6

Grundgesetze and the Sense/Reference Distinction

Kevin C. Klement

6.1 INTRODUCTION

Frege started his work on the intellectual project culminating in his *Grundgesetze der Arithmetik*—that of deriving the basic principles of arithmetic from purely logical laws—quite early in his career. Indeed, in the Foreword to *Grundgesetze* (viii–x), Frege claims to have already begun working on it at the time of his 1879 *Begriffsschrift*, and it is certainly evident in intermediate works as well. He also claims, however, that the project was delayed significantly, in part because of modifications to his logical language brought on by certain philosophical changes of mind, mentioning the sense/reference distinction in particular. This makes it likely that the sense/reference distinction itself was the product of reflection on the proper formulation and interpretation of his logical language. The *Grundgesetze* is therefore the most natural and important test case for the application of this theory of meaning. While Frege's explicit discussion of the sense/reference distinction within *Grundgesetze* itself is quite limited (restricted primarily to §2 and §32), a close examination of how the theory is related to the work sheds considerable light on it, especially when it comes to Frege's account of the structure and identity conditions of complex senses. It may also be important for determining what, if anything, is intellectually salvageable from the project of *Grundgesetze* in light of its internal inconsistency, shown most notably by Russell's paradox. This contradiction is usually blamed on *Grundgesetze*'s Basic Law V governing value-ranges. But as we shall see, if the commitments of the sense/reference distinction had been codified within the formal system of *Grundgesetze*, it is likely that they would lead to other Cantorian-style contradictions independent of Law V.

6.2 SENSE AND REFERENCE AND THE CHANGES TO FREGE'S LOGICAL LANGUAGE

The principal arguments of Frege's 1892 'Über Sinn und Bedeutung' are well-known to any student of the philosophy of language. It opens with discussion of puzzles regarding identity, with the famous examples:

(1) the morning star = the morning star

(2) the morning star = the evening star

Frege argues that the expressions on the two sides of (2) have the same meaning in one sense, but different meanings in another: they have the same reference (*Bedeutung*) but not the same sense (*Sinn*). Frege further argues that the entire propositions (1) and (2) also each have a reference, determined by the references of their parts, as well as a sense, determined by the senses of their parts. Because the corresponding parts of (1) and (2) have the same references, the entire propositions must have the same reference as well. Frege identifies the reference of an entire proposition with its truth-value, either the True or the False. Both (1) and (2) refer to the True. However, the senses of (1) and (2) differ, because not all of their corresponding parts express the same sense. Frege calls the sense of a complete proposition a thought (*Gedanke*), and claims that the cognitive value of a proposition depends on the thought it expresses. (1) expresses a trivial thought, but (2) expresses an informative one.

Frege also invokes the distinction to explain why what would ordinarily be co-referential expressions cannot replace one another *salva veritate* in all contexts by means of his account of indirect speech or *oratio obliqua*. In certain contexts, such as in what we would now call propositional attitude reports, Frege claimed that words shift from having their ordinary sense and reference to having an *indirect* sense and reference, where the indirect reference of an expression is identified with its customary sense. Because (1) and (2) have different senses, they have different references when they appear in the context of (3) and (4):

(3) Ptolemy believes that <u>the morning star = the morning star</u>.

(4) Ptolemy believes that <u>the morning star = the evening star</u>.

Because the underlined portions of (3) and (4) have different references in that context, the complete propositions may have a different reference, that is, a different truth-value.

In the foreword to *Grundgesetze*, Frege mentions certain changes to his logical language owed to the adoption of this theory. In his earlier *Begriffsschrift* (§8), Frege employed a sign '$\equiv$' for what he called "identity of content". In *Grundgesetze*, he now employs the usual sign for forming equations in mathematics, '$=$', for identity. Prior to adopting the sense/reference distinction, Frege had not been convinced that an equation such as

(5) $2 + 2 = 2^2$

was to be interpreted as expressing an outright identity rather than a weaker relation of, e.g., equality of "numerical magnitude" or somesuch, a view Frege found in certain of his contemporaries, such as Weierstrass (see vol. II, §151;

Frege, 1914, 213). With the new theory of meaning, he was free to understand (5) as a case just like (2).

Frege puts even greater emphasis on the other change to his logical language, adding that it represents "a deep reaching development in my logical views" (x). This involves not a new or modified symbol, but a re-interpretation of the old ones. In Frege's notation, a proposition has the form:

$$\vdash\!\!-\, A$$

The vertical bar at the left of the proposition is called the "judgment stroke" and it is used to mark that the content that follows is asserted or judged as true. The portion that follows

$$-\!\!-\, A$$

on its own is supposed to represent only the content of the proposition. In Frege's function-based language, this is taken as a term, but a term for what? In his earlier work, Frege called the horizontal here the "content stroke" and claimed that $\ulcorner\!-\!\!-\, A\urcorner$ on its own is to mean something like "the circumstance that A" (Frege, 1972, §2). But this makes the sign '—' itself somewhat obscure. The other symbols in Frege's logic stand for functions. If this is a function, what sort of function is it, and how could its value be a "circumstance" if the argument term, A, does not *already* designate a "circumstance"? With the sense/reference distinction in place, Frege now regards terms in his logical language as standing for their references, and in the case of those terms to which one might add the judgment stroke, their references are truth-values. '—' stands for a function whose value is the True if the input is the True, and whose value is the False if the input is the False or an object that is not a truth-value. Other logical constants, such as Frege's negation and conditional strokes, are also interpreted as functions having truth-values as output. Rather than the rather obscure "circumstances" it becomes much clearer what the references of various expressions in the logical language are taken to be, and what possible arguments need to be considered when a function is defined.[1] As Frege himself puts it, "[o]nly a thorough engagement with the present work can teach how much simpler and more precise everything is made by the introduction of the truth-values" (x).

Treating expressions of the form $\ulcorner\!-\!\!-\, A\urcorner$ as names of truth-values also allows Frege to employ the sign '=' to do duty for a material biconditional, such as in Basic Law IV:

$$\begin{array}{l}\vdash\!\!\top\!\!-\, (-\!\!-\, a) = (-\!\!-\, b)\\ \;\;\llcorner\!\!\top\, (-\!\!-\, a) = (\,\top\!\!-\, b)\end{array}$$

[1]This is important, for example, when it comes to Frege's argument that all names in his language have a reference in §31 of *Grundgesetze*. Because every proper name stands for either a truth-value or a value-range, Frege thinks that by identifying truth-values with certain value-ranges, Basic Law V will settle the truth conditions for any identity statement in the language.

Simply put, if the truth-value of A is not the same as the negation of B, then A and B have the same truth-value. It follows that if both $\ulcorner \vdash A \urcorner$ and $\ulcorner \vdash B \urcorner$ are theorems in *Grundgesetze*, then so is the identity statement

$$\ulcorner \vdash (— A) = (— B) \urcorner.$$

It cannot be concluded from this that $\ulcorner — A \urcorner$ and $\ulcorner — B \urcorner$ have the same sense (express the same thought), unless one were willing to accept the unintuitive result that all theorems have the same cognitive value. So here we have many more cases of identity statements similar to (2) and (5).

Not only was the theory of sense and reference developed largely in order to support certain aspects of *Grundgesetze*'s logical language, there are certain claims Frege makes about the nature of senses or meaning generally that are in effect *only* true of his logical language, or another logical language like it. Frege holds that thoughts have parts corresponding to the parts of the propositions that express them. To what extent this is always true in the logical language of *Grundgesetze* is something we shall explore further in the remainder. In ordinary language, Frege would admit exceptions to this. Ordinary language contains words that do not affect the sense expressed, but only the emotional coloring of the statement. Frege gives examples such as "regrettably" and "fortunately" (Frege, 1918a, 356). Moreover, ordinary language sentences with very different grammatical forms may express the same thought. Frege regards the distinction between grammatical subject and grammatical predicate as unimportant for the structure of the thought expressed so that the switch from active to passive voice does not change the thought (Frege, 1897b, 141; 1892b, 188). In his logical language, the subject/predicate distinction gives way to a function/argument analysis, where it is not clear that there is a parallel phenomenon. Elsewhere (Frege, 1923b, 259), Frege gives these as examples of different ordinary language expressions that express the same thought:

(6) If something is a man, it is mortal.

(6′) All men are mortal.

