HAROLD T. HODES

THE COMPOSITION OF FREGEAN THOUGHTS

(Received 14 November, 1980)

What sort of entity is a proposition?[1] When are propositions identical? How are they constructed? These and related questions crop up quickly when we begin to theorize about language. The tradition of the last hundred years offers several well-worn answers to these questions.

For example, Carnap[2] isolates a notion of intension according to which two (eternal) sentences express the same intension iff they are L-equivalent. Perhaps we should say that propositions are such intensions. In a similar vein, practitioners of possible-world semantics suggest that two sentences express the same proposition iff their biconditional is a necessary truth. As Carnap was well aware, these proposals individuate propositions rather coarsely – too coarsely, many say, to be the objects of propositional attitudes. Carnap recognized that the pre-theoretic relation of synonymy cut more finely than L-equivalence and seemed to correctly individuate the objects of propositional attitudes. So he proposed to reconstruct the informal notion of synonymy as the 'rigorous' notion of intensional isomorphism restricted to sentences. Supposing that each atomic significant unit is associated with an intension, intensional isomorphism was defined recursively as follows: atomic units are intensionally isomorphic iff they express the same intensions; non-atomic units are intensionally isomorphic iff they are constructed in the same way out of immediate constituents which are appropriately intensionally isomorphic. Obviously this definition is to be taken relative to the syntax of a particular language.

Frege had discussed these issues using his own terms. In spite of terminological differences, we would do well to compare him with Carnap. For according to Michael Dummett[3], the relation obtaining between two sentences when they express the same Fregean thought is, or is very much like, Carnap's notion of intensional isomorphism. Frege considered a sentence to be a complex expression whose sense, the thought it expresses, is determined

Philosophical Studies **41** (1982) 161–178. 0031-8116/82/0412–0161$01.80

by the senses of its constituents. How are we to understand this use of 'is determined by'? Dummett says:

> ...we grasp the sense of a whole sentence by grasping the senses of the constituent expressions, and, of course, observing how they are put together in the sentence.... We might well, of course, associate the same reference with an expression in some other way, but we could not associate the same sense with it without viewing the expression as having a parallel structure. ...The sense of a complex sentence is thus actually composed of the senses of its constituents.[4]

Later he amplifies this:

> Frege's notion of sense, as applied to complex expressions, involves a very narrow criterion of identity. Frege says that the sense of a complex expression, including a sentence, is composed out of the senses of its constituents. 'Composed out of' is a metaphor; but it is used deliberately by Frege to convey something stronger than the nonmetaphorical 'determined by'. The value of a number-theoretic function is determined by the arguments of that function; but the number which is the value can be conceived otherwise than as the value of that function for those arguments. To say that the sense of a sentence is composed out of the senses of its constituent words is to say, not merely that, by knowing the sense of the words, we can determine the sense of the sentence, but that we can grasp that sense only as the sense of a complex which is composed out of parts in exactly that way; only a sentence which had exactly that structure, and whose primitive constituents corresponded in sense pointwise with those of the original sentence, could possibly express the very same sense.[5]

There is, however, ample evidence that Frege denied the view here attributed to him. I shall argue that Frege thought thoughts to be compositionally polymorphous, that a single thought may be built up in different ways out of different constituent senses. A sentence, or an analysis of a sentence, only suggests, or displays, a particular way in which the thought expressed may be composed from simpler senses. Furthermore, I'll show that on the basis of natural assumptions about what incomplete entities exist, to deny that thoughts are compositionally polymorphous involves an unpalatable multiplication of distinct thoughts and of ambiguities in sentences expressing these thoughts. Avoidance of this consequence was a likely motivation for Frege's acceptance of polymorphous composition. Finally, I claim that the polymorphous composition of thoughts is a necessary condition for the joint tenability of two of Frege's central doctrines: that arithmetic truths are validities of logic, and that numbers are objects; in fact, it underlies the entire project sketched in the *Grundlagen*.

The most trivial example of polymorphous composition is provided by double negation. The sign for negation refers to an incomplete entity. If thoughts are not compositionally polymorphous, the occurrence of a negation sign in a sentence requires that its referent be presented by a constituent of the thought which the sentence expresses. So if a sentence '*B*' contains no component referring to that incomplete entity, '*B*' and 'not (not (*B*))' must express distinct thoughts. Actually, this applies to any sentence '*B*'. But in 'Compound thoughts', Frege says: " '(not (not (*B*))' has the same sense as '*B*' ".[6] Was this a slip which, on further consideration, Frege would have retracted? Or does it involve a sort of polymorphousness of a very limited, and so uninteresting, sort? Since a 'Yes' to either question may seem reasonable, I want to consider further consequences of supposing that thoughts are either compositionally unique, or that their polymorphous composition is of the rather mild sort exemplified above.

