intuition that existence-contexts are indeed special, and that what prevents

[34] We have found Frege's own account to be thoroughly unsatisfactory. We could with good conscience argue thus: as 'Pegasus' does not refer, 'There is at least one object identical to Pegasus' is not true; therefore its negation is true. For the time being I ask the reader to suspend her dissatisfaction. Dissatisfaction suspended, we note that the Fregean must say that the predicate 'identical to Pegasus' expresses a sense but fails to refer: no concept answers to the sense expressed.

[35] Leonard Linsky, 'Frege and Russell on vacuous singular terms', in M. Schirn (ed.), *Studien zu Frege/Studies on Frege*, vol. III (Stuttgart: Frommann, 1976), pp. 97–115, at p. 112.

[36] If to be is to be the value of a bound variable, existence comes cheap in ordinary usage.

[37] Linsky, *Names and Descriptions*, p. 6.

"Pegasus does not exist" from being meaningless is the fact that the denotationless name "Pegasus" is not devoid of sense'.[38]

Whatever the merits of Linsky's proposal, it has two significant demerits. The first is that in classical logic the sentence ϕa containing the singular term a is logically equivalent to $\neg\exists x(x = a \,\&\, \neg\phi x)$. So either we junk classical quantification theory or thoughts about objects have the same truth-conditions (which may, for the Fregean, mean that they express the same thought)[39] as sentences about the modes of presentation of those objects. Neither is a happy place to end up.

The second demerit is that it fails to take account explicitly of what Frege says about existence presuppositions. In 'On sense and reference' we find,

If anything is asserted there is always an obvious presupposition that the simple or compound proper names used have a reference. If therefore one asserts 'Kepler died in misery', there is a presupposition that the name 'Kepler' designates something; but it does not follow that the sense of

[38] Linsky, 'Frege and Russell', p. 112. Linsky does not restrict the induction of oblique contexts by 'exists' to negative existentials. Some support for Linsky's position might be drawn from Anthony Kenny's observations:

> If a man uses a proper name, then he implies that it has a bearer, that is to say, that the object which he means exists. If someone says 'Satan exists' or 'Satan does not exist' then he does not imply, but respectively asserts or denies, that Satan exists. It follows that he is not using 'Satan' in these sentences as a proper name … [W]hether Satan exists or not, 'Satan' is not used as a proper name either in 'Satan exists' or in 'Satan does not exist'. Neither of these sentences, moreover, is about Satan, whether or not he exists. (Anthony Kenny, : 'Oratio Obliqua', *Aristotelian Society Supplementary Volume*, 37 (1963), pp. 127–146, at p. 141.

But why suppose it is existence statements that are special, why not the embedded identity? It seemed to Bas van Fraassen that 'we cannot plausibly reject that '$t = t'$' is false when t has a referent and t' does not.' He offers the example: that Santa Claus does not exist is sufficient reason to conclude that the president of the US is not Santa Claus ('Singular terms, truth-value gaps, and free logic', *Journal of Philosophy*, 63 (1966), pp. 481–95).

[39] See *Gg*, §32; 'Compound thoughts', *CP*, pp. 393 and 405; Letter to Husserl, 9 December, 1906, *PMC*, pp. 70–1; 'A brief survey of my logical doctrines', *PW*, p. 70; but see Charles Parsons, 'Review article: Gottlob Frege *Wissenschaftlicher Briefswechsel*', *Synthese*, 52 (1982), pp. 325–43, at pp. 328–9, and Jean van Heijenoort, 'Frege on sense identity', *Journal of Philosophical Logic*, 6 (1977), pp. 103–8.

the sentence 'Kepler died in misery' contains the thought that the name 'Kepler' designates something. If this were the case the negation would have to run not

> Kepler did not die in misery

but

> Kepler did not die in misery, or the name 'Kepler' has no reference.
>
> That the name 'Kepler' designates something is just as much a presupposition for the assertion
>
> Kepler died in misery

as for the contrary assertion. Now languages have the fault of containing expressions which fail to designate an object (although the grammatical form seems to qualify them for that purpose) because the truth of some sentence is a prerequisite. ('On sense and reference', pp. 168–9)

It is, then, a presupposition of the sentence 'Pegasus does not exist' that the name 'Pegasus' designate something, and hence that 'Pegasus exists' is true. Thus the sentence 'Pegasus does not exist' cannot be (truthfully) asserted: one cannot acknowledge it as true for in the very attempt to do so one must accept its contrary. Frege's observation concerning negations of sentences containing singular terms is correct to the extent that we do not normally make explicit existence assumptions when we negate a sentence. (Russell's theory of descriptions is a case in which the distinction between wide and narrow scope negations is in this respect abnormal.)

Taking the observation at face value, Frege has a strong motive for insisting, as he does on numerous occasions, that in a perfected scientific language every properly constructed (closed) singular term refers, for only then is it the case that the language is used in accordance with the presuppositions of (its) correct usage. We can see from the way we say what the negation (or, more generally, the contrary) of a given sentence says that we take for granted that the proper names of our language refer. It is, then, no surprise if we run into difficulties when those presuppositions fail. Frege, never being one to cast doubt on the correctness of classical logic, takes the presuppositions as well founded. A language rid of imperfections must accord with those presuppositions: well-formed proper names of the language must – as a matter of logic! – refer.

5 THE PERFECTED LANGAUGE OF A DEMONSTRATIVE SCIENCE

If we are to use a language in accord with the presuppositions for its use, all singular terms of the langauge must refer:

A logically perfect language (*Begriffsschrift*) should satisfy the conditions, that every expression grammatically well constructed as a proper name out of signs already introduced shall in fact designate an object, and that no new sign shall be introduced as a proper name without being secured a reference. ('On sense and reference', p. 169)

As Frege indicates here, there are general methods for forming proper names; all such names must be assured of a reference. The best known *variable-binding term-forming operators* (vbtos) generate definite descriptions and set abstracts. Let α be a vbto. Then, for any predicate $\phi(x)$, $\alpha x\phi(x)$ is a singular term (in Frege's terminology: a proper name).