(6″) Every man is mortal

Certainly, an analysis of the *surface* syntax (at least) of these three sentences would identify different syntactic parts, and the first one on the list involves an explicit conditional whereas the other two do not. Yet, Frege would translate these into his logical language identically:

$$\vdash\!\overset{\mathfrak{a}}{\smile}\!\!\begin{array}{l} \top\, \mathit{Mortal}(\mathfrak{a}) \\ \llcorner\, \mathit{Human}(\mathfrak{a}) \end{array}$$

Similarly, Frege denies that in a case such as (6) that the 'something is a man' clause expresses an independent thought, despite making up a grammatically independent component. He calls such clauses "quasi-propositions" (Frege, 1906a, 189; Frege, 1906b, 199). The reason for this is that in its translation into his logical language, the component part 'Human(𝔞)' contains a variable bound outside of the antecedent clause, and thus is not a complete closed term. Only the logical translation of these sentences depicts accurately the structure of the thought they all express. As Frege puts it, "[f]or discerning logical structure it is better to use letters than to rely on the vernacular" (Frege, 1906a, 190).

6.3 PRINCIPLES OF COMPOSITION: THE ISSUE

Frege's theory of sense and reference gives rise to a number of interpretive difficulties. Some of the more important involve the precise relationship between senses and references. Frege claims that a sense contains a "mode of presentation" of the reference, but exactly how a sense presents its reference, and whether or not all senses of names present their reference in roughly the same way, is not entirely clear.[2] Unfortunately, *Grundgesetze* provides little help here. One controversial issue of interpretation for which consideration of the relation of the theory to *Grundgesetze* is, I think, useful, involves the structure and composition of thoughts and other complex senses. This relates very closely to the issue of under what conditions two expressions in Frege's logical language express the same sense.

Frege says in many places that the sense of an entire proposition (the thought it expresses) is *composed* of the senses of the component expressions within that proposition. In *Grundgesetze* itself, Frege writes:

> Now, the simple or complex names of which the name of a truth-value consists contribute to expressing the thought, and this contribution of the individual name is its sense. If a name is part of the name of a truth-value, then the sense of the former name is part of the thought expressed by the latter.
>
> (§32; cf. Frege, 1918b, 378; 1923a, 390–1; 1914, 225)

(Recall that in *Grundgesetze*, a proposition is simply a name of a truth-value with the judgment stroke added.) Elsewhere, Frege admits that talk of a part/whole relationship here is figurative or metaphorical (Frege, 1923a, 390). Thoughts and other senses are abstract objects, existing in a "third realm" apart from the physical and mental (Frege, 1918a, 363). However, there is an analogy with physical complexity: if the physical object o_1 has parts p_1, p_2, and p_3, then if object o_2 is the same as o_1, it must have those parts too. The cash value of the part/whole metaphor in the realm of sense then has to do with sense identity. At least this much is true: two complex expressions with corresponding parts that have the same senses will have the same sense as well. This motivates the following principle, which I believe is relatively uncontroversial among Frege scholars that he accepts:

[2] In the philosophy of language literature, for example, Frege is often given as an example of a descriptivist about names, or one who holds that all names pick out their bearers in virtue of the bearer satisfying some descriptive content uniquely. While there is some evidence in favor of this reading, it is far from decisive. For discussion, see Klement (2010a, 881–2).

(WC) If A expresses the same sense as B, and $\ulcorner F(B) \urcorner$ is a complex expression containing B in one or more places where $\ulcorner F(A) \urcorner$ contains A, then $\ulcorner F(A) \urcorner$ and $\ulcorner F(B) \urcorner$ also express the same sense.

I dub this (WC) for "weak compositionality". Frege clearly accepts not only this, but the corresponding principle for references.

Part of the interpretive controversy involves whether or not Frege would accept something stronger, such as the *biconditional* form of (WC). Notice that this would entail accepting not just (WC), but its converse (contraposed for clarity):

(Conv) If A expresses a different sense from B, and $\ulcorner F(B) \urcorner$ is a complex expression containing B in one or more places where $\ulcorner F(A) \urcorner$ contains A, then $\ulcorner F(A) \urcorner$ and $\ulcorner F(B) \urcorner$ also express different senses.

Clearly, Frege cannot accept the corresponding principle to (Conv) for references. The numeral '7' differs in reference from '8', but both '— $7 > 5$' and '— $8 > 5$' refer to the True. Matters are less clear with senses.

There is textual evidence that Frege endorsed (Conv). In his notes for Ludwig Darmstädter, for example, he writes:

> If in a proposition or part of a proposition one constituent is replaced by another with a different reference, the different proposition or part that results does not have to have a different reference from the original; on the other hand, it always has a different sense. If in a proposition or part of a proposition one constituent is replaced by another with the same reference but not with the same sense, the different proposition or part that results has the same reference as the original, but not the same sense.
> (Frege, 1919, 255)[3]

Elsewhere, he hedges only slightly, adding an "in general":

> Now if, in a combination of signs, '$\Phi(A)$' which has a reference, a sign 'A' is replaced by another sign 'Δ' with the same reference, then obviously the new combination of signs '$\Phi(\Delta)$' will refer to the same thing as the original '$\Phi(A)$'. But if the sense of 'Δ' deviates from the sense of 'A', then in general the sense of '$\Phi(\Delta)$' will also deviate from the sense of '$\Phi(A)$'. (Frege, 1897a, 241)

Finally, in *Grundgesetze* itself, Frege suggests that the fact that replacing one expression with another with the same reference but a different sense in a true proposition will yield another true proposition expressing a distinct thought is important for understanding the progress of knowledge:

> If the group '$3 + 5$' is used in a proposition of contentual arithmetic, then we can put the sign '8' in its place without jeopardizing truth since both designate the same object, the same number proper, and everything that holds good of the object designated by '$3 + 5$' must therefore also hold good of the object designated by '8'. Further, progress in knowledge will be made in many instances of such replacement, in so far as the

[3] Here, and throughout, I have revised the translation of various key words to match the new translation of *Grundgesetze*.

senses of the co-referential signs, and thereby also the thoughts expressed by the two propositions, will be different. (*Grundgesetze*, vol. II, §104)

Concerning the belief puzzles, it is natural to suppose that the fact that the 'the morning star' and 'the evening star' differ in sense by itself should *guarantee* that (1) and (2) express different thoughts. But notice that without (Conv), this may not be so. Arguably, then, (Conv), or something very much like it, is part of the core of the sense/reference distinction.[4]

However, there are several different interpretations of Frege's views on the structure and identity conditions of thoughts. Some of them are compatible with (Conv), some are incompatible. These interpretations can be arranged in order of the stringency of the conditions they impose under which two expressions express the same sense.

6.4 FOUR VIEWS ON THE STRUCTURE OF THOUGHTS

6.4.1 The Coarse-Grained View: Logical Equivalence

According to the interpretation that adopts the least stringent criterion for sense identity, something like the following represents Frege's view in *Grundgesetze*:

(Coarse) Two closed terms A and B express the same sense of the formal language of *Grundgesetze* if and only if $\ulcorner\vdash A = B\urcorner$ can be proven using the logical axioms and rules of *Grundgesetze* alone.

Given Frege's own account of analyticity (Frege, 1950, §3), this would make two expressions synonymous just in case it is an analytic truth that they are co-referential. Notice here that A and B may be names of truth-values, in which case the thoughts they express will be the same, under this criterion, provided that A and B are logically equivalent.

Although it is often discussed, it is difficult to find unambiguous support for such a criterion among Frege interpreters.[5] Frege himself sometimes writes things that give fuel to such a reading, giving examples of pairs of logically equivalent propositions which he does claim to express the same thought, such as $\ulcorner A\urcorner$ and $\ulcorner$not not $A\urcorner$ (Frege, 1897b, 152) and $\ulcorner A\urcorner$ and $\ulcorner A$ and $A\urcorner$ (Frege, 1923a, 404). Although we should be cautious in taking a view expressed in personal correspondence as definitive, it is worth noting that in a letter to Husserl, Frege suggests something *very* close to (Coarse) for thoughts:

[4] This is further argued in Klement (2002, ch. 3; 2010b, 165), Heck and May (2011, 135–40).

[5] It is, for example, one of three alternatives (Alternative (2)) regarding sameness of sense considered in Church's 'Logic of Sense and Denotation' (see Church, 1951), though Church does not attribute this view to Frege. Perhaps the closest is Sluga, who argues that one motivation for the sense/reference distinction was to make room for basic laws of logic taking the form of identity statements, which Sluga seems to suggest would only be acceptable to Frege if the two sides expressed the same sense (Sluga, 1980, 150–4). I cannot find unambiguous support for (Coarse) in Sluga's writing, however.