In what follows, I use 'λ' as a device for predicate abstraction, to clarify the function-argument structures in analyses of sentences. I depart from the conventions of the λ-calculus by permitting the abstraction of genuine two-place predicates of the form '$\lambda x, y \cdot T$'; if 'T' has no free variable other than 'x' and 'y', '$(\lambda x, y \cdot T)(a, b)$' represents a well-formed sentence co-referential with '$T(x/a, y/b)$'. Does this notation violate Frege's dictum that reference to concepts may occur only in complete sentences? If we say that the 'x' in '$\lambda x \cdot$ (x is prime)' is bound, as is usually done, it is tempting to construe:

(0) '$\lambda x \cdot$ (x is prime)' refers to $\lambda x \cdot$ (x is prime)

as a complete, and true, sentence. There is no need to quibble over the words 'bound variable', provided we regard (0) as a non-sentence and '$\lambda x \cdot$ (x is prime)' as an incomplete expression. Of course we should not say that '$\lambda x \cdot$ (x is prime)' considered as such refers to an incomplete entity, but only that in '$\lambda x \cdot$ (x is prime) (2)', that is, in '2 is prime', it does so. I'll occassionally speak loosely, hoping that the reader will not 'begrudge a pinch of salt'.

Frege developed his own notation, essentially the standard notation for a fragment of second order logic, which presumably he considered adequate to the demands of his doctrines. Why then introduce predicate abstraction? In part this is my point: standard notation can represent all needed distinctions within Frege's realm of sense. The distinctions in the realm of reference, for which we'll need predicate abstraction, do not correspond to analogous distinctions in the realm of sense.

Consider these three ways of decomposing the sentence 'John loves Mary' into function and argument:

(I)

John loves Mary

$\lambda x \cdot (x \text{ loves Mary})$ John

$\lambda y \cdot \lambda x \cdot (x \text{ loves } y)$ Mary

(II)

John loves Mary

$\lambda x \cdot (\text{John loves } x)$ Mary

$\lambda y \cdot \lambda x \cdot (y \text{ loves } x)$ John

(III)

John loves Mary

$\lambda x, y \cdot (x \text{ loves } y)$ John Mary

The expressions '$\lambda x \cdot (x$ loves Mary)' and '$\lambda x \cdot$ (John loves x)' are clearly not coreferential; and neither '$\lambda y \cdot \lambda x \cdot (x$ loves $y)$' nor '$\lambda y \cdot \lambda x \cdot (y$ loves $x)$' is co-referential with '$\lambda x, y \cdot (x$ loves $y)$'. Thus the senses expressed by these expressions differ. If thoughts are uniquely composed, the sentence in question must be triply ambiguous, with each analysis displaying a distinct thought which it may be taken to express.

At their second levels, analysis (I) and (II) invoke functions whose values are functions. To my knowledge, Frege nowhere introduces such functions. But he does distinguish analyses (I) and (II) as far as the first level; and this suffices for my point. For example, in the *Begriffsschrift*:

This distinction [viz. between monadic function and argument] has nothing to do with conceptual content; it concerns only our way of looking at it [i.e., our analysis of that content]. In the manner of treatment just indicated, 'hydrogen' was the argument and 'being lighter than carbon dioxide' the function; but we can equally look at the same conceptual content in such a way that 'carbon dioxide' is the argument and 'being heavier than hydrogen' is the function.[7]

In a letter, which the editors of the *Nachlass* suppose was to Anton Marty, dated 1882, Frege recognizes analyses of form (III) and explicitly recognizes the polymorphous composition of the relevant thoughts:

I do not believe that for any judgeable content there is only one way in which it can be decomposed, or that one of these possible ways can claim objective preeminince. In the inequality $3 > 2$ we can regard either 2 or 3 as the subject. In the former case we have the concept 'smaller than 3', in the latter case we have the concept 'greater than 2'. We can also regard '3 and 2' as a complex subject. As a predicate we have the relation of the greater to the smaller.[8]

The grammar of a sentence may suggest one analysis of the thought expressed; in fact, the speaker may intend this.

The speaker intends the subject to be taken as the principle argument; the next in importance often appears as the object. Language has the liberty of arbitrarily presenting one or another part of the proposition [recte: sentence] as the principle argument by a choice between inflextions and words...[9].

But such intentions are not relevent to the thought the speaker intends to express.

Passivization and other forms of conversion of binary predications raise similar problems. Analyzing 'Mary is loved by John' in the pattern exemplified by (III), we obtain:

(IV)

Mary is loved by John

$\lambda x, y \cdot (x$ is loved by $y)$ Mary John

But '$\lambda x, y \cdot (x$ is loved by $y)$' is not co-referential with any of the predicates displayed in analyses (I)–(III). In numerous places Frege insists that passivization preserves sense. Thus 'John loves Mary' and 'Mary is loved by John' express the same thought. If thoughts have unique composition, (IV) displays a fourth thought which these sentences may be taken to express.