In *Grundgesetze*, terms for courses-of-values are introduced this way (in the notation $\grave{\varepsilon}\Phi\varepsilon$) (*Gg*, §9).[40] What Frege calls his 'substitute for the definite article' is introduced a little differently, as a function mapping objects to objects:

if $\xi = \grave{\varepsilon}(\varepsilon = \Delta)$, for some object Δ then $\backslash\xi = \Delta$;
otherwise $\backslash\xi = \xi$. (*Gg*, §§11 and 31)[41]

The only objects Frege has previously introduced are the two truth-values, the True and the False, and courses-of-values. In §10 he has argued that the True and the False may be identified with courses-of-values. The net effect is therefore the same as introducing a vbto meeting these constraints:

[40] It is tempting to put '$\{x: \varphi(x)\}$', in a more modern notation, for Frege's '$\grave{\varepsilon}\Phi\varepsilon$' but that, while largely harmless, would be misleading in that Frege's extension of a concept is closer to the graph of a set's characteristic function than to the set itself.

[41] Frege's Basic Law VI says only: $\forall x(x = \backslash\grave{\varepsilon}(\varepsilon = x))$. That is, the Law says *nothing* about how $\backslash\xi$ is to be interpreted when ξ is not of the form $\grave{\varepsilon}(\varepsilon = \Delta)$ for some properly formed name Δ.

if a unique object falls under the concept $\phi(\xi)$ then '$\grave{\ }x\phi(x)$' denotes that object;

otherwise '$\grave{\ }x\phi(x)$' denotes $\grave{\varepsilon}\varphi\varepsilon$.

From the Fregean perspective, vbtos are second-level functions mapping concepts to objects. Like all functions, they must be defined for all possible arguments – all concepts – and they must be well defined, that is, the function must assign the same object to co-extensive concepts, for co-extension is the analogue for concepts of identity between objects. In the Fregean scheme, the vbto α satisfies these two axioms, where we read the second-order quantifiers as quantifying over concepts:

α-Existence: $\forall X \exists z(z = \alpha y Xy)$;

α-Extensionality: $\forall X \forall Y(\forall x(Xx \leftrightarrow Yx) \rightarrow \alpha y Xy = \alpha y Yy)$.

We should note that these are trivially consistent with respect to standard second-order semantics, for we may take a one-element domain, $D = \{0\}$, and a function $a{:}\wp(D) \rightarrow D$ which assigns 0 to both subsets of D. We should note too that, as George Boolos observed, α-Extensionality may be considered a logical truth: given extensional semantics, α is interpreted as a function from subsets of the domain to the domain and $\forall x(Xx \leftrightarrow Yx)$ is satisfied under an assignment of values to the second-order variables if, and only if, X and Y are assigned the same subset of the domain.[42] Lastly, we should note that in standard second-order logic α-Existence is a consequence of α-Extensionality as $\forall X \forall x(Xx \leftrightarrow Xx)$ is a logical truth.

6 BASIC LAW V

Introducing courses-of-values, in §3 of *Grundgesetze*, Frege says

I use the words

'the function $\Phi(\xi)$ has the same *course-of-values* as the function $\Psi(\xi)$'

generally to denote the same as the words

'the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument'.

[42] George Boolos, 'Frege's theorem and the Peano postulates', *The Bulletin of Symbolic Logic*, 1 (1995), pp. 317–26, at p. 322.

This is Basic Law V informally stated. There is further discussion of courses-of-values in §§9 and 10, and a formal statement in §20. Later on he splits the law into Va and Vb. Va is α-Extensionality for the extension-of-concept vbto $\grave{\varepsilon}$. It is unexceptionable. As Montgomery Furth says, 'This is no news to us; it merely follows from the extensionality of concepts'.[43] Vb is the converse of Va.

Converse of α-Extensionality: $\forall X \forall Y(\alpha y Xy = \alpha y Yy \rightarrow \forall x(Xx \leftrightarrow Yx))$.

This says that the vbto α stands for a one–one function from concepts to objects.

7. CANTOR'S THEOREM

In the 1890–1 volume, the first volume, of the *Jahresbericht der deutschen Mathematiker-Vereinigung* Cantor published a proof of what is now widely known[44] as Cantor's Theorem: every set is of lower cardinality than the set of its subsets. Emphasizing its pertinence to Fregean concerns, we note that Cantor's original proof was phased not directly in terms of subsets but in terms of functions defined on a given set and taking only two values (0 and 1).[45]

Cantor's theorem can be phrased in two equivalent ways:

(1) There is no function from a set X onto the set of all its subsets.
(2) There is no one–one function from the set of all subsets of X into the set X.

The orthodox textbook proof of Cantor's Theorem, following Cantor's original, is a proof of (1).

43 Montgomery Furth, 'Editor's Introduction', in Frege, *The Basic Laws of Arithmetic; Exposition of the System*, ed. and trans. M. Furth (Berkeley and Los Angeles: University of California Press, 1964), pp. xlv–xlvi.

44 Ernst Zermelo, 'Untersuchungen über die Grundlagen der Mengenlehre I', *Mathematische Annalen*, 65 (1908), pp. 261–81, translated as 'Investigations in the foundations of set theory I' in J. van Heijenoort (ed.), *From Frege to Gödel: A Source Book in Mathematical Logic, 1879–1931* (Cambridge, Mass.: Harvard University Press, 1967), pp. 199–215.

45 Cantor's proof is reproduced in Michael Hallett, *Cantorian Set Theory and Limitation of Size* (Oxford: Oxford University Press, 1984), p. 77.

> Let f be a function from X into the set of all subsets of X. The set $Y = \{x \in X: x \notin f(x)\}$ is a well-defined subset of X. But no element of X is mapped to Y by f. For suppose, to the contrary, that $f(y) = Y$; we ask, Does y belong to $f(y)$? If $y \in f(y)$, *i.e.* if $y \in Y$, then, by the definition of Y, $y \notin f(y)$; conversely, however, if $y \notin f(y)$ then y meets the defining condition for membership of Y and so $y \in Y$, i.e. $y \in f(y)$. Thus $y \in f(y)$ if, and only if, $y \notin f(y)$ – a contradiction.