> It seems to me that an objective criterion is necessary for recognizing a thought again as the same … and here I assume that neither of the two propositions contains a logically self-evident component part in its sense. If both the assumption that the content of A is false and that of B true and the assumption that the content of A is true and that of B false lead to a logical contradiction, and if this can be established without knowing whether the content of A or B is true or false, and without requiring other than purely logical laws for this purpose, then … what is capable of being judged true or false in the contents of A and B is identical, and this alone is of concern to logic, and this is what I call the thought expressed by both A and B. (Frege, 1980, 70–1)

Both $\ulcorner A$ and not $B\urcorner$ and $\ulcorner B$ and not $A\urcorner$ will lead to contradictions just in case A and B are logically equivalent, so this generates something very much like (Coarse).

Clearly, (Coarse) is incompatible with (Conv). The propositions $\ulcorner A$ or not $A\urcorner$ and $\ulcorner B$ or not $B\urcorner$ will be logically equivalent whatever the sense or reference of A and B. Thus, one can replace A with B in the former without changing the sense of the whole whether or not A and B differ in sense. Similarly, every instance of the law of identity $\ulcorner A = A\urcorner$ is logically equivalent with any other, so one can replace A with any other expression, synonymous or not, and get a result with the same sense if (Coarse) were true.

Proponents of (Coarse) might respond that Frege's talk of the senses of component expressions being parts of the senses of complex expressions is best read as an endorsement only of (WC). Geach (1976, 444), for example, notes that Frege also at times speaks of the reference of a component expression as being part of the reference of a complex expression (Frege, 1892a, 165), and as we have already seen, Frege can only hold something like (WC) for references, not something stronger. However, in later writings, Frege recants and suggests that the whole/part metaphor is only appropriate for the realm of sense. For example, in his notes for Darmstädter:

> We can regard a proposition as a mapping of a thought: corresponding to the whole-part relation of a thought and its parts we have, by and large, the same relation for the proposition and its parts. Things are different in the domain of reference. We cannot say that Sweden is a part of the capital of Sweden. (Frege, 1919, 255)

And also in Carnap's notes from Frege's 1914 lectures we find simply:

> The reference of the parts of a proposition are not parts of the reference of the proposition.
>
> However: The sense of a part of the proposition is part of the sense of the proposition. (Reck and Awodey, 2004, 87)

Since Frege never abandoned the principle corresponding to (WC) for references, it seems instead that he had in mind something stronger than that for senses.

In addition to (Conv), there are other commitments Frege holds, and uses to which he intends to put the theory of sense, which would make adopting (Coarse) in full generality highly problematic. Notice that according to the theory of indirect sense and reference, if two propositions express the same thought customarily, they will have the same indirect reference in a belief report. Thus, Frege is committed to the view that if $\ulcorner A \urcorner$ and $\ulcorner B \urcorner$ have the same sense, anyone who believes A also believes B, as he acknowledges:

> Now two propositions A and B can stand in such a relation that anyone who recognizes the content of A as true must thereby also recognize the content of B as true and, conversely …
>
> So one has to separate off from the content of a proposition the part that alone can be accepted as true or rejected as false. I call this part the thought expressed by the proposition. (Frege, 1906b, 197–8)

On this view, Frege would be committed to holding that anyone who believed $\ulcorner A \urcorner$ must believe anything equivalent, such as $\ulcorner$if not both A and not B then B does not imply not $A\urcorner$. Worse, as we saw in Section 6.2, for any two theorems of Frege's logic, the identity statement between them is also provable. Hence, if Frege's logicist project were a success, every two propositions of arithmetic would be logically equivalent, and thus, according to (Coarse), would express the same sense. Then, to believe one would be to believe any other, which certainly flies in the face of Frege's understanding of arithmetic as consisting of informative analytic truths (Frege, 1950, §88). Moreover, if A and B in (Coarse) are not limited to terms for truth-values—and it would seem arbitrary to so limit it—this criterion would mean that some of the very examples Frege gives of identity statements where the two halves have different senses, including ordinary mathematical equations such as (5), would be reduced to trivial instances of self-identity where the two sides have the same sense.

Some of these untoward consequences might be ameliorated by adopting perhaps a modified or narrower version of (Coarse). Indeed, the passage from the letter to Husserl quoted earlier strictly only supports a restricted version. Notice that there Frege assumes that "neither of the two propositions contains a logically self-evident component part". It is possible he added this assumption to block the result that all logical or analytical truths would express the same thought. However, if we interpret the criterion to mean that A and B express the same thought whenever both are contingent, contain no sub-proposition which is not contingent, and are logically equivalent, it hardly seems like much of an improvement. Moreover, the lack of parallel between how logically necessary thoughts and how contingent thoughts are understood as being composed leads to some very odd results. For example, while this approach may block the result that '$2+2=2^2$' has the same sense as '$2+2=2+2$' since these are logical truths, it would still have the result that 'Brenda has $2+2$ brothers' and 'Brenda has 2^2 brothers' express the same thought, since these are contingent but nonetheless logically equivalent (assuming the logical truth of '$2+2=2^2$'). Such a view of sense is very unappealing.

6.4.2 Intermediate Views: Concepts, their Extensions and Recarvings

This brings us to a family of interpretations I shall call *intermediate* views. What these readings have in common is a conception of sense according to which not all pairs of logically equivalent expressions are synonymous, but nonetheless some are, even in cases in which the members of the pair do not have the same syntactic form in the language of *Grundgesetze*. These readings gain traction by drawing attention to passages in which Frege makes such claims as (i) the complete proposition is the primary vehicle of meaning and one arrives at the meanings of its parts through analyzing it rather than the reverse, and (ii) the same thought can be analyzed in a variety of ways, making it appear that the thought itself does not have a definite logical structure but only comes to appear as such through our acts of analysis or precisification. These passages are relatively more common in Frege's early writings, but appear later on as well (e.g., Frege, 1919, 252; 1892b, 188–9; 1906a, 187).

Intermediate views may take a variety of forms, but often there is an emphasis on what are now called 'abstraction principles', in which the identity conditions for members of a sortal concept are fixed by reference to the holding of an equivalence relation for related entities. An example from Frege's *Grundlagen* is:

(7) the direction of line a = the direction of line b iff a is parallel to b

Of this principle, Frege claims that we can arrive at one half of the biconditional from the other by "carv[ing] up the content in a way different from the original way" (Frege, 1950, §64). This passage is prior to Frege's adoption of the sense/reference distinction, and exactly what Frege's attitude was or would have been to the idea of "recarving content", especially with regard to abstraction principles, in his later work is a difficult matter we cannot fully address here. Our interest here lies in the formal language of *Grundgesetze*, and in it there is only one principle which may qualify as an abstraction principle—the infamous Basic Law V:

$$\vdash (\dot{\varepsilon} f(\varepsilon) = \dot{\alpha} g(\alpha)) = (\smile\!\!\!\!\!\!^{\mathfrak{a}}\;\, f(\mathfrak{a}) = g(\mathfrak{a}))$$

The law in effect claims that the value-range of f is the value-range of g just in case the functions f and g have the same value for every argument. A value-range might be thought of as nothing more than the argument-value mapping generated by a function, and so one might argue that the two halves of this principle have the same meaning in a stronger sense than having the same truth-value. Indeed, when discussing a particular instance of Basic Law V (giving specific values to f and g) in his *Function und Begriff*, Frege claims that the two halves "express the same sense, but in a different way" (Frege, 1891, 143). The two halves here have obviously different syntactic forms (one is just an identity, the other is quantified, etc.), yet they seem to have a closer relation than other pairs of logical equivalents.

Frege is explicit that no other similar abstraction principles are needed so long as Law V is adopted (vol. II, §67), so apart from instances of Law V itself, it is somewhat unclear what other cases there may be of syntactically different propositions expressing the same thought on intermediate views. While there are many authors on Frege who endorse an intermediate view,[6] I'm not aware of any who develop the position sufficiently in order to state a criterion similar to (Coarse) (or (Fine) discussed later), at least not as an interpretation of Frege. One attractive option, which would include Law V as an instance of a more general phenomenon, would be to claim that all propositions (or some well-defined subset) about value-ranges or extensions are equivalent to similar propositions referring directly to their defining functions or concepts. Indeed, it could be argued that there is textual evidence in Frege's writings for such a position. In his 'Über Begriff und Gegenstand', Frege argues that phrases of the form 'the concept F', in virtue of being complete names, must refer to objects rather than concepts. He suggests that for every concept $F(\ )$, there is an associated object which correlates with it, or "goes proxy for it" when this kind of phrase is used (Frege, 1892b, 186). A thought about a concept can then be expressed by a proposition reworded or rephrased to be about the proxy object instead. For example, the following ordinary language examples are claimed to express the same thought (Frege, 1892b, 188–9):

(8) There is a square root of 4.