These passages predate Frege's recognition of the distinction between sense and reference.[10] So his talk of 'conceptual contents' or of 'concepts' involves equivocation between the realms of sense and reference. The previous remarks should be taken as applying to the former realm. In 'On concept and object' (p. 47) Frege confesses that the sense-reference distinction requires a reformulation of some earlier material. But there is no reason to think that he'd retract any of the claims just presented. In fact, a few paragraphs later, he goes on to say:

...a thought can be split up in many ways, so that now one thing, now another, appears as subject and predicate. The thought itself does not yet determine what is to be regarded as the subject. If we say 'the subject of this judgment', we do not designate anything definite unless as the same time we indicate a definite kind of anlysis; as a rule we do this in connexion with a definite wording.[11]

Does a thought whose expression requires use of a quantificational expression have a unique composition? A passage from the *Begriffsschrift* may seem to give an affirmative answer:

We attach no importance to the various ways that the same conceptual content may be regarded as a function of this or that argument, so long as function and argument are completely determinate [i.e. so long as the function and argument are, or are represented by, bound variables]. But if the argument becomes indeterminate, as in the judgment:

'whatever arbitrary positive integer you may take as an argument for "being representable as the sum of four squares", the proposition remains true', then the distinction between function and argument becomes significant as regards the content. ...the whole proposition splits up into function and argument as regards its own content, not just as regards our way of looking at it.[12]

This passage is obscure. Frege did not yet see quantificational expressions as referring to level-two concepts. Consequently he describes bound variables as indeterminate subjects of predication, a confusion for which he later faults other writers. But the point made here seems to be only this: in quantified sentences, the variable bound by a given quantifier cannot be moved around without changing the thought expressed. 'Everyone loves John' cannot be rephrased so that the quantifier binds they 'y' in 'x loves y'. This passage does not indicate that the following are unacceptable analyses of that sentence:

$$(\forall x)\, x \text{ loves John}$$
$$\lambda X^1 \cdot \forall(X^1) \qquad \lambda x \cdot (x \text{ loves John})$$

$$(\forall x)\, x \text{ loves John}$$
$$\lambda y \cdot ((\forall x)\, x \text{ loves } y) \qquad \text{John}$$

In 'On concept and object'. he addresses this question:

For example, the thought we are considering [that expressed by 'There is at least one square root of 4'] could also be taken as asserting about the number 4: 'The number 4 has the property that there is something of which it is the square', Language has means of presenting now one, now another, part of the thought as the subject; one of the most familiar is the distinction between active and passive forms. It is thus not impossible that one way of analysing a given thought should make it appear as a singular judgement; another, as a particular judgement; and a third, as a universal judgement.[13]

For the moment, let's focus on the problem posed by analyses (III) and (IV). Insisting on the unique composition of thoughts, we could say that they lay out the composition of two distinct thoughts; or we could say that they lay out two ways of composing a single thought. Are these the only alternatives? An opponent might suggest that there is a third alternative: that these two trees really represent a single analysis of a single thought, and that 'John loves Mary' and 'Mary is loved by John' involve reference to only one binary relation. Is that relation the loves relation or the is-loved-by relation? Is it the referent of '$\lambda x, y \cdot (x \text{ loves } y)$' or of '$\lambda x, y \cdot (x \text{ is loved by } y)$'? The suggestion continues: "This question is ill-formed and ill-motivated. A binary relation is like a sharpened pencil with an eraser at the other end. We may view the pencil in two orientations: with the sharpened point on the left and the eraser on the right, or vice versa. Similarly, in the sentences we are considering, a single relation is presented in different orientations; there is a unique sense in the thought in question which presents a relation, and it presents a unique relation. By virtue of the syntax of the sentence, that sense may present that relation in different orientations." We might reply that it would take different senses to present a relation in different ways. After all, a sense contains the 'mode of presentation' of whatever it presents; and to present something in different orientations is to present it in different ways. This reply assumes gratuitiously that ways of presentation are individuated very finely by the notion of sense. Perhaps the Fregean should allow certain differences in the way in which an entity is presented to be irrelevant to sense.

To see the difficulty with this approach, let's press our opponent to explain what is ill-formed about our question. '$\lambda x, y \cdot (x$ loves $y)$' and '$\lambda x, y \cdot (x$ is loved by $y)$' are definitely not co-referential. So the relation in question is either the first, the second, or neither. This approach only offers us an appealing metaphor about pencils, not the diagnosis of an unreasonable request. To answer 'Both' is to choose polymorphous composition; to choose one relation over another is to suppose that analyses (III) and (IV) are not on a par. Now syntacticians have argued that the active and passive voices are not on a par, that 'Mary is loved by John' is derived from 'John loves Mary' and not vice versa. Making this supposition, one might claim that both sentences involve reference to the referent of '$\lambda x, y \cdot (x$ loves $y)$' rather than that of '$\lambda x, y \cdot (x$ is loved by $y)$'. Of course Frege has nothing like this in mind, as the previous quotation from the letter to Marty shows. This approach gains implausibility if we consider cases involving forms of conversion different from passivization, e.g., 'Hydrogen is lighter than carbon dioxide'.