With a proof of (1) in hand, (2) may quickly be derived via a proof by contradiction.

> Suppose that there is a one–one function h from $\wp(X)$ into X. As it is one–one it has an inverse: for any x in X, there is at most one subset of X mapped to x by h. The inverse maps some not necessarily proper subset Y of X *onto* $\wp(X)$. Pick an arbitrary subset of X and map any remaining members of X – those not in Y – to that subset. We then have a function from X onto $\wp(X)$, in contradiction to (1).

For present purposes it is of more interest to prove (2) directly. We do this twice over.

> Let g be a function from $\wp(X)$ into X. Let Y be the subset $\{x \in X: \exists Z \subseteq X\ [x = g(Z) \text{ and } x \notin Z\ ]\}$, a well-defined subset of X. Let $y = g(Y)$.
>
> I_A If $y \notin Y$ then y satisfies the condition for membership of Y, *i.e.* $y \in Y$. This suffices to establish that $y \in Y$.
>
> I_B But now, think what this, *i.e.* $y \in Y$, says. On the one hand, $y = g(Y)$ and $y \in Y$. On the other, for some subset Z of X, $y = g(Z)$ and $y \notin Z$. Thus g is not one-one as two distinct subsets of X are mapped to y.
>
> II_A Suppose that g is one–one and that $y \in Y$. For some Z, $y = g(Z)$ and $y \notin Z$. But as g is one-one, $Z = Y$. So, if $y \in Y$ then $y \notin Y$. This establishes that $y \notin Y$.
>
> II_B But think what this says: for any subset Z of X that gets mapped to y by g, y is a member of Z. But Y is one such and y, as just demonstrated, is not a member of it. Contradiction. – g cannot be one–one.

We have here *three* proofs: $I_A + I_B$; $II_A + II_B$; and $I_A + II_A$, for the latter two give us, respectively, $y \in Y$ and $y \notin Y$, a contradiction, on the assumption that g is one–one.

To bring out the role of the clause 'Y is a well-defined subset of X', here's a neat little exercise in set theory:

> The co-finite subsets of $\mathbb{N}$ are those subsets of $\mathbb{N}$ whose complements with respect to $\mathbb{N}$ are finite. The set of all finite and co-finite subsets of $\mathbb{N}$ is countably infinite. Let $X_0, X_1, \ldots, X_n, \ldots$ be some enumeration of this set. Show that the set
>
> $Y = \{n \in \mathbb{N}: n \in X_n\}$
>
> is neither finite nor co-finite.
>
> PROOF Consider the complement of Y, $\mathbb{N} - Y = \{n \in \mathbb{N}: n \notin X_n\}$, which is finite or co-finite as Y is co-finite or finite. If it is either finite or co-finite then $\mathbb{N} - \mathrm{Y} = X_m$, for some $m \in \mathbb{N}$. But then, $m \in X_m$ if and only if $m \in \mathbb{N} - \mathrm{Y}$ if and only if $\mathrm{m} \notin X_m$.

So $\mathbb{N} - \mathrm{Y}$, and hence Y itself, is neither finite nor co-finite.

The point is that here we can have a one–one correspondence between $\mathbb{N}$ and the family of finite and co-finite subsets of $\mathbb{N}$, exactly because the 'diagonalizing set' $\mathbb{N} - Y$ doesn't belong to that family.

Cantor's Theorem, published at around the time Frege was finishing the writing of the first volume of *Grundgesetze*, provides a stark warning. Sadly, it was a warning to which Frege was blind.

8 THE PARADOX

With a vbto α in our (second-order) language, we may form the predicate containing one free variable

$$\exists \mathrm{X}(\mathrm{x} = \alpha y Xy \; \& \; \neg Xx),$$

which we shall abbreviate as $\Psi(x)$. On the assumption that it is a suitable substituend for the second-order quantifiers, i.e. on the assumption that this predicate does refer to a concept, we can show, as a matter of logic, that $\Psi(\alpha y \Psi(y))$. We shall conduct the proof in a weak second-order free logic, so that we explicitly mark assumptions that a formula is an appropriate substitution instance for a quantifier – semantically, that a formula does refer to an entity falling within the range of the quantifiers, be that at first- or second-order.[46] We take the natural deduction rules for quantifiers and identity in free logic from Tennant.[47] The proof looks like this:

$$\dfrac{\dfrac{\forall X\exists z(z=\alpha yXy) \quad \exists!\psi}{\exists z(z=\alpha y\psi(y))}\,\forall_2\text{-E} \qquad \dfrac{\dfrac{}{t=\alpha y\psi(y)}\,1 \quad \dfrac{}{\exists!t}\,1}{\exists!\alpha y\psi(y)}\,=\text{-E}}{\exists!\alpha y\psi(y)}\,1\ \exists_1\text{-E}$$

$$\dfrac{\dfrac{\dfrac{\exists!\alpha y\psi(y)}{\alpha y\psi(y)=\alpha y\psi(y)}\,=\text{-I} \qquad \dfrac{}{\neg\psi(\alpha y\psi(y))}\,2}{\alpha y\psi(y)=\alpha y\psi(y)\ \&\ \neg\psi(\alpha y\psi(y))}\,\&\text{-I} \qquad \exists\psi}{\exists X(\alpha y\psi(y)=\alpha yXy\ \&\ \neg X(\alpha y\psi(y)))}\,\exists_2\text{-I}$$

$$\text{========================= Definition}$$

$$\dfrac{\psi(\alpha y\psi(y))}{\psi(\alpha y\psi(y))}\,2\ \text{CRA}$$

From α-Existence and the referential assumption $\exists!\Psi$ – that the predicate $\exists X(x = \alpha yXy\ \&\ \neg Xx)$ refers to a concept – we have derived $\Psi(\alpha x\Psi(x))$. Only the last step, an application of a weak form of classical *reductio ad absurdum*, is essentially classical.