(8′) The concept square root of four is realized.

(8″) The number four has the property that there is something of which it is the square.

Arguably, something similar to this natural language phenomenon occurs in the formal language of *Grundgesetze* as well.[7] There, Frege (§§25, 34–5) explains how a second-level concept can be "represented" by means of a first-level concept by defining a first-level concept which an object falls under just in case it is the value-range of a function to which the second-level concept applies. To aid with this, he defines a binary function $\frown$ as follows:

$$\Vdash \backslash\dot{\alpha}\Big(\neg\forall\mathfrak{g}\,\neg\big(u=\dot{\varepsilon}\mathfrak{g}(\varepsilon)\rightarrow\mathfrak{g}(a)=\alpha\big)\Big) = a\frown u$$

Although its use is in fact more general (see Klement, 2003b, 19–20), $\frown$ in many ways works like a membership sign $\in$ in Frege's logic. Thus, he has as an important theorem (§55):

(theorem 1) $\vdash f(a) = a\frown\dot{\varepsilon}f(\varepsilon)$

[6] Examples may include Garavaso (1991), Currie (1985), Geach (1975), Hodes (1982), Blanchette (2012), and perhaps Sluga (1980) (though see note 5) among others.

[7] For discussion of whether or not the "proxy objects" of Frege (1892b) can simply be identified with value-ranges or extensions of concepts, see, e.g. Klement (2012).

Although it is, again, more general, this entails for example that a is included in the extension of $F(\ )$ just in case $F(a)$ is the True. One might then claim that the *Grundgesetze* "translations" of (8), (8′) and (8″) are, respectively:

$$\vdash\!\top\!\overset{\mathfrak{a}}{\smile}\!\top\ \mathfrak{a}^2 = 4$$

$$\vdash\!\top\!\overset{\mathfrak{a}}{\smile}\!\top\ \mathfrak{a}\frown\dot{\varepsilon}(\varepsilon^2 = 4)$$

$$\vdash 4\frown\dot{\varepsilon}(\top\!\overset{\mathfrak{a}}{\smile}\!\top\ \mathfrak{a}^2 = \varepsilon)$$

Perhaps a natural refinement of an intermediate view would include the suggestion that these three propositions, and collections of propositions differing from each other in the same sort of way, also express the same thought. More generally, one might hold that the two halves of any instance of (theorem 1) express the same sense.

There are, however, problems with this approach similar to the problems for the coarse view. Notice that the process of replacing talk of a concept with talk about its extension or value-range can continue *ad infinitum*. From the above examples, one can move to similarly-modified propositions such as:

$$\vdash \dot{\varepsilon}(\top\!\overset{\mathfrak{a}}{\smile}\!\top\ \mathfrak{a}^2 = \varepsilon)\frown\dot{\alpha}(4\frown\alpha)$$

And then to, e.g.:

$$\vdash \dot{\alpha}(4\frown\alpha)\frown\dot{\omega}(\dot{\varepsilon}(\top\!\overset{\mathfrak{a}}{\smile}\!\top\ \mathfrak{a}^2 = \varepsilon)\frown\omega)$$

The process can be repeated until one has a proposition that runs all the way down the block and back. It seems strange to think that all such propositions could have the same cognitive value, or that to believe one is tantamount to believing any of the others.

Furthermore, one must remember that $\frown$ is a *defined* sign, and Frege is clear that a defined sign takes on the complex sense of its definiens (§27). With the definition unpacked (theorem 1) reads:

$$\vdash f(a) = \backslash\dot{\alpha}\left(\top\!\overset{\mathfrak{g}}{\smile}\!\top\!\!\begin{array}{l} \mathfrak{g}(a) = \alpha \\ \llcorner\ \dot{\varepsilon}f(\varepsilon) = \dot{\varepsilon}\mathfrak{g}(\varepsilon) \end{array}\right)$$

Holding that the two halves of an instance of this express the same sense seems to reject wholly the idea that the senses of component expressions are parts of the sense of the whole. Where did the senses of all the logical signs that appear on the right here "go" on the left? (Recall that logical constants refer to concepts and relations for Frege—they are not syncategorematic.) Holding it in this case, and not holding it in general for all identity statements one can prove in *Grundgesetze* seems arbitrary. Notice, for example, that this was not the only possible definition for $\frown$. In *Grundgesetze*, one can also establish, e.g.:

$$\vdash f(a) = \backslash\dot{\alpha}\left(\overset{\mathfrak{g}}{\smile}\!\top\!\!\begin{array}{l} \mathfrak{g}(a) = \alpha \\ \llcorner\ \dot{\varepsilon}f(\varepsilon) = \dot{\varepsilon}\mathfrak{g}(\varepsilon) \end{array}\right)$$

Postulating a difference in sense between the two halves of instances of this equation but not for the two halves of instances of (theorem 1) seems completely *ad hoc*, and holding that we have the same sense in all such cases seems to push straight back to the coarse view.

Moreover, this view is still incompatible with (Conv). For example, consider these instances of (theorem 1):

$$\vdash (7 \leq 7) = (7 \frown \dot{\varepsilon}(\varepsilon \leq 7)) \tag{9}$$

$$\vdash (7 \leq 7) = (7 \frown \dot{\varepsilon}(7 \leq \varepsilon)) \tag{9'}$$

If every instance of (theorem 1) has the same sense expressed on the two sides of the identity, then, since the left sides of these two instances are the same, the two right sides would, by transitivity of sameness of sense, also express the same sense. But notice that '$7 \frown \dot{\varepsilon}(\varepsilon \leq 7)$' and '$7 \frown \dot{\varepsilon}(7 \leq \varepsilon)$' differ from each other by the replacement of '$\dot{\varepsilon}(\varepsilon \leq 7)$' for '$\dot{\varepsilon}(7 \leq \varepsilon)$'. These two subexpressions have different references, so *a fortiori*, they have different senses. But this is not, on the view considered here, enough to make the two right sides of the above express different thoughts, and so (Conv) must be false on this interpretation.

There is not much exegetically beyond the one claim about an instance of Law V in *Function und Begriff* to support the kind of intermediate position we have been exploring here. When Frege discusses Law V in *Grundgesetze*, he simply writes that the two halves are co-referential, and does not claim that they express the same sense (§3; cf. §10, vol. II, §38, and Afterword). Similarly, in discussing instances of (theorem 1) in §34, he only claims that $\Delta \frown \dot{\varepsilon}\Psi(\varepsilon)$ and $\Psi(\Delta)$ are co-referential, not that they express the same sense. It may have been that the passage in *Function und Begriff* was a slip or temporary position he later reconsidered. In light of the difficulties mentioned, it seems uncharitable to saddle Frege with this sort of position as his considered view.

6.4.3 The Fine-Grained View: Language-Proposition Isomorphism

On the interpretation I (and others[8]) endorse, the structure of a Fregean thought is indeed isomorphic to the proposition that would be used to express it in the logical language of *Grundgesetze*. For reasons seen at the end of Section 6.2, such an isomorphism need not hold in natural language. Since *Grundgesetze* is an austere language with a small stock of primitive expressions, none of which involves the sense of any other, a *Grundgesetze* proposition written without signs introduced by definition is almost a perfect model of the structure of the thought expressed. No meaningful sign can be added or removed without altering the sense. The only difference in expression which would be irrelevant to the thought expressed would be the choice of one variable in place of another. This suggests the following:

[8] See also Klement (2002, ch. 3) and (2010b). I believe my position is, aside from small points of detail, the same as Dummett's (see, e.g., Dummett (1974), (1981)); the view of Heck and May (2011) seems rather close as well.

(Fine) Two closed terms A and B of the formal language of *Grundgesetze* written without signs introduced by definition express the same thought if and only if they differ from one another by at most arbitrary choice of bound variable.

So '$\dot{\alpha}(\alpha = \alpha)$' expresses the same sense as '$\dot{\varepsilon}(\varepsilon = \varepsilon)$', but any change beyond such trivial variable swaps would result in a distinct sense—even, e.g., '$\dot{\alpha}(\top\!\top\, \alpha = \alpha)$' would differ. On this view, because the two sides of Law V are syntactically different, they must be taken as expressing different senses.