How much polymorphous composition must Frege permit? At first glance, a sentence such as 'Socrates is wise' may be analyzed in only one way: into the proper name 'Socrates' and the predicative part '$\lambda x \cdot (x$ is wise)'. But, by that maneuver which construes quantification over objects as involving the application of a level two concept to a level-one concept, Frege is also committed to a curious reduplication of simple concept-object predications into all levels of the type hierarchy. For example, 'Socrates is wise' may be analyzed as containing an expression for the Socrates quantifier, the result of 'removing' '$\lambda x \cdot (x$ is wise)' from 'Socrates is wise.' We will write this as '$\lambda X^1 \cdot$

X^1 (Socrates)'.[14] We now have two analyses:

(V)

Scocrates is wise

$\lambda x \cdot (x \text{ is wise})$ Socrates

(VI)

Socrates is wise

$\lambda X^1 \cdot X^1 (\text{Socrates})$ $\lambda x \cdot (x \text{ is wise})$

Does this show that the sentence 'Socrates is wise' really expresses two possible thoughts, one containing a sense which presents the person Socrates, the other containing a sense which presents a level two concept intimately connected with that person? If so, we should actually admit to an infinite ambiguity; for we may repeat this pattern, associating with the referent of '$\lambda x \cdot (x$ is wise)' a level-three concept and so forth.

We may think of the entities of higher types which appear in these analyses as follows. Socrates has a 'shadow' in each even level above level one; the referent of '$\lambda x \cdot (x$ is wise)' casts a 'shadow' in each odd level above level one. Given any analysis of the sort under consideration, it will display two constituents, a function and an argument, one of even level and the other of odd level; there is a natural one-one correspondence between these constituents and these shown by analysis (V), matching concepts of even level with the referent of '$\lambda x \cdot (x$ is wise)', and those of odd level with Socrates. This suggests a novel way to avoid an infinity of distinct thoughts, while preserving the thesis that thoughts have unique compositions. We could say that both 'Socrates' and '$\lambda X^1 \cdot X^1$ (Socrates)', though syntactically quite different, express the same sense. But, and here we depart from the Fregean letter, senses do not present unique entities. Rather, they primarily present one entity, secondarily another entity of different type, and so forth. So the sense expressed by 'Socrates' primarily presents a person, secondarily an entity of level two, etc. The tree of senses determined by the thought expressed by 'Socrates is wise' may be unique; there is no unique corresponding tree of referents, but rather an ordered infinitude of such trees. This cost may seem slight.

However, a sentence like 'Socrates is wise' may be analyzed so as to display three rather than two constituents. That sentence may be rephrased as 'Socrates falls under the concept *wisdom*', suggesting an analysis in which an expression standing for a binary relation is completed by two further expressions. Frege recognizes such analyses:

> When we say 'Jesus falls under the concept *man*', then, setting aside the copula, the predicate is:
> 'someone falling under the concept *man*'
> and this means the same as:
> 'a man',
> But the phrase
> 'the concept *man*'
> is only part of this predicate.[15]

Of what type is this binary relation? Frege claims that expressions such as 'the concept *wisdom*' refer to objects. In this case, a reasonable candidate would be the class of all wise objects; understood in this way, we read 'falls under' as 'is a member of'. The type of this relation is (0,0). Alternatively, we may take the relation in question to have type (0, 1), and to be applied to Socrates and the referent of '$\lambda x \cdot (x$ is wise)', in that order. Representing the latter relation as '$\lambda x, X^1 \cdot F(x, X^1)$, we may parse 'Socrates is wise' as 'F(Socrates, $\lambda x \cdot (x$ is wise))'. Speaking loosely, this relation is the 'real' falling-under relation between types 0 and 1. This relation has its obvious analogues of types $(n, n + 1)$, as Frege was well aware:

> Second-level concepts, which concepts fall under, are essentially different from first-level concepts, which objects fall under. The relation of an object to a first-level concept that it falls under is different from the (admittedly similar) relation of a first-level to a second-level concept.[16]

To countenance the falling-under relation is not to suppose that the copula, or any syntactic constituent of 'Socrates is wise', refers to it. Nonetheless,

(VII)

Socrates is wise

$\lambda x, X^1 \cdot F(x, X^1)$ Socrates $\lambda x \cdot (x$ is wise)

is an analysis distinct from (V).