This proof is a natural deduction free-logic variation on the proof Frege himself gives in the Appendix to Volume II of *Grundgesetze*. He then goes on to parallel step I_B above. He summarizes the result:

[46] Stewart Shapiro and Alan Weir, '"Neo-Logicist" logic is not innocent', *Philosophia Mathematica*, 8 (2000), §§IV and V, pp. 160–89, use free logic at first order but not second.

[47] Neil Tennant, *Natural Logic* (Edinburgh: Edinburgh University Press, 1978), pp. 167–8. Tennant takes (the first-order) $\exists!a$ as an abbreviation for $\exists x(x = a)$. One need not, for the rules for the quantifiers and identity allow one to prove their logical equivalence. In the present setting, in which we have taken over the rules at second order too, and in which the significance of '$\exists!\varphi$' has yet to be fully worked out, it is best that we take $\exists!$ as primitive.

In other words: for every second-level function of one argument of type 2 there are two concepts such that, taken as arguments of this function, they determine the same value, but also such that this value does fall under the first concept and does not fall under the second. (*Gg* II, Appendix)

As Frege realizes only too well, the result is quite general: it holds for any vbto α in the extensional/Fregean framework.[48]

We can give a formal version of the proof II_A. It's more involved – see next page – but, as we shall see shortly, there is a point to considering the $I_A + II_A$ proof of Cantor's Theorem.

Making explicit the assumption that the predicate $\exists X(x = \alpha y Xy \mathbin{\&} \neg Xx)$, which we have abbreviated $\Psi(x)$, refers to a concept, what the formal proofs give us is:

(i) α-Existence + $\exists!\Psi \vdash \Psi(\alpha x \Psi(x))$;

(ii) α-Existence + Converse of α-Extensionality + $\exists!\Psi \vdash \neg\Psi(\alpha x \Psi(x))$.

Hence the combination of α-Existence, Converse of α-Extensionality, and the assumption that the predicate $\exists X(x = \alpha y Xy \mathbin{\&} \neg Xx)$ refers to a concept is inconsistent.[49] Moreover, in view of the publication of Cantor's Theorem in 1890, this inconsistency was foreseeable.[50]

[48] Famously, in 'On concept and object', Frege tells us that 'the concept *horse*', being complete or saturated, names an object, not a concept. The expression 'The concept' is therefore a vbto. It requires rather strong Fregean nerves then to countenance that there must be two concepts, $\varphi(\xi)$ and $\chi(\xi)$, such that

the concept φ = the concept χ

but the object that we may variously refer to as 'the concept φ' or 'the concept χ' falls under one of these concepts but not the other. If, as seems plausible, we take it that 'the concept φ' names the extension of the concept φ – see Tyler Burge, 'Frege on extensions of concepts, from 1884 to 1903', in Burge, *Truth, Thought, Reason: Essays on Frege* (Oxford: Clarendon Press, 2005), pp. 273–98, at pp. 283–4 – this is, of course, merely a restatement of the application that led Frege to rethink extensions in the Appendix to volume II of *Grundgesetze*.

[49] The use of second-order quantification in obtaining the contradiction is essential. It is known that the (standard, hence also the free) first-order fragment of the system of *Grundgesetze* is consistent. See Terence Parsons, 'On the consistency of the first-order portion of Frege's logical system', *Notre Dame Journal of Formal Logic*, 28 (1987), pp. 161–8.

[50] As J. N. Crossley points out ('A note on Cantor's Theorem and Russell's paradox', *Australasian Journal of Philosophy*, 51 (1973), pp. 70–1, at p. 71), a derivation

$$\dfrac{\dfrac{}{\exists!\phi}\,2 \qquad \dfrac{\forall X\,\forall Y(\alpha yXy = \alpha yYy \rightarrow \forall x(Xx \leftrightarrow Yx))\quad \exists!\psi}{\forall Y(\alpha y\,\psi(y) = \alpha yYy \rightarrow \forall x(\psi(x) \leftrightarrow Yx))}}{\alpha y\,\psi(y) = \alpha y\phi(y) \rightarrow \forall x(\psi(x) \leftrightarrow \phi(x))}$$

$$\dfrac{\dfrac{}{\alpha y\,\psi(y) = \alpha y\,\phi(y)\ \&\ \neg\phi(\alpha y\,\psi(y))}\,2}{\alpha y\,\psi(y) = \alpha y\,\phi(y)}$$

$$\dfrac{\alpha y\,\psi(y) = \alpha y\phi(y) \rightarrow \forall x(\psi(x) \leftrightarrow \phi(x)) \qquad \alpha y\,\psi(y) = \alpha y\,\phi(y)}{\forall x(\psi(x) \leftrightarrow \phi(x))}$$

$$\dfrac{\dfrac{\forall X\exists z(z = \alpha yXy)\quad \exists!\psi}{\exists z(z = \alpha y\,\psi(y))} \qquad \dfrac{\dfrac{}{t = \alpha y\,\psi(y)}\,1 \quad \dfrac{}{\exists!t}\,1}{\exists!\alpha y\,\psi(y)}}{\exists!\alpha y\,\psi(y)}\,1$$

$$\dfrac{\forall x(\psi(x) \leftrightarrow \phi(x)) \qquad \exists!\alpha y\,\psi(y)}{\psi(\alpha y\psi(y)) \leftrightarrow \phi(\alpha y\,\psi(y))}$$

$$\dfrac{\dfrac{}{\alpha y\,\psi(y) = \alpha y\,\phi(y)\ \&\ \neg\phi(\alpha y\,\psi(y))}\,2}{\neg\phi(\alpha y\,\psi(y))}$$

$$\dfrac{\psi(\alpha y\psi(y)) \leftrightarrow \phi(\alpha y\,\psi(y)) \qquad \neg\phi(\alpha y\,\psi(y))}{\neg\psi(\alpha y\psi(y))}$$