This interpretation has the advantage of explaining claims made by Frege such as:

> The world of thoughts has a model in the world of propositions, expressions, words, signs. To the structure of the thought there corresponds the compounding of words into a proposition ... (Frege, 1918b, 378)

It also seems fully compatible with the suggestion that believing A and believing B are one and the same when $\ulcorner A \urcorner$ and $\ulcorner B \urcorner$ express the same sense. While the "calculation" involved in moving between, e.g., a proposition and its double negation may seem so trivial that we hardly notice it, it is still a form of calculation not different in kind from that involved in recognizing $2 + 2$ and 2^2 as the same number. It seems that a difference in sense is warranted in both cases. (Fine) is also fully compatible with (and indeed demands) (Conv), which, we argued at the end of Section 6.3, seems crucial to the theory of sense and reference as a whole.

Perhaps the greatest challenge facing this interpretation involves explaining why Frege sometimes claims that the same thought is capable of different analyses or decompositions into parts. However, this is not as difficult as it seems. At least in his later work, when Frege makes such claims he nearly always gives as his example a complex proposition which is capable of different function/argument analyses (and we have already considered all the exceptions). The simplest kind of case involves a proposition consisting of multiple proper names and a phrase for a relation:

> If several proper names occur in a proposition, the corresponding thought can be analyzed into a complete part and an unsaturated part in different ways. The sense of each of these proper names can be set up as the complete part over against the rest of the thought as the unsaturated part. (Frege, 1906a, 192)

Frege gives as example here the thought expressed by 'Jupiter is larger than Mars', which he tells us, can be divided either into the complete sense of 'Jupiter' and the unsaturated sense of '( ) is larger than Mars' or into the complete sense of 'Mars' and unsaturated sense of 'Jupiter is larger than ( )'. If we were to put this proposition into the logical language of *Grundgesetze*, we might add a relation sign '$\succ$' and constants '♃' and '♂' for Jupiter and Mars, and write '$\vdash$ ♃ $\succ$ ♂'. As Frege makes clear, an unsaturated part of a thought is the sense of a function. Frege claims that in addition to primitive function signs, one can also obtain a function sign by removing a name from a complex term (§26). (In informal discussion, Frege then often inserts ‘ξ’ or ‘ζ’ in the gap where the name was removed to indicate where an argument is needed—see §1.) If we remove ‘♃’ from ‘♃ ≻ ♂’, we get the function-name ‘ξ ≻ ♂’. If we remove ‘♂’ we get ‘♃ ≻ ξ’. It is just as correct to regard the complex term ‘♃ ≻ ♂’ as having the parts ‘♃’ and ‘ξ ≻ ♂’ as it is to regard it as having ‘♂’ and ‘♃ ≻ ξ’ as its parts. Hence, it is equally correct to regard the parts of the thought it expresses as corresponding to either syntactic analysis. None of this in any way poses a problem for (Fine). Whether we place ‘♃’ in the argument spot of ‘ξ ≻ ♂’ or ‘♂’ in the argument spot of ‘♃ ≻ ξ’ we arrive back at ‘♃ ≻ ♂’, and nothing here shows that any other closed expression has the same sense as ‘♃ ≻ ♂’. These different possible *partial* “decompositions” of the same thought also do not show that there is not a unique *final* analysis of the simple parts of this thought, since these complex function expressions themselves have parts, including the doubly unsaturated sense of ‘≻’. Either way one begins the decomposition, in the end, one will arrive at the same simple parts (cf. Dummett, 1981, ch. 15).

In even more complex propositions, an even greater number of partial function/argument analyses are possible. For example, Frege writes:

> We should mention that, strictly speaking, it is not in itself that a thought is singular, but only with respect to a possible way of analysing it. It is possible for the same thought, with respect to a different analysis, to appear as particular (Christ converted some men to his teaching). (Frege, 1906a, 187)

Passages such as these have led commentators away from the conclusion that thoughts can be considered structured wholes or be identified as having certain parts independent of our acts of analyzing them. But this is overreaching. Again, this can be explained as what would amount to one and the same formal language representation of the same thought being amenable to multiple function-argument partial decompositions. A “translation” of ‘Christ converted some men to his teaching’ would have a name for Christ in it, and thus could be analyzed into that name along with a complex first-level function-name; on that decomposition the thought would appear singular. But the translation would also contain an existential quantifier, and thus could be decomposed into a second-level function-name along with a first-level function as argument, in line with Frege’s general understanding of quantification (see *Grundgesetze*, §22). On this analysis, it would appear particular. Again, however, nothing here goes to show that two syntactically different formal language expressions could express the same sense. It is worth noting that elsewhere, Frege makes it clear that when a proposition involves generality or quantification, generality can be considered a property the thought itself has (Frege, 1923b, 259).

6.4.4 The Ultra-Fine-Grained View

We lastly come to a rather different interpretation of the structure of thoughts, to my knowledge only endorsed by Landini (2012, 130–44). His reading makes senses even more finely grained than the view endorsed in the previous section. At first blush, this may seem impossible: how could senses be more fine-grained than *Grundgesetze* terms? Landini does not deny (Fine); however, he holds that there are senses or thoughts which, despite being formed by "re-composing" component senses which are expressed in *Grundgesetze*, cannot actually be expressed in its language without further syntactic resources. On his reading, the sense of a closed term has as its parts *only* the senses of the primitive expressions making it up. Treating '3', '>' and '2' for the moment as primitive signs, the thought expressed by '$3 > 2$' has only the senses of '3', '>' and '2' as parts, and not, additionally, the senses of '$3 > \xi$' or '$\xi > 2$'. Landini contends that '$3 > 2$' has a unique syntactic analysis. It contains only a single binary function sign and two object names. It cannot with equal justice be read as containing "complex" function-names such as '$3 > \xi$' and '$\xi > 2$'. Landini contrasts his view with what he calls the "liberal view" of semantic compositionality which identifies a greater number of parts within a complex sense.

It is not of course that Landini denies that there are such concepts as *being something three is greater than* and *being greater than two*, or even that it is possible to refer to such concepts. It is rather that he believes that there are no *Grundgesetze* expressions which do refer to these concepts. To have such, *Grundgesetze*'s syntax would have to be expanded to allow such expressions as '$[\xi > 2](3)$' and '$[3 > \xi](2)$' akin to the modern Lambda Calculus's '$\lambda x.(x > 2)(3)$' and '$\lambda x.(3 > x)(2)$' (Landini, 2012, 141–2). The thought expressed by '$[\xi > 2](3)$', although "recomposed" at least in part of senses that are found in *Grundgesetze*, cannot, as a whole, be expressed in *Grundgesetze* as such.[9] This contrasts with the liberal view that already sees '$3 > \xi$' and '$\xi > 2$' as parts of '$3 > 2$' so that the thoughts that would be expressed by '$[\xi > 2](3)$' and '$[3 > \xi](2)$' were the language to be so enriched syntactically, are, in effect, already expressed by '$3 > 2$' (and so are identified with each other). In this sense, the "liberal view" adopts less fine-grained criteria.

Unfortunately, Landini does not provide any textual evidence that favors his view over the more liberal view. Moreover, Landini cannot provide the same kind of explanation that we gave in the previous section for why Frege himself often claims that the same thought can be divided into complete and incomplete part, or into function and argument, in different ways (see also Frege 1903, 281; 1904, 291; 1906a, 191; 1906b, 202; 1980, 101). Landini dismisses such passages on the grounds that they occur in Frege's non-technical works and do not apply to Frege's own formal language. For *Grundgesetze* itself, Landini insists that "expressions made with parametric letters" (by which he means such function-names as '$\xi > 2$' and '$3 > \xi$') are not

[9] The difference between Landini's interpretation and mine, then, is quite similar to the difference between the criteria of identity of senses taken in Alternatives (0) and (1) of Church's Logic of Sense and Denotation. I have argued that Alternative (1) makes for a closer approximation to Frege's views elsewhere (Klement, 2002, 104–5; 2010b).

