

Carving Content at the Joints

STEPHEN YABLO

To cite this article: STEPHEN YABLO (2008) Carving Content at the Joints, Canadian Journal of Philosophy, 38:sup1, 145-177, DOI: [10.1353/cjp.2011.0038](https://doi.org/10.1353/cjp.2011.0038)

To link to this article: <https://doi.org/10.1353/cjp.2011.0038>



Published online: 01 Jul 2013.



Submit your article to this journal [↗](#)



Article views: 40



View related articles [↗](#)



Citing articles: 1 View citing articles [↗](#)

*Carving Content at the Joints*¹

STEPHEN YABLO

1. The Problem

Here is Frege in *Foundations of Arithmetic*, §64:

The judgment 'Line a is parallel to line b ', in symbols: $a \parallel b$, can be taken as an identity. If we do this, we obtain the concept of direction, and say: 'The direction of line a is equal to the direction of line b .' Thus we replace the symbol \parallel by the more generic symbol $=$, through removing what is specific in the content of the former and dividing it between a and b . We carve up the content in a way different from the original way, and this yields us a new concept. (Frege 1997, 110–11)

Something important is going on in this passage. But at the same time it borders on incoherent. For Frege is saying at least the following:

- (1) ' $\text{dir}(a) = \text{dir}(b)$ ' has the same content as ' $a \parallel b$ '
- (2) reflecting on that can lead one to the concept of direction.

Why doesn't (2) contradict (1)? (2) has a neophyte acquiring the concept of direction – and so presumably a grasp of the content of ' $\text{dir}(a) = \text{dir}(b)$ ' – by reflecting on a certain content-identity. But then it is hard to see how the postulated content-identity can really obtain; Leibniz's Law would seem to forbid it. If one grasps content X at a certain time, and content $X =$ content Y , then one grasps content

1 This paper was originally presented at the Arché conference 'Does Mathematics Require a Foundation?' (August 2002). Thanks to Bob Hale, Crispin Wright, Kit Fine, Neil Tennant, John Burgess, Gideon Rosen and Matti Eklund for very useful comments.

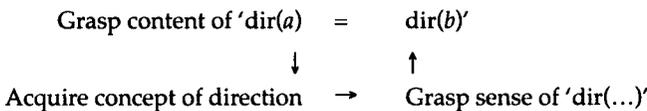
Y at that time. The neophyte grasped the content of ' $a \mid b$ ' before encountering (1), so if that content is also the content of ' $\text{dir}(a) = \text{dir}(b)$,' she must have grasped the content of ' $\text{dir}(a) = \text{dir}(b)$ ' before encountering (1) as well.

I know of only one good way of getting around this. The neophyte *did* grasp the content of ' $\text{dir}(a) = \text{dir}(b)$ ' before encountering (1); she just failed to know it *as the content of an identity-sentence*. She doesn't know it as the content of an identity-sentence until she acquires the concept of direction: perhaps knowing it that way is acquiring the concept of direction.

2. Sense

What should content be, for this way around the problem to work? One natural hypothesis is that *content is sense*; and Frege certainly says things that suggest this. But the suggestion is problematic, if we take Frege at his word that the sense of part of a sentence is part of the sense of the whole.

Remember, the neophyte has to grasp the content of ' $\text{dir}(a) = \text{dir}(b)$ ' *before* acquiring the concept of direction. So if content is sense, she must be able to grasp the *sense* of ' $\text{dir}(a) = \text{dir}(b)$ ' before acquiring the concept of direction. If she lacks the concept of direction, though, how is she supposed to grasp the sense of direction-terms? And if she does not grasp the sense of direction-terms, how is she supposed to grasp the sense of ' $\text{dir}(a) = \text{dir}(b)$ '? The problem is that each of the indicated achievements presupposes the one 'before' it:



Frege's strategy does not appear to work, then, if content is sense. What else could it be?

The downward-facing arrow is compulsory, for the passage clearly states that the concept of direction is acquired by grasping the content of the direction-sentence. The left-to-right arrow is compulsory too, for grasping the sense of 'direction-of' is appreciating that it expresses the relevant concept. The upward-facing arrow is not forced on us,

though. Why shouldn't grasping the content of a sentence leave one still undecided about its sense?