$$\dfrac{\dfrac{}{\psi(\alpha y\,\psi(y))}\,3}{\exists X(\alpha x\,\psi(x) = \alpha xXx\ \ \&\ \neg X(\alpha x\,\psi(x)))}\ \text{(=====) Def'n}$$

$$\dfrac{\exists X(\alpha x\,\psi(x) = \alpha xXx\ \ \&\ \neg X(\alpha x\,\psi(x))) \qquad \neg\psi(\alpha y\psi(y))}{\neg\psi(\alpha y\,\psi(y))}\,2$$

$$\dfrac{\neg\psi(\alpha y\,\psi(y))}{\neg\psi(\alpha y\,\psi(y))}\,3$$

Proof that $\forall X\exists z(z = \alpha xXx)$, $\forall X\,\forall Y(\alpha yXy = \alpha yYy \rightarrow \forall x(Xx \leftrightarrow Yx))$, $\exists!\psi \vdash \neg\psi(\alpha y\,\psi(y))$,
where $\psi(x)$ abbreviates $\exists X\,(x = \alpha yXy\ \&\ \neg Xx)$.

Foreseeable, not foreseen. John Burgess offers this account:

The explanation is *not* that Frege rejected Cantor's results. A sufficient explanation is that Frege (like so many others) was largely *unaware of the bearing* of Cantor's cardinality theorems on the issues that concerned him. If he had pondered that bearing, he would surely have begun by translating Cantorian jargon into Fregean jargon. He would then immediately have seen that the Cantorian greater cardinality theorem says that there are more Fregean 'concepts' than Fregean 'objects'. He would then immediately have seen that this contradicts an axiom of the Fregean system, according to which there is a distinct 'object' associated with each 'concept,' namely, the 'class' that is its 'extension.' He would then surely have gone on to ponder whether or not the Cantorian proof can be reproduced within the Fregean system. He would then surely have seen that it can, and would thus have seen that his system is inconsistent.[51]

What adds pathos is that in his review of Frege's *Grundlagen* in 1885, Cantor had warned against taking extensions of concepts as the building blocks. Cantor already held then that there could be no set of all sets. Opinions divide on whether Cantor's warning was obscurely put or Frege simply negligent in, apparently, failing to understand it.[52]

of Russell's paradox is easily obtained from Cantor's proof of his theorem: if we take the domain of all sets to be, itself, a set, then, for that set V, we must have that $\wp(V)$ is V itself, for, on the one hand, all sets are contained in V, and, on the other, every set is a subset of V (as all its members are sets), and hence that the identity function is, *per impossibile*, a function from V onto the set of all its subsets. The set one then constructs in the course of Cantor's proof that there is no such function is the set $\{x \in V: x \notin x\}$, the Russell set. But by 1890 Cantor knew that the collection of all sets was, in the terminology he would later use, an absolutely infinite and inconsistent multiplicity, so he would not have carried out this application of his proof. (See Michael Hallett, *Cantorian Set Theory and Limitation of Size* (Oxford: Oxford University Press, 1984), chs. 1 and 4.

51 John P. Burgess, 'Frege and arbitrary functions', in W. Demopoulos (ed.), *Frege's Philosophy of Mathematics* (Cambridge, Mass.: Harvard University Press, 1995), pp. 89–107, at pp. 101–2.

52 Contrast Hallett, *Cantorian Set Theory*, pp. 126–7, and W. W. Tait, 'Frege versus Cantor and Dedekind: On the concept of number', in Tait (ed.), *Early Analytic Philosophy: Frege, Russell, Wittgenstein* (Chicago: Open Court, 1997), pp. 213–48, esp. §12.

9. WAYS OUT, EXPLORED AND UNEXPLORED

In recent years much highly productive effort has been spent in exploring weakened versions of Frege's theory.[53] What concerns me here is what responses were open to Frege in the light of his philosophy in the years from 1890 to Russell's delivery of his bombshell. It seems to me that there are two responses open to Frege, neither of which is there any evidence he considered.

The first is prompted by our having made the concept-existence assumption explicit in using a free-logic framework in deriving inconsistency. So one can read the proof as showing that α-Existence and Converse of α-Extensionality jointly entail that the predicate $\exists X(x = \alpha yXy \mathbin{\&} \neg Xx)$ does not refer to a concept. Formally, it is quite consistent to take that line, for in free logic, where the existence presuppositions of classical logic figure as *refutable* assumptions, α-Existence, α-Extensionality, and Converse of α-Extensionality are consistent. The little set-theoretic exercise with finite and co-finite sets shows this: take the domain to be the set of natural numbers, take concepts to be finite and co-finite subsets of that domain and, under some enumeration of the finite and co-finite subsets, take $\alpha x\varphi(x)$ to be the index of the set to which φ is mapped (and similarly for assignments of concepts to the second-order variables).

Under what conditions does a predicate not refer to a concept? In *Grundgesetze* Frege is quite explicit on this: $\varphi(\xi)$ denotes a function (concept) if, whenever 'ξ' is replaced by a name that denotes, the resulting sentence denotes (§§29 and 31). In volume II of *Grundgesetze* we get something perhaps a little different:

> Any object Δ that you choose to take either falls under the concept Ψ or does not fall under it; *tertium non datur*. (*Gg* II, §56)[54]

[53] See the papers collected in §§II and III of Demopoulos, *Frege's Philosophy of Mathematics*, and John P. Burgess, *Fixing Frege* (Princeton: Princeton University Press, 2005.)

[54] Frege seems to be skittering between objectual and substitutional readings of his quantifiers. On Frege's reading of quantifiers see Leslie Stevenson, 'Frege's two definitions of quantification', *Philosophical Quarterly*, 23 (1973), pp. 207–23, which goes some way to explaining why objectual and substitutional readings may not be so far apart for Frege. At this point it is first-order quantifiers that concern us. Dummett remarks that at second order Frege's 'formulations make it likely that he thought of his function-variables as ranging only over those

Immediately before this Frege says,

The law of excluded middle is really just another form of the requirement that the concept have a sharp boundary.