"genuine function terms of the object language" (Landini, 2012, 137). While Frege does not state recursive syntactic rules for his object language with quite the same precision expected by contemporary logicians, he did, I believe, adequately make it clear what does and does not count as an object language expression. Indeed, in §§26–30 he explicitly discusses what he calls "the correct formation of a name", and includes under this rubric what he calls "complex function-names". In §30, he discusses "two ways to form a name". In the case of function-names with one argument, these may be formed *either* by filling one of the argument places of a function-name with two arguments, or else by taking away one of the names occurring within an already formed complex proper name. These sections make evident that Frege considered complex function-names as part of his language, and they are included within the scope of the claim made two sections later (§32; quoted earlier) that when a simple or complex name occurs in the name of a truth-value the sense of the component name is part of the thought thereby expressed. It is true that the "parametric letters" 'ξ' and 'ζ' do not occur in *Grundgesetze* propositions. Frege is explicit that 'ξ' is not really a part of the function-name '$\xi > 2$'; its role is simply to hold open the "gap" in what might otherwise be written '$\ > 2$'. Its sense is likewise "incomplete" or "unsaturated". It would be misleading to write '$[\xi > 2](3)$' as this obscures the fact that the incomplete sense becomes complete when the argument sense saturates it. The result is better written '$3 > 2$', which, then, is precisely the same result one gets when the sense of '$3 > \xi$' is filled by the sense of '2'. Thus, the ultra-fine-grained interpretation does not do justice to Frege's understanding of the senses of functions as unsaturated.

Consider again (9) and (9′) from Section 6.4.2. These are both derived from (theorem 1) by means only of the replacement rule for free variables. For (9), the object variable 'a' is replaced by '7' and function variable 'f' is replaced by '$\xi \leq 7$'. To get (9′), the 'a' is again replaced by '7', and 'f' is instead replaced by '$7 \leq \xi$'. Notice again that the left sides of the two resulting equations are the same. The very same closed term '$7 \leq 7$' can be obtained from the open term '$f(a)$' by replacing 'f' with *different* function-names. Many deductions in *Grundgesetze* rely on the fact that identical complex expressions can result by instantiating different variables to different values. Quite obviously, the same expression '$7 \leq 7$' cannot be read as expressing different thoughts depending on how it was so arrived at. The only plausible interpretation is to regard the thought it expresses as analysable into a functional part and argument part in different ways. Landini notes that instead of using a replacement rule for free function variables which allows their replacement by complex function expressions, Frege *could* have given a deductively equivalent system using a comprehension schema for functions. But this is irrelevant. How Frege *actually* proceeded is what is important for interpreting his views on sense-identity.

6.5 THE OVERABUNDANT THIRD REALM

While I find fine-grained readings of the nature of senses most amenable to Frege's own understanding of his theory of meaning, they give rise to problems of their own. Unfortunately, Frege would not have been in a position to appreciate these problems while composing *Grundgesetze* in the 1890s. The finer-grained one understands the identity conditions of senses to be—that is, the more discriminations between senses one makes—the greater the number of senses one will be committed to. But exactly how populous is the third realm, and mightn't there be problems with any reasonable answer? The logical system of *Grundgesetze* is inconsistent due to Russell's paradox among others, but it becomes consistent if Frege's Basic Law V is dropped or even suitably weakened (see Heck, 1996). However, an overabundance of thoughts or other senses threatens to reintroduce problems of the same stripe.

The central problem is one Russell tried to warn Frege about in a letter written a few months after the famous one disclosing Russell's paradox. Russell despairs that "[f]rom Cantor's proposition that any class contains more subclasses than objects we can elicit constantly new contradictions" (Frege, 1980, 147) and goes on to describe a problem for his own theory of mind- and language-independent propositions, suggesting that *mutatis mutandis* it might also be a problem for Frege's theory of thoughts. By Cantor's power-class theorem, there ought to be more classes of propositions than propositions. However, it seems possible to generate a distinct proposition for every class of propositions, such as the proposition that every member of that class is true. Cantor's diagonalization procedure will then generate a contradiction. Russell formulated the contradiction using a different logical notation, based on Peano's. Therein the truth-value and sense of a proposition were not distinguished in a manner to Frege's liking, and, as a result, Russell never got his point across.[10]

However, this is a warning of which Frege should have taken notice. As Russell also attempted to demonstrate in a follow-up letter (Frege, 1980, 160), it is not necessary to speak of classes here—one may speak of (propositional) functions or Fregean concepts instead. A first-level concept, for Frege, is a function whose value for every object as argument is always a truth-value. Because there are two truth-values, the number of possible first-level concepts should be 2^n where n is the number of objects. By argumentation due to Cantor, $2^n > n$ even when n is infinite. (For what amounts to more or less the same reason, Frege himself argued in his afterword to vol. II of *Grundgesetze* on Russell's paradox that no second-level function could have a distinct object as value for each possible first-level function as argument.) However, Fregean thoughts are objects, and isn't it possible to come up with a distinct thought, or possibly more than one distinct thought, for every concept? For each sense

[10] For further discussion of their breakdown in communication on this matter, see Klement (2001).

of a concept, for example, consider the thought that everything falls under that concept formed with that sense as part. There seem to be at least as many such thoughts as there are concepts. In that case, it would be impossible for 2^n to be greater than n.

Cantor's diagonal process then leads us to the following contradiction. Let $\mathbb{K}$ be a concept a thought t falls under just in case t is a thought of the form *every object is* F for some concept F, but t does not fall under the concept F of which it asserts the generalization. For example, the thought *every object is a cat* is not itself a cat, and so it falls under $\mathbb{K}$. The thought *every object is self-identical*, on the other hand, is self-identical, and so does not fall under $\mathbb{K}$. However, now consider the thought *every object is* $\mathbb{K}$; call this thought $\mathbb{E}$. Does $\mathbb{E}$ fall under $\mathbb{K}$? If it does not, then it does, and if it does, it does not. Contradiction.

Although this contradiction does not utilize classes, value-ranges or extensions, and so is independent of Basic Law V, it cannot be formulated in the consistent portion of Frege's extant logic because the language has no way of speaking about senses or thoughts instead of their references. In other words, it has no way of referring to senses in the way that ordinary language expressions do when they appear in indirect contexts. As Frege admits in his exchange with Russell "in my concept-script I did not yet introduce indirect speech because I had as yet no occasion to do so" (Frege, 1980, 149). Later he adds that "[t]o avoid ambiguity, we ought really to have special signs in indirect speech, though their connection with the corresponding signs in direct speech should be easy to recognize" (Frege, 1980, 153). In the present essay we cannot fully examine how indirect speech ought to be added to the language of *Grundgesetze* but a brief sketch may be enough to show that the sort of diagonal paradox of thoughts Russell warned of is a real danger.[11]

To capture Frege's commitment to senses formally one might add special quantifiers and variables for senses. Here I use Roman letters (for free variables) and Gothic letters (for bound variables) with asterisks for that purpose, e.g., a^*, F^*, $\mathfrak{a}^*$, $\mathfrak{f}^*$, etc. Borrowing from the approach of Church (1951), for each logical constant, in an expanded language, one might make use of a hierarchy of constants written with superscripts indicating their position in the hierarchy of senses. Thus, we might write '$=^{[0]}$' for the normal identity relation, '$=^{[1]}$' to refer to the sense of '$=^{[0]}$', and '$=^{[2]}$' to refer to the sense of '$=^{[1]}$' and so on. (It would be natural to omit the superscript when it is [0]; the use of multiple dots or asterisks could achieve the same purpose.) There is some disagreement in the secondary literature about how exactly to understand the sense of a function expression.[12] Here we only assume that expressions that refer to them also have "argument spots" which when filled by other names of senses create complex expressions that refer to the complex senses formed by the saturation of the function sense. For example, if '$\succ^{[1]}$' refers to the sense of

[11]A much fuller attempt is undertaken in my (2002, ch. 5).

[12]For a discussion of this issue, see my (2002, 65–76) and (2010b).

'$\succ$', '$\text{♃}^{[1]}$' to the sense of '♃' and '$\text{♂}^{[1]}$' to the sense of '♂' then '$\text{♃}^{[1]} \succ^{[1]} \text{♂}^{[1]}$' would refer to the complex sense of '$\text{♃} \succ \text{♂}$'. In that way, for any closed term A, one could create a name for the sense of $\ulcorner A \urcorner$ by raising the superscripts on every constant in A by 1. I shall abbreviate this as $\ulcorner A^{[+1]} \urcorner$.