The obvious way to arrange for that is to make content *coarser-grained* than sense, though presumably still finer-grained than reference. Then everything hangs together just right:

- (1) ' $\text{dir}(a) = \text{dir}(b)$ ' shares *something* with ' $a \parallel b$,' but
- (2) the shared something is content, not sense;
- (3) the shared content can be carved in two ways,
- (4) corresponding to the sentences' two senses;
- (5) we start out knowing one carving, then learn the other;
- (6) the directional carving teaches us the concept of direction.

This has got to be the way to go. But every step raises questions, the main ones being (a) what is content? and (b) what are carvings? I want to sketch a conception of carving that deals with some of these questions in a way broadly congenial to the Fregean platonist program in the foundations of mathematics.

3. Sketch and Motivation

I can't claim *too* much of a Fregean basis for the proposal to follow. I would, however, like to relate the proposal to what are popularly regarded as some Fregean concerns and themes.

One such theme is that semantic theory begins with sentences and their truth values. Words and other subsentential expressions have their semantic values too; but these are constrained mainly by the requirement that they predict the expression's contribution to truth-value. Semantic values are whatever they have to be to deliver the right sentence-level results. One gets from the one to the other by a kind of abduction or inference to the best explanation.

A second theme concerns logical structure or form. There is no way of abducting semantic values that doesn't take a stand on the sentence's logical form. Without the form, one doesn't know what types of semantic value are needed (object, truth-function, m -place concept, etc.), and one doesn't know by what mode of combination they are supposed to deliver the truth-value. A given bunch of semantic values – say, Brutus, Caesar, and the relation of killing

– might yield various truth-values depending on how they are combined.

Now, given what was said about semantic *values* being slaves to truth-value, it might seem that the *structures* the values are slotted into should likewise be whatever best serves the needs of truth-value prediction. A certain type of sentence may *look* atomic (or whatever), as ‘The dodo bird still exists’ looks atomic. But if it proves hard to predict truth-values on that hypothesis, we may decide that its real, underlying, structure, is something quite different.

This leads directly to a third theme. I said you might *expect* Frege to make structure a slave to truth-value prediction. But you’d be wrong, or anyway not entirely right. He certainly does want to impute structures that generate the appropriate truth-values. But the enterprise is subject to a heavy constraint imposed by the actual words employed and how they are ordered.

This is just the familiar point that Frege tries very hard – much harder than Russell, to make the obvious comparison – not to run roughshod over grammatical appearances. He is concerned to understand the sentence as he finds it. Semantic structures ought if possible to *parallel* sentential structure, even if it takes lots of theoretical work to find a parallel semantic structure that does the job. Call that Frege’s *empiricism about structure*.

To review, our three main players are *truth-value*, *semantic structure*, and *semantic value*. Of these three, the *first* is treated as given, and the *second and third* are reached abductively, subject to the constraint that semantic structures should not be imputed that run roughshod over the grammatical appearances.

This idea of respecting grammatical appearances is fair enough in itself, but it creates a tension in Frege’s system. What if we spot an alternative, not entirely parallel, semantic structure, that works better somehow than the structure we were initially inclined to impute? (I propose to be vague for the moment about what *better* might mean here. The alternative structure predicts truth-values more efficiently, perhaps, or in a way that illuminates entailment relations with other sentences, or ...).

If we are Frege, we cannot make like Russell and adopt this alternative structure despite its grammatical implausibility. But we also don’t want to just wave the alternative structure goodbye. Because

while, on the one hand, we find fault with it for not running parallel enough to the sentence's apparent grammatical form, at a deeper level we may feel that it is *the sentence's fault* for not running parallel enough to this excellent structure. That we employ the sentence we do, and not one with the alternative structure, is *too bad* in a way. There would be *advantages* to using a sentence of the second sort rather than the one we do use.

What is the Fregean to do when this sort of feeling comes over him? I want to suggest that this is where the need for alternative carvings arises. Frege's *empiricism* (about structure) tells him that *S*'s truth-value is generated in one way. His *rationalism* tells him that it might better have been generated another way, the way *S**'s truth-value is generated. Content (re)carving gives Frege an outlet for the second feeling that lets him stay true to the first. The conflict is resolved by noting that *S*'s content can also be carved the *S** way, followed perhaps by a recommendation that *S** be used instead of *S* in situations where, as Quine later put it, greater theoretical profundity is professed.