We now may see the failing of the last-mentioned approach: that the thought in question contains two immediate constituent senses, one primarily presenting Socrates, secondarily presenting a level two concept, etc., the other primarily presenting a level one concept, etc. The point is that there is no third sense which we could suppose presents the referent of '$\lambda x, X^1 \cdot F(x, X^1)$'. Analysis (VII) identifies three constituent senses of the same level, showing that the dodge suggested three paragraphs ago does not avoid attributing polymorphous composition to the thought in question.

We will now look at the role of polymorphous composition in Frege's doctrines about number. In the *Grundlagen* Frege claims that an attribution of number may be viewed as a statement about a level one concept. 'There are four moons of Jupiter' may be construed as the completion of a numerical quantifier expression by a predicate:

(1) $(\exists ! 4x)(x \text{ is a moon of Jupiter})$.

This is turn may be expressed within the usual sort of first-order quantificational language with identity:

(2) $(\exists x)(\exists y)(\exists z)(\exists u)(x \text{ is a moon of Jupiter } \& \ldots \& \, x \neq y \, \& \ldots \& \, (\forall v)(v \text{ is a moon of Jupiter} \supset (v = x \text{ or} \ldots)))$.

But the use of attributive constructions "can always be got around".[17] That is the same thought may be expressed by an equation:

(3) The number of moons of Jupiter = 4.

Frege then turns to equations such as:

(4) The number of moons of Jupiter = the number of burners on my stove.

This sentence is analyzable into two proper names and an expression for the identity relation. But it may also be analyzed as a statement that the type (1, 1) relation of equinumerosity holds between two level one concepts. Letting '$\lambda X^1, Y^1 \cdot E(X^1, Y^1)$' represent that relation, we may rephrase (4) as

(5) $E(\lambda x \cdot (x \text{ is a moon of Jupiter}), \lambda x \cdot (x \text{ is a burner of my stove}))$.

(Ultimately, Frege assimilates (3) to the form of (4) by assigning to '4' the sense of 'the number of natural numbers less than or equal to 3'.) He then goes on to show how the thought expressed by (5) may be expressed within a second-order quantificational language as

(6) $(\exists X^{(1,1)})\, A(X^{(1,1)}, \lambda x \cdot (x \text{ is a moon of Jupiter}), \lambda x \cdot (x \text{ is a burner on my stove}))$,

where '$A(X^{(1,1)}, X^1, Y^1)$' spells out '$X^{(1,1)}$ is a one-one relation between all the X's and all the Y's'.

With each of these two examples I have introduced some novel notation to point out three analyses of a single thought. These analyses are ordered with respect to a sort of conceptual priority. Imagine an ideal Fregean mathematics student, Sally, whose cognitive development reflects the order of Frege's definitions. She begins with an understanding of a second-order quantificational language. Thus she can grasp the thoughts expressed by (2) and (6). Regarding (2) and (6) as contextual definitions of '$(\exists!\,4x)$' and '(Ex)' respectively, she 'subtracts' the sense of '$\lambda x \cdot (x$ is a moon of Jupiter$)$' from the thought expressed by (2), to grasp the sense expressed by '$(\exists!\,4x)$' in (1); and by a similar trick, she grasps the sense of '(Ex)' in (6). This way of coming to grasp senses may be likened to solving an equation in one variable, that is, going from the value of a function to the argument at which the function takes that value.[18] Elsewhere Frege speaks of carving up "the content in a way different from the original way, and this yields up a new concept".[19] This sort of stunt relies on the polymorphous composition of the thoughts involved.

But in the *Grundgesetze* didn't Frege explicitly reject contextual definition?

Given the reference of an expression and of a part of it, obviously the reference of the remaining part is not always determined. So we may not define a symbol or word by defining an expression in which it occurs, whose remaining parts are known. For it would first be necessary to investigate whether – to use a readily understood metaphor from algebra – the equation can be solved for the unknown, and whether the unknown is unambiguously determined.[20]

We must revise our picture of Sally's cognitive gymnastics. Pressing the algebraic metaphor, she must solve an infinite system of simultaneous equations in one unknown which, in fact, has a unique solution. To grasp the sense of '$\lambda X^1 \cdot (\exists!\,4)(X^1)$', she must also 'subtract' the sense of '$\lambda x \cdot (x$ is a burner on my stove$)$' from the thought that there are four burners on my stove, etc. If the definiens is incomplete, the sort of definition which Frege espouses is a schema representing the remains of a new way of breaking up an infinitude of thoughts with which se started:

Of course names of functions, because of their characteristic 'unsaturatedness', cannot stand alone on one side of a defining equation; their argument-places must always be filled up somehow or other. In my ideography, as we have seen, this is done by means of italic letters, which must then occur on the other side as well. In language, instead of these, there occur pronouns and particles... which indicate indefinitely.[21]