Lastly we need a way of expressing the relationship between a sense and the reference it picks out or presents (assuming there is one). Here we shall employ '$\rhd$' (adapted from Parsons, 2001) as a new primitive relation sign, but we can informally explain its semantics as follows:

$$x \rhd y = \begin{cases} \text{the True,} & \text{if } x \text{ is a sense that presents the object } y \text{ as reference,} \\ \text{the False,} & \text{otherwise.} \end{cases}$$

(Strictly, we are adding not merely '$\rhd$', but an entire hierarchy also with '$\rhd^{[1]}$', '$\rhd^{[2]}$', etc.) This is a first-level relation, and thus only appropriate for presenting the relationship between the sense of an object and that object. For the relation between senses of functions and the functions they present, we shall need a higher-level relation. Given Fregean customs regarding higher-level relations, this might be written:

$$f^*(\alpha) \underset{\alpha}{\rhd} f(\alpha)$$

(Here the α is bound by the symbol with it as subscript and fills the argument spots of its relata.) Rather than taking this as primitive, one might suggest this definition:

$$\Vdash \left(\overset{\mathfrak{a}}{\smile}\overset{\mathfrak{a}^*}{\smile} \begin{array}{l} f^*(\mathfrak{a}^*) \rhd f(\mathfrak{a}) \\ \llcorner\, \mathfrak{a}^* \rhd \mathfrak{a} \end{array} \right) = \left(f^*(\alpha) \underset{\alpha}{\rhd} f(\alpha) \right)$$

This definition may prove controversial.[13] However, it matters little in what follows whether we take this as a defined or primitive sign.

Let us now see whether or not a paradox such as that involving $\mathbb{K}$ threatens. The connection with the issue of the identity conditions of senses arises in the following way. Consider again thoughts of the form *everything is F*. In our expanded language, the sense of the quantifier '$\overset{\mathfrak{a}}{\smile} \ldots \mathfrak{a} \ldots$' is written '$\overset{\mathfrak{a}^*}{\smile}{}^{[1]} \ldots \mathfrak{a}^* \ldots$', and so we may write '$\overset{\mathfrak{a}^*}{\smile}{}^{[1]} F^*(\mathfrak{a}^*)$' as a "Roman marker" (cf. *Grundgesetze*, §17) of such a thought. A fine-grained understanding of the identity of senses seems to require the following:

$$\vdash \overset{\mathfrak{f}^*}{\smile}\overset{\mathfrak{g}^*}{\smile}\overset{\mathfrak{f}}{\smile}\overset{\mathfrak{g}}{\smile} \begin{array}{l} \overset{\mathfrak{a}}{\smile}\, \mathfrak{f}(\mathfrak{a}) = \mathfrak{g}(\mathfrak{a}) \\ \llcorner\, \left(\overset{\mathfrak{a}^*}{\smile}{}^{[1]} \mathfrak{f}^*(\mathfrak{a}^*) = \overset{\mathfrak{a}^*}{\smile}{}^{[1]} \mathfrak{g}^*(\mathfrak{a}^*) \right) \\ \llcorner\, \mathfrak{g}^*(\alpha) \underset{\alpha}{\rhd} \mathfrak{g}(\alpha) \\ \llcorner\, \mathfrak{f}^*(\alpha) \underset{\alpha}{\rhd} \mathfrak{f}(\alpha) \end{array} \qquad \text{(ID)}$$

[13]For a defense of a definition like this, along with discussion of some complications, see Klement (2002, 119, 150) and (2010b).

This states that if F^* is a sense of the function F and G^* is a sense of the function G and the thought that *everything is* F (as presented through F^*) is identical to the thought that *everything is* G (as presented through G^*), then the functions F and G coincide, i.e., have the same value for every argument (see Frege, 1892–95, 121). This seems demanded by a fine-grained account (one endorsing something like (Fine)) in the following way. An expression of the form $\ulcorner \overset{\mathfrak{a}}{\smile} \Phi(\mathfrak{a})\urcorner$ will differ by at most arbitrary choice of bound variable from another expression of the same form $\ulcorner \overset{\mathfrak{a}}{\smile} \Psi(\mathfrak{a})\urcorner$ just in case $\ulcorner\Phi(\ )\urcorner$ and $\ulcorner\Psi(\ )\urcorner$ differ from each other by at most that much. This would mean that $\ulcorner\Phi(\ )\urcorner$ and $\ulcorner\Psi(\ )\urcorner$ would express the same sense as well. Senses of functions, like all senses, are determinate in the sense that they present at most one reference. So if $\ulcorner\Psi(\ )\urcorner$ and $\ulcorner\Phi(\ )\urcorner$ have the same sense, they must refer to the same function. Generalizing the point beyond those thoughts of this form expressible in language to all thoughts of this form, one obtains (ID) above.

The paradoxical concept $\mathbb{K}$ which a thought falls under just in case it is a thought of this form which does not fall under the concept it generalizes may now be defined:

$$\Vdash \left(\begin{array}{l} \top\!\overset{\mathfrak{f}^*}{\smile}\!\overset{\mathfrak{f}}{\smile}\!\!\top\!\top\; \mathfrak{f}(x) \\ \qquad\qquad\;\;\llcorner\; x = \overset{\mathfrak{a}^*}{\smile}{}^{[1]}\mathfrak{f}^*(\mathfrak{a}^*) \\ \qquad\qquad\;\llcorner\; \mathfrak{f}^*(\alpha) \underset{\alpha}{\rhd} \mathfrak{f}(\alpha) \end{array} \right) = \mathbb{K}(x) \tag{K}$$

We now let $\mathbb{K}^{[+1]}$ be the result of raising the superscript on every constant in the definition above by 1. Thus $\mathbb{K}^{[+1]}$ is the sense of '$\mathbb{K}$' and we have:

$$\vdash \mathbb{K}^{[+1]}(\alpha) \underset{\alpha}{\rhd} \mathbb{K}(\alpha) \tag{K^{+1}}$$

This allows us to define the thought $\mathbb{E}$, that *every object is* $\mathbb{K}$:

$$\Vdash \left(\overset{\mathfrak{a}^*}{\smile}{}^{[1]}\mathbb{K}^{[+1]}(\mathfrak{a}^*)\right) = \mathbb{E} \tag{E}$$

The definitions (K) and (E) and the principles (K^{+1}) and (ID) alone (along with the consistent part of Frege's logic) suffice to prove the contradictory results:

$$\vdash \mathbb{K}(\mathbb{E})$$
$$\vdash\!\!\top \mathbb{K}(\mathbb{E})$$

This deserves a full proof,[14] but we will make do here with a sketch. Suppose $\mathbb{K}(\mathbb{E})$ for *reductio*.[15] By the definition (K), this means that there is a function F and sense thereof F^* such that $F^*(\alpha) \underset{\alpha}{\rhd} F(\alpha)$ and $\mathbb{E} = \overset{\mathfrak{a}^*}{\smile}{}^{[1]}F^*(\mathfrak{a}^*)$

[14] For a full deduction of what amounts to the same contradiction, see Klement (2002, 166–9).

[15] Frege himself does not employ proofs using an assumption such as *reductio* proofs in his system. But it is easily shown that any such proof could be transformed into a full proof.

and $\top F(\mathbb{E})$. Unpacking the definition of $\mathbb{E}$, this means $\overset{\mathfrak{a}^*}{\smile}{}^{[1]}\mathbb{K}^{[+1]}(\mathfrak{a}^*) = \overset{\mathfrak{a}^*}{\smile}{}^{[1]}F^*(\mathfrak{a}^*)$. Then, by ($K^{+1}$) and (ID) and the results so far, $\overset{\mathfrak{a}}{\smile}\, \mathbb{K}(\mathfrak{a}) = F(\mathfrak{a})$. But then, since $\top F(\mathbb{E})$, we get $\top \mathbb{K}(\mathbb{E})$ which contradicts our assumption. So by *reductio*, $\vdash\!\top \mathbb{K}(\mathbb{E})$. But, now, applying the definition (K), removing the double negation, and instantiating $\mathfrak{f}^*$ and $\mathfrak{f}$ to $\mathbb{K}^{[+1]}$ and $\mathbb{K}$ respectively, we get, through (K^{+1}) and an instance of the principle of identity, $\vdash \mathbb{K}(\mathbb{E})$.