Consider in this light the discussion in paragraph 57 of adjectival versus singular statements of cardinality. 'Since what concerns us here is to define a concept of number that is useful for science, we should not be put off by the attributive form in which number also appears in our every day use of language. This can always be avoided.... the proposition "Jupiter has four moons" can be converted into "The number of Jupiter's moons is four," [where] "is" has the sense of "is equal to," "is the same as"' (Frege 1997, 106–7). Yes, we do use the attributive form (that's the empiricism talking), but it is possible, and Frege suggests more perspicuous, to convey the information with an identity statement. This occurs shortly before paragraph 64 on content carving and foreshadows his principal application of the notion.

The *motivation* for carving is normative, or ameliorative: it would be *better* if the job had been given to *S** instead of *S*. This is not to say that the notion of carving is itself normative. Rather, the normative claim has a factual presupposition, and carving speaks to that presupposition. *S** can't do the job better than *S* unless there is a thing they both do, each in its own way. *Content* is Frege's word for the thing they both do; *carving* is his word for their different ways of doing it. All of that is perfectly factual. I'll be suggesting, however, that the

factual notion comes into clearer focus if we remember the ameliorative motivation.

4. Roads to Content

What is this thing *content*, the expressing of which is the job *S* and *S** have in common? One thing we know is that it lies somewhere between truth-value and sense. I want to sketch a route up to content from truth-value, and a route down from sense.

Up from Truth-Value

A semantic theory is *materially adequate* if it assigns semantic values and semantic structures that together yield sentences' actual truth-values. It may seem that the fundamental semantic project for a given language is to come up with a materially adequate interpretation of the language. But on reflection, that is not quite right.

A theory could be materially adequate for the wrong reasons. It could be a matter of *luck* that it succeeds in making all the truths come out true and the falsehoods false; the theory is able to cope only because the facts take a particularly tractable form. An accidentally adequate theory would not have been acceptable to Frege. It is hard to imagine him saying, 'Let's hope it snows tonight, for if not, then truth-values will be distributed in a way I am powerless to predict, given the structures I have assigned.'

Imagine that we mistakenly treated 'something' as a name. This would make it very hard to assign semantic values to basic expressions so as to generate the correct truth-values. For let *F* be a predicate that is true of some things but not of everything. 'Something is *F*' and 'Something is not-*F*' ought to both come out true. Whatever value *s* we assign to the 'name,' though, this is not the result we obtain; we can make $\sim Fs$ true only by making *Fs* false.

Well, but we might get lucky. It might happen that every predicate of the language was satisfied either by everything or by nothing. Then we wouldn't *want* *Fs* and its negation to both come out true. It wouldn't matter what we assigned to *s*; if *Fs* was true (false) on one assignment, it would be true (false) on all of them. Our theory would escape refutation by pure dumb luck.

A good theory should not depend for its material adequacy on lucky accidents. That is essentially to say that one needs a *policy* of semantic value assignment that, no matter how the world turns out, assigns values of a type that, plugged into the relevant structures, takes one to the correct destination, true or false. The policy should specify for each expression E , and every situation w in which we might find ourselves, that E 's semantic value on the w -hypothesis is so and so.

What determines if the policy is a good one? Well, sentences have to come out with the right truth *profiles*. That is, we first ask what truth-value a sentence S *deserves* on the hypothesis that we are in situation w . Then we ask what truth-value it *receives* if basic expressions are assigned values according to the policy. A successful assignment policy has these always coming out the same.

Good, but where do contents come in? They are already in. An *assignment policy* is no different from a series of functions – one per expression E – taking circumstances w to E 's semantic value SV in w . A *truth profile* for sentence S is no different from a function taking circumstances w to S 's truth-value in w . Value profiles are my candidates for the role of E 's content and truth profiles are my candidate for the content of sentence S .

This identification having been made, the project of assigning semantic values to basic expressions that yield the expected truth-values *non-accidentally* is the same as the project of assigning *contents* to basic expressions that yield the expected *contents* for sentences. To whatever extent the first project is Frege's, the second is too, albeit formulated in a way he might not recognize or appreciate.