For example, the definitions we want may be easily provided:

$$(\exists!\, 4x)\, X^1 \equiv (\exists x)(\exists y)(\exists z)(\exists u)(X^1\ (x)\ \&\ldots\&\ x \neq y\ \&\ldots\&$$
$$(\forall v)(X^1\ (v) \supset (x = y \text{ or}\ldots)));$$
$$(Ex)(X^1, Y^1) \equiv (\exists Z^{(1,1)})\, A\, Z^{(1,1)}, X^1, Y^1).$$

Frege goes on to try to lead Sally through yet more spectacular cognitive gymnastics. Sections 62–65 of the *Grundlagen* are motivated by the question:

How, then, are numbers to be given to us, if we cannot have any ideas of them?[22]

This is the central question for numerical realists. For whether one thinks of reference in modern 'causal' terms, or, like Frege's contemporaries, in terms of 'imagination' and 'having ideas', it may seem strange that objects as abstract as numbers should ever fall within our referential ken. This concern motivates Sections 58–61.[23] Reformulating Frege's question as "How are we to grasp the senses of expressions which refer to numbers?", we see that Frege's answer is again "By contextual definition". The route to this achievement will run through a prior grasping of the thought expressed by sentences like (4). In Section 62 Frege says:

> In our present case, we have to define the sense of the proposition
>
> 'The number which belongs to the concept *F* is the same as that which belongs to the concept *G*',
>
> that is to say, we must reproduce the content of this proposition in other terms, avoiding use of the expression
>
> 'The number which belongs to the concept *F*'.[24]

In the next section:

> Our aim is to construct the content of a judgement which can be taken as an identity such that each side is a number.

The construction in question is his definition of equinumerosity; the 'taking as' is the step from grasping the thought expressed indifferently by (4), (5) and (6), to grasping the senses of 'The number of moons of Jupiter' and 'The number of burners on my stove'; it is the analogue within the realm of sense the familiar algebraic maneuver of factoring by an equivalence relation. Sally already grasps the thought expressed by (4) and the sense of '='; she divides the former by the latter and ends up grasping the senses of two number-words. Of course we must think of the feat as performed simultaneously for all sentences of the form 'The number of F's = the number of G's'. It will not do

> ...to define two things [*recte*: expressions] with one definition. ...One equation alone cannot be used to determine two unknowns.[25]

But we can sometimes solve an infinitude of equations for an infinitude of unknowns. So perhaps this sort of 'backwards' definition does fix the senses of number-words, in spite of its oddity:

> We are therefore proposing not to define identity specially for this case, but, taking the concept of identity as already known, to arrive by its means at that which is to be regarded as identical. Admittedly, this seems to be a very odd kind of definition.[26]

Notice that this odd kind of definition requires that the thought expressed by (4) and (5) be composed in two quite different ways: by the senses of two predicate expressions and a sense presenting the equinumerosity relation; and by the senses of two names and a sense presenting the identity relation. If we follow Dummett in maintaining that

> ...we can grasp that sense only as the sense of a complex which is composed out of parts in exactly that way,

then the senses of the names in question could not be fixed in this roundabout manner. Indeed, the thought (4) expresses could not be grasped without first grasping the senses of the number-names (4) contains.

Does this show that the polymorphous composition of thoughts is central to Frege's view of arithmetic? The previous considerations may be less impressive when we recall that in Section 66 Frege concludes that the previous account falls short of its goal. Division in the realm of sense does not fully fix the sense (or fix the full sense) of either 'The number of moons of Jupiter' or 'The number of burners on my stove'. If an expression has a determinate sense, any well-formed sentence in which all other constituents have determinate sense will itself express a determinate thought. But Frege says that we have not yet fixed the thought expressed by

(7) The number of moons of Jupiter = England.

'The number of moons of Jupiter' still has a sense-gap which is not manifested in (4) but which appears in (7). In Section 68 Frege claims to close this gap by "introducing extensions of concepts into the matter", defining 'The number of moons of Jupiter' as: the extension of the predicate 'the X^1's are equinumerous with the moons of Jupiter'.

Of course the same problem now arises for names of extensions, or more generally, names of courses of value for functions: how can instances of the axiom schema of extensionality fix determinate senses for those names? In the *Grundgesetze*, Section 10, Frege's reconstructed language only contains names of courses of value and of truth values. He identifies the truth values with certain courses of value, and then claims that all equations *are* between names for courses of value. This shows Frege's notion of sense to be distinctively holistic (in a way which Dummett seems to rule out): for if we extend the language by adding even a single name whose sense has been somehow fixed first, we can make the previously determinate senses of the old names non-determinate. For example, since 'England' presumably does not refer to a course of value,

(7′) $\{x \mid x \text{ is a moon of Jupiter}\} = \text{England}$,

would not have a determinate sense in the extended language. This certainly suggests that the introduction of extensions in the *Grundlagen* does not really succeed in closing even the sense gap displayed by (7); for the background language is meant to be an appropriately cleaned-up version of the language we really use.