There are a variety of ways that someone with broadly Fregean views might respond to problems such as this. We cannot canvass them all here.[16] It seems clear nevertheless that Frege should have paid closer heed to Russell's warning. Interestingly, the first formulation of Church's Logic of Sense and Denotation was independently shown to fall prey to a Cantorian paradox of senses quite similar to this by John Myhill (1958), and so the difficulty is sometimes referred to as "the Russell-Myhill Antinomy". Notice that there is little wiggle room, especially for someone attracted to a relatively fine-grained understanding of senses and thoughts. Definitions such as (K) and (E) are theoretically dispensable, and so the problem cannot be blamed on them. Unless one were to claim that the function-name $\mathbb{K}$ simply has no sense, one cannot avoid something like (K^{+1}) either. This leaves only (ID) which it would be very hard indeed to deny without doing real violence to the fine-grained understanding of senses. I have tried to argue in the previous sections that a close reading of how Frege's theory relates to the project of *Grundgesetze* more or less forces us in the direction of a fine-grained reading. But now it appears that these aspects of Frege's theory of meaning, intended to provide philosophical support for the logical language of *Grundgesetze*, actually threaten to undermine even what remains salvageable from the system in the wake of Russell's paradox and related contradictions: a sad state of affairs.

It should be noted for the record that it is not at all clear that this antinomy is avoided on intermediate readings either, but it would depend on the details of the view. Most of the examples we saw in Section 6.4.2 of syntactically different expressions that express the same sense according to the intermediate readings involve pairs of expressions where, in one, a function expression is used, and in the other, an expression for its value-range is used instead. If those are *the only* cases in which syntactically different expressions can express the same sense, then, because nothing in the above involves value-ranges, the intermediate readings would not fare any better.

Perhaps one might suppose that there is something more systemically wrong with the way that I have sketched how Frege's logical language could be expanded to include the commitments of the theory of sense and reference, but I believe it to be deeply rooted in Frege's own approach to these issues. May those with a better approach improve upon it.

[16] For a discussion of several responses to these and related problems, see Klement (2002, ch. 7) and (2003a).

REFERENCES

Blanchette, Patricia (2012). *Frege's Conception of Logic*. Oxford: Oxford University Press.

Church, Alonzo (1951). 'A Formulation of the Logic of Sense and Denotation', in *Structure, Method and Meaning: Essays in Honor of Henry M. Sheffer*, edited by Paul Henle, Horace M. Kallen and Susanne K. Langer. New York: Liberal Arts Press, pages 3–24.

Currie, Gregory (1985). 'The Analysis of Thoughts', *Australasian Journal of Philosophy* 63:283–98.

Dummett, Michael (1974). *Frege: Philosophy of Language*. Cambridge, Mass.: Harvard University Press.

Dummett, Michael (1981). *The Interpretation of Frege's Philosophy*. Cambridge, Mass.: Harvard University Press.

Frege, Gottlob (1891). 'Function and Concept', in Frege (1984), pages 137–56.

Frege, Gottlob (1892a). 'On Sense and Meaning', in Frege (1984), pages 157–77.

Frege, Gottlob (1892b). 'On Concept and Object', in Frege (1984), pages 182–94.

Frege, Gottlob (1892–95). 'Comments on Sense and Meaning', in Frege (1979), pages 118–25.

Frege, Gottlob (1897a). 'On Mr. Peano's Conceptual Notation and My Own', in Frege (1984), pages 234–48.

Frege, Gottlob (1897b). 'Logic', in Frege (1979), pages 126–51.

Frege, Gottlob (1903). 'On the Foundations of Geometry: First Series', in Frege (1984), pages 273–84.

Frege, Gottlob (1904). 'What is a Function?', In Frege (1984), pages 285–92.

Frege, Gottlob (1906a). 'Introduction to Logic', in Frege (1979), pages 185–96.

Frege, Gottlob (1906b). 'A Brief Survey of My Logical Doctrines', in Frege (1979), pages 197–202.

Frege, Gottlob (1914). 'Logic in Mathematics', in Frege (1979), pages 203–50.

Frege, Gottlob (1918a). 'Thoughts', in Frege (1984), pages 351–72.

Frege, Gottlob (1918b). 'Negation', in Frege (1984), pages 373–89.

Frege, Gottlob (1919). 'Notes for Ludwig Darmstädter', in Frege (1979), pages 253–7.

Frege, Gottlob (1923a). 'Compound Thoughts', in Frege (1984), pages 390–406.

Frege, Gottlob (1923b). 'Logical Generality', in Frege (1979), pages 258–62.

Frege, Gottlob (1950). *The Foundations of Arithmetic*. Evanston: Northwestern University Press. Translated by J. L. Austin; first published as *Die Grundlagen der Arithmetik*. Breslau: W. Köbner, 1884.

Frege, Gottlob (1972). *Conceptual Notation and Related Articles*. Oxford: Clarendon Press. Translated by T. W. Bynum; first published as *Begriffsschrift, eine der arithmetischen nachgebildete Formelsprache des reinen Denkens*. Halle: L. Nebert, 1879.

Frege, Gottlob (1979). *Posthumous Writings*. Chicago: University of Chicago Press. Translated by Peter Long and Roger White; first published as *Nachgelassene Schriften*. Hamburg: Felix Meiner, 1969.

Frege, Gottlob (1980). *Philosophical and Mathematical Correspondence*. Chicago: University of Chicago Press. Translated by Hans Kaal; first published as *Wissenschaftlicher Briefwechsel*. Hamburg: Felix Meiner, 1976.

Frege, Gottlob (1984). *Collected Papers on Mathematics, Logic and Philosophy*. New York: Basil Blackwell. Edited by Brian McGuinness; originally published as *Kleine Schriften*. Hildesheim: Georg Olms, 1967.

Frege, Gottlob (2013). *Basic Laws of Arithmetic. Derived Using Concept-Script*, vol. I and II. Ed. and trans. Philip A. Ebert and Marcus Rossberg. Oxford: Oxford University Press; originally published as *Grundgesetze der Arithmetik*. Jena: H. Pohle, 1893/1903.

Garavaso, Pieranna (1991). 'Frege and the Analysis of Thoughts'. *History and Philosophy of Logic* 12:195–210.

Geach, Peter (1975). 'Names and Identity', in *Mind and Language*, ed. S. Guttenplan. New York: Oxford University Press.

Geach, Peter (1976). Review of Dummett, *Frege: Philosophy of Language*, *Mind* 85:436–49.

Heck, Richard Kimberly (1996). 'The Consistency of Predicative Fragments of Frege's *Grundgesetze der Arithmetik*', *History and Philosophy of Logic* 17:209–20. (Orig. publ. under the name "Richard G. Heck, Jr".)

Heck, Richard Kimberly and Robert May (2011). 'The Composition of Thoughts', *Noûs* 45:126–66. (Orig. publ. under the name "Richard G. Heck, Jr".)

Hodes, Harold T. (1982). 'The Composition of Fregean Thoughts', *Philosophical Studies* 41:161–78.

Klement, Kevin C. (2001). 'Russell's Paradox in Appendix B of the *Principles of Mathematics*: Was Frege's Response Adequate?', *History and Philosophy of Logic* 22:13–28.

Klement, Kevin C. (2002). *Frege and the Logic of Sense and Reference*. New York: Routledge.

Klement, Kevin C. (2003a). 'The Number of Senses', *Erkenntnis* 58:302–23.

Klement, Kevin C. (2003b). 'Russell's 1903–05 Anticipation of the Lambda Calculus', *History and Philosophy of Logic* 24:15–37.

Klement, Kevin C. (2010a). 'Gottlob Frege', in *The Routledge Companion to Nineteenth Century Philosophy*, edited by Dean Moyar. Abingdon: Routledge, pages 858–86.

Klement, Kevin C. (2010b). 'The Senses of Functions in the Logic of Sense and Denotation', *The Bulletin of Symbolic Logic* 16:153–88.

Klement, Kevin C. (2012). 'Frege's Changing Conception of Number', *Theoria* 78:146–67.

Landini, Gregory (2012). *Frege's Notations: What They Are And How They Mean*. London: Palgrave Macmillan.

Myhill, John (1958). 'Problems Arising in the Formalization of Intensional Logic', *Logique et Analyse* 1:78–82.

Parsons, Terence (2001). 'The Logic of Sense and Denotation: Extensions and Applications', in *Logic, Meaning and Computation: Essays in Memory of Alonzo Church*, edited by C. A. Anderson and Michael Zelëny. Dordrecht: Kluwer, pages 507–44.

Reck, Erich and Steve Awodey, eds. (2004). *Frege's Lectures on Logic: Carnap's Student Notes. 1910–1914*. Peru, Ill.: Open Court.

Sluga, Hans (1980). *Gottlob Frege*. London: Routledge and Kegan Paul.