Down from Sense

The basic work of a sentence is to be true or false. Of course, the sentence is true or false because of its sense, or the thought expressed. But there is liable to be more to the thought than is needed to determine its truth-value (examples in a moment). One might want to abstract away from this excess and limit attention to those aspects of the thought that are potentially relevant to truth.

This is not so different from what Frege himself does when he abstracts away from tone and colour and from the 'hints' given by

words like 'still' and 'but.' According to Frege, 'Alfred has still not arrived' *hints* that Alfred is expected, but this has no bearing on truth-value – Alfred's turning out not to be expected would not make the sentence false – so Frege leaves it out of the thought. The idea is to continue Frege's project of purging whatever is truth-irrelevant and focusing on what is left. *One* natural stopping point is the thought, but it is possible to go further.

Take 'Today is sunny' (uttered on August 16, 2002) and 'August 16, 2002, is sunny' (this is Michael Beaney's example). What do they share? Not sense, because accepting one does not rationally commit me to accepting the other. More than truth-value, though, because truth-value ignores that the thoughts stand or fall together. Beaney remarks that

An obvious candidate is Frege's early notion of conceptual content, which, if a metaphysical gloss could be put on the notion, might be best characterized as referring to [Umstände or] 'circumstances' (Frege 1997, 34–35; see pp. 52–53, *Begriffsschrift 2*, for Umstände).

The sameness of content here seems well captured by our idea of senses whose differences are guaranteed in advance not to make a difference to truth-value. The thought expressed by 'Today is sunny,' uttered on August 16, cannot differ in truth-value from the one expressed by 'August 16 is sunny.' Likewise the thoughts expressed by 'Gustav Lauben is thinking,' spoken by Gustav Lauben, and 'I am thinking,' written by Lauben at the same time.

Of course these indexicality-based examples can hardly serve as a model of the relation between 'lines *a* and *b* are parallel' and 'the direction of *a* is identical to the direction of *b*.' An example from Davidson (1979) comes a bit closer. Suppose that everything has exactly one shadow, and each shadow is had by one thing. Associated with each name *a* there is a name *a*[#] that stands for the *a*-object's shadow. To each predicate *P* corresponds a predicate *P*[#] that is true of a shadow just if the object casting the shadow is *P*. If *S* is an atomic sentence *Pa*, let *S*[#] be *P*[#]*a*[#]. Clearly *S* and *S*[#] differ in sense; only one involves the concept of shadow. But the difference is of no possible relevance to truth-value. However matters might stand, *S* is true if and only if *S*[#] is. They therefore agree in content, if content is the truth-relevant aspect

of sense. There are other ways of pulling the same basic trick, such as Quinean proxy functions.

A third example comes from the 'slingshot' argument – an argument which has been seen as clarifying and/or consolidating Frege's reasons for rejecting a level of significance between sense and truth-value, such as content is supposed to be. How does the argument go? Let S and T be both true or both false. Then it seems that the following ought to be a significance-preserving sequence:

1. S
2. $0 =$ the number which is 0 if S and 1 if not- S .
3. $0 =$ the number which is 0 if T and 1 if not- T .
4. T

1. and 2. differ in sense, since only one involves the concept of a number. But this difference is (so one might claim) of no possible relevance to truth-value, hence the sentences have the same content. The same applies to 3. and 4. If 2. and 3. share a content, we are sunk, because all truths then wind up with the same content.

If content is the aspect of sense bearing on whether a sentence is true or false, the question of whether 2. and 3. share a content boils down to this: is the substitution of T for S potentially relevant to truth-value? Given our stipulation that S and T are both true or both false, it may seem the answer is *No*. But this is to confuse *epistemic* relevance – what might *for all we know* change the truth value – with the intended semantic notion – what is in a *position* to change the truth value, even if we happen to know that in this case it does not. The slingshot argument gives no reason for rejecting content as we are beginning to conceive it here.

5. What Content Is

Two routes to content have been described. The first was a route up from truth-value, where the upward pressure was exerted by the *non-accidental* character of a theory's success at predicting truth-values. The second was a route down to content from sense, where the downward pressure came from our desire to bleach out aspects of sense with no possible bearing on truth.