Rather than pursue this interesting tangent, I return to the main thread: does the negative conclusion of Section 66 show that Frege does, or should, take (4) and (5) to express distinct thoughts? If so, we must choose between three accounts. (i) Contrary to the first impressions on which we were relying previously, before the final definition of number words in terms of extensions, (4) involved as much as a sense gap as (7). (ii) (4) and (5) do express the same thought; but contrary to the suggestion carried by the syntactic form of (4), that thought cannot be analyzed as involving reference to numbers, where these are taken to be objects. (iii) Before our adoption of the final definition, (4) expressed a unique thought identical with the one expressed by (5); afterwards (4) is ambiguous between the former thought and one which does involve reference to numbers.

As an interpretation of Frege, (iii) seems to be obviously unreasonable. (ii) may seem appealing in view of Frege's belief that the analysis suggested by the syntax of (6) is prior to that suggested by (5), and that that is in turn prior to the analysis suggested by the syntax of (4). Furthermore, this priority is objective, part of the "true order of things", as Frege makes clear when considering the analogous case concerning directions of lines. He considers defining 'parallel' in terms of directions, and says:

Only the trouble is, that this is to reverse the true order of things. For surely everything geometrical must be given originally in intuition. But now I ask whether anyone has an intuition of the direction of a line. ...The concept of direction is only discovered at all as a result of a process of intellectual activity which takes its start from intuition.[27]

(ii) supposes that the maximally prior analysis is ontologically privileged. If Frege's project is to provide definitional extensions of an underlying formalization of second-order logic this would be right. But then number names would really be non-referring, as Russell supposed definite descriptions to be. Frege's insistence on the objecthood of numbers belies this interpretation. The *Grundlagen* was intended to show how number-words manage to have determine sense and thereby refer to numbers, not how they may be construed as eliminable abbreviations, like our talk of sakes.

We are forced back to (i). But (i) is contrary to the underlying motivation for logicism. Let's step back from the details of Frege's project and consider why one might suppose that the truths of arithmetic are really validities of logic. Consider '2 + 5 = 7'. What thought does this express? Roughly: if one takes two objects and then five more objects, one has seven objects. Using numerical quantifiers, which may be construed as abbreviations, this expresses a valid statement of second-order logic:

$$(8)\qquad (\forall X^1)(\forall Y^1)((\exists!\,2x)\,X^1\,(x)\,\&\,(\exists!\,5x)\,Y^1\,(x)\,\&\,-(\exists x)(X^1\,(x)\,\&\,Y^1\,(x)))\supset(\exists!\,7x)(X^1\,(x)\text{ or }Y^1\,(x))).$$

This might suggest that in '2 + 5 = 7' the numerals are not genuine names. As just emphasized, Frege robustly resists this conclusion. This conclusion is resistable only if the thought (8) expresses is composed in at least two ways: the way displayed in (8), and the way suggested by the syntax of '2 + 5 = 7'. Logicism with numerical realism requires a non-trivial sort of polymorphous composition for arithmetic truths. Thus Fregean thoughts are not individuated by a relation remotely like Carnap's intensional isomorphism.[28]

This is not to imply that are individuated by a relation like L-equivalence or necessary equivalence. It is central to the viewpoint of the *Grundlagen* that many statements about number be analytically true, in fact provable, when correctly reconstructed in the appropriate formal system. Thus the biconditionalization of two such sentences is also analytically true and provable. Nonetheless, Frege clearly thinks that analytically true sentences may express distinct thoughts. Why do '$2 + 5 = 7$' and 'not (not ($2 + 5 = 7$))' express the same thought, whereas '$2 + 5 = 7$' and '$3 + 5 = 8$' do not? To point out that this is pretheoretically reasonable is not to answer the question. Frege is committed to the existence of a line between these sorts of cases. His notion of a thought, of the content of a judgment and of other sorts of propositional attitudes, is the pre-theoretical everyday notion which is so deeply engrained in so-called common-sense psychology. He does not draw this line; in fact he is tremendously vague about its location.

Perhaps ordinary talk about thoughts and propositional attitudes, and philosophical talk which employs these pre-theoretic notions, involve divided reference, between Carnapian or possible-world intensions and entities individuated by a relation like intensional isomorphism. Diagnoses of divided reference can often explain difficulties faced and things said by theorists who employ the terms which divide their reference.[29] Perhaps the Fregean notion of a thought is a hybrid, born of confusions created by divided reference.