A concept without sharp boundary is, he says, 'wrongly termed a concept'. There are other places in his writings where Frege makes similar stipulations.[55] We now see how to read our marker for the second-order existence assumption in our proof of contradiction:

'$\exists!\Psi$' means $\forall x(\Psi(x) \vee \neg\Psi(x))$.

Of course, if the Law of Excluded Middle is part of our free logic, this does no good. More precisely, we would have that α-Existence and the Converse of α-Extensionality are inconsistent in a second-order classical free logic. But to have the Law of Excluded Middle as part of our logic wouldn't be to play the game, if we wish to turn existence presuppositions into explicitly formulated existence assumptions.[56]

Our proof of $\neg\Psi(\alpha y\Psi(y))$ from α-Existence, Converse of α-Extensionality and $\exists!\Psi$ uses no essentially classical rule. Not so our proof of $\Psi(\alpha x\Psi(x))$ from α-Existence and $\exists!\Psi$; it uses classical *reductio ad absurdum*. But it uses it only once, as the last step in the proof. Instead we could use *reductio ad absurdum* to obtain a proof of $\neg\neg\Psi(\alpha x\Psi(x))$, which suffices for obtaining a contradiction from α-Existence, Converse of α-Extensionality and $\exists!\Psi$, the latter now construed as $\forall x(\Psi(x) \vee \neg\Psi(x))$.

In this setting we read our proof as a proof of

$$\neg\forall x(\exists X(x = \alpha yXy \;\&\; \neg Xx) \vee \neg\exists X(x = \alpha yXy \;\&\; \neg Xx)).$$

The predicate $\exists X(x = \alpha yXy \;\&\; \neg Xx)$ cannot, on pain of contradiction, denote a concept. As our α-Existence claim is confined to

functions that could be referred to by functional expressions of his symbolism' (Dummett, *Frege: Philosophy of Mathematics*, p. 220).

55 'Function and concept', p. 148; 'The argument for my stricter canons of definition', *PW*, p. 152; 'Logic in mathematics', pp. 229, 241, 243.

56 Frege considers, but rejects, failure of the Law of Excluded Middle. He does so because he sees its failure as indicating that extensions of concepts would not be proper objects. He does not consider that the fault could lie with the predicate used. (See the Appendix to *Gg*, vol. II.)

concepts, it does not apply to this predicate – this predicate is not an allowed substituend – and the known route to paradox is blocked.

Can anything more general be said? In first-order intuitionist logic, $\neg\forall x(\varphi x \vee \neg\phi x)$ is formally consistent (although $\exists x\, \neg(\phi x \vee \neg\phi x)$ is not). Now, our current reading of ∃! limits the range of the second-order quantifiers to what, in intuitionist terms, are decidable properties. *If* the predicate $\exists X(x = \alpha y Xy \;\&\; \neg Xx)$ denotes a decidable property, paradox ensues. But even though the variable X ranges over decidable properties, it is not immediately evident that $\exists X(x = \alpha y Xy \;\&\; \neg Xx)$ is itself decidable. Providing a semantic model in which the denotation of $\exists X(x = \alpha y Xy \;\&\; \neg Xx)$ falls outside the range of second-order quantifiers appropriately limited in range so as to secure truth of $\forall X \forall x(Xx \vee \neg Xx)$ turns upon fine points in the interpretation of second-order quantifiers in the model-theory of second-order intuitionist logic. One thought against decidability of $\exists X(x = \alpha y Xy \;\&\; \neg Xx)$ would be that, just as the domain of individuals can increase from lesser to greater states of information (earlier to later nodes) in Kripke models for first-order intuitionist logic, so too can the range of the second-order quantifiers when we make the move to second order: as information increases one learns of new decidable properties (or learns of old ones that they are decidable).

All of this may seem, even if feasible, desperately ad hoc. All I wish to claim for it is that it has its roots in Frege's pre-paradox writings.

A second route is also licensed by those writings, and perhaps more so than Frege realized. Consider, for a moment, Frege's stipulations regarding his surrogate for definite descriptions. The surrogate behaves as it should when exactly one object falls under the concept used in constructing the description: it denotes that object. When less than or more than one object falls under the concept, it denotes the extension of the concept. *This* isn't a matter of getting anything right: it's just a stipulation that ensures description-terms always have a reference. The same attitude is to the fore when, in 'Function and concept', Frege says that we must 'lay down rules from which it follows, e.g., what "⊙ + 1" is to mean, if "⊙" refers to the Sun' ('Function and concept', p. 148). He follows this injunction with the comment, 'What rules we lay down is a matter of comparative indifference.' In principle it is open to Frege to behave just as cavalierly in the case when it is determined that a predicate does not pick out an extension, cases of the kind Cantor was well aware of, cases of the kind that emerged from Cantor's, Russell's and Burali-Forti's paradoxes. In a logically perfect language, set-abstract terms that *seem* to pick out those 'impossible sets' must be assigned some reference, but need not denote extensions of the predicates occurring in them. Basic Law Vb then needs to be qualified. It prescribes 'normal behaviour', when set-abstracts do refer to, so to say, the right extensions. 'Abnormal' predicates may be assigned the same extension even though not co-extensive – we know there must be some predicates for which this happens. The problem with such an approach is in determining the range of the 'abnormal'. (This is not to say that the 'abnormal' cases must be explicitly taken care of in the revised basic law: compare Basic Law VI, which describes only the well-behaved cases for definite descriptions.)