One hopes, of course, that the two routes will converge on the same point. And this appears to be the situation; for to say of the thoughts expressed by *S* and *T* that their differences are of no possible relevance to truth value is to say that they are both true, or both false, *no matter what*. This gives us a first rough definition of what is involved in sharing a content.

(CONTENT – intuitive) *S* and *T* share a content iff the thoughts they express, although perhaps different, differ in a way that makes no (possible) difference to truth.

One can imagine ways of elaborating this. One could ask, e.g., that it be *knowable a priori* that *S* and *T* have the same truth value no matter what. And one could ask that it be a priori knowable *independently of any intelligence one might possess about what the sentences' truth-values actually are*.²² But these more elaborate approaches, although they get us a same-content relation on sentences, do not get us all the way to contents considered as entities in their own right. (The most we could hope for is equivalence classes of sentences.) Because the present approach calls for *contents*, we cannot afford to be so fancy. In this paper, 'same truth-value no matter what' means: true in the same cases.

This might be thought un-Fregean for the following reason. Cases sound a lot like worlds, and we are told that 'Frege has no notion of metaphysically possible worlds distinct from this world,' and indeed rejects 'metaphysical modality' altogether (Levine 1996, 168).

But to say that he had no use for *metaphysical* modality is not to say that Frege rejects modality altogether. He accepts a priority, it would seem, and an epistemicized form of analyticity (uninformativeness). More to the present point, the notion of sense is implicitly modal. Frege explains sense in more than one way, but there is a clear modal element in sense qua *mode of determination* (Frege 1997, 22–23).

People sometimes object to the mode of determination account that sense cannot determine reference *all by itself*. If it did, then merely understanding a sentence would put you in a position to know its truth-value. This is to read 'determines the referent' as 'leaves no room for other factors, such as the way of the world.' A more plausible

2 This is to allow for differences in content between sentences both of which are knowable a priori, e.g., 'Sisters are siblings' and Fermat's Last Theorem.

reading is 'exhibits the referent as a function of those other factors.' I don't see what it can mean to say that the sense of 'the Evening Star' determines its reference if not that the reference is one thing if, say, the brightest object in the evening sky is Mars, another if it is Venus. Similarly, what could it mean to say that the thought expressed by a sentence determines its truth-value, if not that the truth-value is one thing if the world is this way, another thing if not?

That Frege accepts some sort of modality – call it 'conceptual' modality – might seem no help because possible worlds are suited to the explanation only of metaphysical modality. But this is in fact controversial. Some philosophers maintain that there are two quite different ways of associating sentences with worlds, one of which lines up with conceptual necessity more than metaphysical.

How is this supposed to go? When Kripke talks about the worlds in which *S*, he means the *w*'s that answer to the description that *S* gives of our world. I will call that *satisfaction*. A world *satisfies S* iff it would have been that *S*, had *w* obtained. When Fregeans talk, to the extent that they can be induced to talk, about worlds in which *S*, they mean the *w*'s such that if this turns out to be *w*, then *S*. I will call that *verification*. A world *verifies S* iff *S* holds on the supposition that *w* really does obtain.

Consider a world *w* where Venus appears in the evening but the planet appearing in the morning is Mars. This world doesn't *satisfy* 'Hesperus isn't Phosphorus,' because it is not true that if certain appearances had had different causes, Hesperus, that is Venus, would have been distinct from Phosphorus, that is, Venus. But the world I mentioned *does verify* 'Hesperus isn't Phosphorus,' for if astronomers have *in fact* misidentified the morning-visible planet – it's really Mars – then Hesperus *isn't* Phosphorus.

The mode of evaluation relevant to sense is verification; to say that thoughts determine truth-values is to say that whether they are true or false depends on what (actually) happens. But then, given that content is a coarsening of sense, the mode of evaluation bearing on content ought to be verification too.³ I find it is easier to keep the *verification* aspect clearly in mind if we speak not of worlds but

3 'But does the proposition "The Earth has two poles" mean the same as "The North Pole is different from the South"? Obviously not. The second proposition could be true without the first being so, and vice versa' (*Grundlagen*, §44).

cases. ('Have you heard? The morning-visible planet turns out not to be Mars.' 'In that case, Hesperus is not Phosphorus.') So the proposal is that

(CONTENT-official) *S* and *T* share a content iff they are true in the same cases.