On the other hand, perhaps the Fregean notion of a thought is on the right track; but we need to know more about actual information-processing within human beings in order to get at the way of individuating the contents of propositional attitudes implicit in our everyday scheme of things. For this rain-check to be redeemed, common sense phychology would have to be remarkably right about matters which are highly theoretical. In this case, the theory would be about the theorizers, so perhaps (only perhaps!) this would not be a miracle. In any case, it is the option on which the Fregean ought to place his money. For it is not a psychologistic option, in the pejorative sense in which Frege used the word. Individuating thoughts in terms of the cognitive workings of human beings would show that the notion of a thought has a certain relativity to particular ways of processing information. But it would not compromise the objectivity of the thoughts themselves, that is, make them into ideas, any more than the institutional basis of our individuation of nations compromises their objectivity.[30]

Cornell University

NOTES

[1] I use the word 'proposition' in what has become its standard philosophical sense. Propositions are expressed by sentences, at least in contexts; they are preserved by correct translations, and are the objects of belief, doubt, etc. And perhaps they do not exist.

[2] Meaning and Necessity (University Press, Chicago, 1966).

[3] Frege: The Philosophy of Language (Duckworth, London, 1973). See p. 228 for a partial disclaimer.

[4] *Ibid.*, pp. 152–153.

[5] *Ibid.*, pp. 378–379. I first heard this interpretation challenged by Warren Goldfarb, in 1975.

[6] See Essays on Frege (University of Illinois Press, Urbana, Chicago, and London, 1968), p. 548. Dummett does seem to recognize this point, see *op. cit.*, p. 325.

[7] Geach and Black (ed.), Translations from the Philosophical Writings of Frege (Basil Blackwell, Oxford, 1970), p. 12.

[8] Gabriel, Hermes (ed.), The Philosophical and Mathematical Correspondences of Frege (University of Chicago, Chicago, 1980), p. 101.

[9] Geach and Black, *op. cit.*, p. 14.

[10] There is a striking adumbration of that distinction in Section 8 of the *Begriffsschrift.*

[11] *Ibid.*, p. 49.

[12] *Ibid.*, p. 14.

[13] *Ibid.*, p. 49.

[14] A word about my use of variables and type-symbols. 0 is the type of objects. Where $\sigma_1, \ldots, \sigma_n$ are types, $(\sigma_1, \ldots, \sigma_n)$ is the type of n-adic relations whose ith place is of type σ_i. $\ulcorner n + 1 \urcorner$ abbreviates $\ulcorner (n) \urcorner$. Lower case variables range over type 0; upper case variables superscripted with a type-symbol range over the type represented by that sumbol.

[15] *Ibid.*, p. 47.

[16] *Ibid.*, p. 51.

[17] Frege, The Foundations of Arithmetic, transl. by J. Austin, (Harper & Row, New York), p. 69.

[18] Of course it is even more like 'extracting' a function from its graph.

[19] Frege, *op. cit.*, p. 75.

[20] Geach and Black, *op. cit.*, p. 170.

[21] *Ibid.*, p. 171.

[22] Frege, *op. cit.*, p. 73.

[23] Frege diagnoses the source of this feeling of strangeness as inattention to this dictum: "Only in a proposition have the words really a meaning". The sort of primacy which Frege here attributes to sentences seems to be connected, though in ways which Frege never makes clear, to the polymorphous composition of thoughts, and also to the incompleteness of incomplete expressions. Consider this passage from the previously excerpted letter to Marty:

> Now I do not believe that concept formation can precede judgements because this would presuppose the independent existence of concepts, but I think of a concept as having arisen by decomposition from a judgeable content. I do not believe that for any judgeable content there is only one way in which it can be decomposed... .

Throughout this letter, confusion between senses which present concepts and concepts themselves runs rampant. But Frege is here clearly concerned with the composition of thoughts. As Dummett has pointed out, in Frege's later writing the previously quoted

dictum disappears, although he does not abandon the thesis that numbers are objects.
[24] Frege, *op. cit.*, p. 73. On the next page, Frege discusses the definition of directions of lines in terms of the relation of paralelism, represented as '//':

> Thus replace the symbol // by the more generic symbol =, through removing what is specific in the content of the former and dividing it between *a* and *b*. We carve up the content in a way different from the original way, and this yields as a new concept. (pp. 74–75)

The concept in question is of a direction of a line.
[25] Geach and Black, *op. cit.*, p. 171.
[26] Frege, *op. cit.*, p. 74.
[27] *Ibid.*, p. 75.
[28] Why does Dummett adopt the interpretation of Frege just criticized? My guess is this: it follows from the account of what it is to grasp a sense which he attributes to Frege. To my knowledge, Frege says very little about the latter issue; he certainly offers us nothing as well developed as what Dummett attributes to him. The conclusions of this paper suggest that Dummett's account should not be attributed to Frege. Perhaps it is overinterpretation to attribute to Frege any account of this matter.
[29] For a nice example of this, see Philip Kitcher's 'Theories, theorists and theoretical change', Philosophical Review (October, 1978).
[30] One might object to this comparison as follows: nations, unlike thoughts, would not exist if people did not exist; but according to this alternative, without people there are no thoughts. I think that this last conclusion does not follow. Ways of doing things exist even if things are never done in those ways.