This too was not the way Frege chose to go. Because of the constructive role played by extensions of concepts in both *Grundlagen* and *Grundgesetze*, Frege took the route of rethinking the very notion of *extension of a concept* in the light of the very general result he obtained in the Appendix to the second volume of *Grundgesetze*: some concepts must have the same extension even though not being co-extensive. It may be true that 'the function $\Phi(\xi)$ has the same course-of-values as the function $\Psi(\xi)$' even though it is not the case that 'the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument'. Frege made the minimal change possible in the light of how he came to that discovery. The only problematic examples he knew of being obtained from $\exists X(x = \alpha y Xy \,\&\, \neg Xx)$ and $\forall X(x = \alpha y Xy \rightarrow \neg Xx)$, he proposed that two concepts $\Phi(\xi)$ and $\Psi(\xi)$ have the same extension if, and only if, the functions $\Phi(\xi)$ and $\Psi(\xi)$ have always the same value for the same argument save with the possible exceptions of the object that is their common extension.

It is known that Frege's specific proposal fails to avoid paradox (as recognized by Lesniewski, Geach and Quine).[57] Dummett says,

[57] See Gregory Landini, 'The ins and outs of Frege's way out', *Philosophia Mathematica*, 14 (2006), pp. 1–25 for a recent, and somewhat wayward, discussion. One enterprise that has attracted a small following away from the mainstream of Frege scholarship is the investigation of something akin to Frege's system in weak logics. The naive comprehension principle (roughly, α-Existence for set abstracts) is known to be consistent in certain weak logics. It is also known

The inconsistency of Frege's *Grundgesetze* system was not a mere accident (though a disastrous one) due to carelessness of formulation. He discovered, by August 1906, that it could not be put right within the framework of the theory, that is, with the abstraction operator as primitive and an axiom governing the condition for the identity of value-ranges: but the underlying error lay much deeper than a misconception concerning the foundations of set theory. It was an error affecting his entire philosophy.[58]

Exactly what Frege realized in late spring or the summer of 1906 is not quite clear. Surmise is aided by the unfinished manuscript of a response to an article of Arthur Schönflies's. A list of headers includes, for parts unwritten,

Concepts which coincide in extension, although this extension falls under the one but not the other.

Remedy from extensions of second level concepts impossible.

Set theory in ruins.[59]

Clearly the hopes of the Appendix to volume II of *Grundgesetze* had been dashed.[60]

One *methodological* error is Frege's belief that in a logically perfect language all properly formed singular terms must refer. As indicated above, there are ways to dilute the consequences of that principle, but it is, nevertheless, ill founded.

It is true that we usually do not use names that we know do not refer (save perhaps ones like 'Santa Claus' that have a recognized social context for their use). Standard logic codifies usage with referential assumptions built in. Frege himself says that in a logically perfect language 'no new sign shall be introduced as a proper name without being secured a reference'. He attempts to secure reference by stipulation. This way of proceeding is very much at odds with not just ordinary but also mathematical practice.

that this need not be the comfort it may at first seem: the naive comprehension principle is consistent in what Petr Hájek calls Basic Fuzzy Logic but the theory is not consistent with the existence of a set of natural numbers obeying a certain, moderately strong schema of mathematical induction (Petr Hájek, 'On arithmetic in the Cantor–Lukasiewicz fuzzy set theory', *Archive for Mathematical Logic*, 44 (2005), pp. 763–82).

58 Dummett, *Frege: Philosophy of Mathematics*, p. 223.

59 'On Schoenflies: *Die logischen Paradoxen der Mengenlehre*', *PW*, p. 176.

60 See further Dummett, *Frege: Philosophy of Mathematics*, pp. 4–6.

Our language contains general means for producing singular terms, e.g., definite descriptions. In *some* sense they are part of our language: they are products of its generative capacity. But by and large they are not part of our *language-in-use*. To take one of Frege's own examples, exactly because there is no least rapidly convergent series, the mathematician has no use for the expression 'the least rapidly convergent series'. True, false beliefs can lead one to use non-referring singular terms, but in use, in conversation say, matters will not run along normal lines if some parties are appraised of the facts that deny the term a reference.

Supported by theoretical claims, notably the aboutness thesis and the functional conception of concepts, connectives and quantifiers, which lead to the unassertibility of singular negative existential claims, Frege mistakes a *defeasible presumption* of reference in ordinary usage for a *presupposition*. He then takes it as given that a properly systematic, logically perfected language must respect that supposed presupposition for all singular terms generable in the language, not just those that have found a use to date. This leads to oddity but is not itself responsible for error. Error comes in the contrast between the treatment of definite descriptions and of terms for extensions of concepts: the wholesale attribution of references with a particular characteristic – satisfaction of Basic Law V – to the latter, the more relaxed who-cares-as-long-as-there-is-a-reference? attitude to description-terms when not exactly one item falls under the concept involved.

10 FREGEAN SET-THEORY: RETAINING A SEMBLANCE OF FREGEAN PREOCCUPATIONS

What were called above Frege's and Dummett's trilemmata show that we cannot maintain all that Frege says about truth. But we can keep a fair amount and a surprisingly large simulacrum of the whole Fregean project, once we reject the functional account of the connectives. To be more exact, we may maintain, with Frege:

I) Truth is unanalysable and *sui generis*.
II) For any assertoric sentence *P*, *P* and 'It is true that *P*' express the same thought.

III) If singular terms are used in the ordinary way in sentences involving simple predications, what one intends to speak of is their reference.

IV) Concepts map objects to truth-values.

V) Simple, i.e. logically unstructured, predicates refer to concepts (if they refer at all).

VI) Our propositional logic is classical (at least for negation, conjunction and disjunction).

VII) Sentences comprising a simple predication and one or more non-referring singular terms are neither true nor false.

VIII) For reasoning within the scope of presumptions of reference, our first-order logic is standard first-order logic.

IX) Basic Law V applies to extensions of concepts (or set abstracts) with second-order quantifiers ranging over sharply defined concepts. I.e., we have α-Existence, α-Extensionality, and the Converse of α-Extensionality for the extension-of-a-concept vbto.

What is unFregean is that

X) We adopt the semantic conception of falsity.

XI) We accept various truisms incompatible with the functional understanding of the logical connectives.

XII) Our general, first- and second-order logic is classical but free.

XIII) For some sentences *P*, it is true that not-*P* even though it is not false that *P*. (This is how Dummett's trilemma is evaded.)

In this setting, '$\neg\exists x(x = a)$' is true when *a* does not refer, which is to the good. Furthermore, the logic being free, there is room to give an inferentialist account of vbtos, as does Tennant.[61]

We have Basic Law V in form. It is consistent provided appropriate constraints are placed on the range of the second-order variables, i.e. on what count as concepts. (The finite–co-finite subsets-of-$\mathbb{N}$ interpretation shows that consistency is attainable.) The hard work

[61] Neil Tennant, 'A general theory of abstraction operators', *Philosophical Quarterly*, 54 (2004), pp. 105–33 (on which see further Peter Milne, 'Existence, freedom, identity, and the logic of abstractionist realism', *Mind*, 116 (2007), pp. 23–53).

goes into what we might call 'the theory of second-order ∃!', which remains to be elaborated, i.e. in specifying concepts. What are the closure conditions of the domain of (extensions of) concepts? Which predicates with one free first-order variable (and no free second-order variable) refer to concepts? Investigation of these topics puts a new spin on the old Quinean saw that second-order logic is set theory.

APPENDIX

We need several clauses to take care of falsity, clauses providing a recursive account:

(i) For any atomic sentence $Rt_1t_2\ldots t_n$
It's false that $Rt_1t_2\ldots t_n$ if, and only if,
$\neg Rt_1t_2\ldots t_n$ and $\exists!t_1$ and $\exists!t_2$ … and $\exists!t_n$.

(This applies as much to identity statements as any other atomic formulas.)

(ii) For any assertoric sentence *P*,
it's false that $\neg P$ if, and only if, *P*.

(iii) For any assertoric sentences *P* and *Q*,
it's false that $P \vee Q$ if, and only if, it's false that *P* or it's false that *Q*.

(iv) For any assertoric sentences *P* and *Q*,
it's false that $P \vee$ Q if, and only if, it's false that P and it's false that Q.

[(v) For any assertoric sentences P and Q,
it's false that P → Q if it's true that P and it's false that Q.]

One feature of this account is to be noted. We have included a minimalist account of reference:

'*t*' refers if, and only if, $\exists!t$

or, equivalently,

'*t*' refers if, and only if, $t = t$.

This, on the face of it, is a rather unFregean thing to do. On the face of it, Frege would want to *explain* the failure of '$t = t$' to be true by saying that '*t*' fails to refer. But, on the other hand, he might be thought to come close to equating the two when he says, People certainly say that Odysseus is not an historical person, and mean by this contradictory expression that the name 'Odysseus' designates nothing, has no reference.[62]

It all rather depends on what he means by 'mean'. He might, after all, just be saying that what people who say that Odysseus is not an historical person really mean to say, what they are trying to express by that – as he calls it, but he's on dodgy ground in his own terms doing so – contradictory formulation, is that the name 'Odysseus' fails to refer.

It remains to extend the definition of falsity to the first-order quantifiers. Here we make matters easy for ourselves by assuming that every object has a name. There may, of course, be singular terms that do not refer.

(vi) For any sentence $\forall x\varphi$

it is false that $\forall x\varphi$ just in case, for some singular term t, $\exists!t$ and it is false that $\varphi[t/x]$.

(vii) For any sentence $\exists x\varphi$

it is false that $\exists x\varphi$ just in case, for every singular term t, if $\exists!t$ then it is false that $\varphi[t/x]$.

Our definition of truth being given by the equivalence scheme, we do not have, as yet, *any* constraints on how either connectives or quantifiers behave with respect to truth. But that is as it should be. If one holds that the equivalence thesis says all there is to say, fundamentally, about truth, one does not look to it to justify one's logic. Rather, one looks to the logic to draw out consequences of the equivalence thesis. If the propositional logic is classical we find that $\neg p$ is true if, and only if, p is not true.

What we do find is that the negation–conjunction–disjunction fragment of classical propositional logic is sound and complete should we aim to obtain Tappolet's truisms. In similar truistic spirit, bearing in mind that some names may not refer, what we'd expect to hold for truth is this:

(vi°) For any sentence $\forall x\varphi$

[62] 'Introduction to logic', p. 191.

it is true that $\forall x\varphi$ just in case, for all singular terms 't', if $\exists!t$ then it is true that $\varphi[t/x]$.

(vii°) For any sentence $\exists x\varphi$

it is true that $\exists x\varphi$ just in case, for some singular term 't', $\exists!t$, and it is true that $\varphi[t/x]$.

The logic we want will fail to be classical precisely because we are not granting that all singular terms refer, equivalently, we are not granting $\exists!t$, equivalently, $\exists x(x = t)$, for all singular terms t. So formulas of that form play a special role. We may adopt, for example, the natural deduction rules as laid out in Tennant's *Natural Logic*, including now his 'denotation rule', or we might take the axiomatic system of Tyler Burge's 'Truth and singular terms'.[63] The denotation rule (or Burge's axiom (A9)) has it that an atomic formula entails $\exists x(x = t)$ for any name occurring in it, and that's what we want: an atomic sentence is true only if all the terms it contains refer. Ignoring the conditional and biconditional, this logic, however formulated, is sound and complete with respect to truth-preservation as determined by our truisms (suppressing any worries, which are certainly not special to this context, issued by the appeal to a substitutional reading of the quantifiers).

(Is it a truism, once we allow truth-value gaps, that a false conclusion may only follow from premises at least one of which is false? Or that a false conclusion cannot follow from true premises? *If* the former, our propositional logic is weakened, for the rules of disjunctive syllogism, double negation introduction, and *ex falso quodlibet* cease to hold.)

[63] Tyler Burge, 'Truth and singular terms', *Noûs*, 8 (1974), pp. 167–81; reprinted in M. Platts (ed.), *Reference, Truth and Reality* (London: Routledge and Kegan Paul, 1980), pp. 309–25. Both treat definite descriptions, which I have been studiously ignoring. Burge treats function symbols.