Contents are sets of cases.

Once again, 'Hesperus = Phosphorus' has a contingent content, since it is not true in all cases. This is what you would hope and expect, if the modality involved was non-metaphysical; it is only in a metaphysical sense that Hesperus could not have failed to be Phosphorus. A sentence that is true in all cases is Evans's 'Julius invented the zip, if any one person did.' This is necessary not in a metaphysical sense – Julius's mother could have had the idea first – but conceptually – it could not turn out that the inventor wasn't Julius.

6. Conflation

So we don't need to worry that contents explained as sets of worlds are objectionably *modal*. Some related worries cannot be laid to rest so easily.

One is that there are not enough contents to go around, with the result that sentences that ought intuitively to be assigned *different* contents will be forced to share. I will call this the Conflation Problem.⁴ Consider 'Julius (if he existed) invented the zip,' 'Sisters are siblings,' and 'There is no largest prime number.' These are true, let's suppose, in all cases, hence in the same cases. Yet one doesn't feel that 'Julius invented the zip' recarves the content of the Prime Number Theorem.

Bob Hale suggests an interesting answer to this objection, though he doesn't accept the answer himself. The objection would succeed, he says,

4 One version of this has been answered, viz., that Kripkean *a posteriori* necessities will share a content – the necessary content – despite conveying very different empirical information. Kripkean *a posteriori* necessities aren't true in the same cases. The counterfactual hypotheses *H* such that *if H is really the case, then gold is a compound* are not the ones such that *if H is really the case, then cats are robots*.

if the claim were that two sentences' having the same content is not only necessary but also sufficient for one to be properly viewed as recarving the content of the other. But the defender of the Fregean account has no need to make so strong a claim: he can claim that coincidence in truth-conditions ... suffices as far as the requirement of identity of content goes, but point out that this does not preclude the imposition of further conditions on the sentences involved. (Hale 1997, 95)

Michael Potter and Timothy Smiley find this baffling: 'Hale is suggesting, twice over, that two sentences can have the very same content but not count as recarvings of that content. This seems to us incomprehensible' (Potter and Smiley 2001, 328; to follow this debate see Hale 2001 and Potter and Smiley 2002).

Once we draw a certain distinction, however, Hale's position is no longer at odds with that of Potter and Smiley. The distinction is between *S*'s *tolerating S**-style recarving, and its *inviting S**-style recarving. Potter and Smiley are right about the first; if two sentences share a content, then each tolerates the recarving of its content provided by the other. But Hale's remarks can be read as directed at the second, and then they seem entirely reasonable. *S invites* an *S**-style recarving of their shared content only if the new carving improves somehow on the original; and most ways of recarving a content are just different, not better. There will be more on this after we consider the Proliferation Problem.

7. Proliferation

There is a way of putting Proliferation that makes it sound just like Conflation. Conflation occurs if

too many thoughts are recarvings of one content.

Proliferation occurs if

one content admits recarving into too many thoughts.

The difference is a matter of emphasis. Conflation puts it on *one content*. Each of the carvings may be in its own way legitimate; but they

shouldn't all be of the same content. Proliferation puts the emphasis on *too many thoughts*. There is no problem about these thoughts' carving the same content, supposing them to be otherwise admissible; but lots of them *aren't* otherwise admissible. This worry arises in a particularly sharp form on the conception of content proposed above.

Suppose that contents are sets of cases, and that *S* and *T* share a content *C*. What is it for them to carve *C* differently? To have a specific example, *S* might be the conjunction of A_1 and B_1 , and *T* the disjunction of A_2 and B_2 . $A_1 \& B_1$ carves its content conjunctively by exhibiting that content as arrived at by taking the intersection of two other contents, those of A_1 and B_1 respectively. $A_2 \vee B_2$ carves that same content disjunctively, because it represents it as obtained by taking the union of A_2 's content with B_2 's. *S* and *T* carve the content differently because they exhibit it as constructed along different lines.

Carvings on this view are semantic etiologies, or constructional histories. I will usually confine myself to immediate history, though ideally one would want to reach further back. A complete constructional history would be a structure tree of the kind found in categorial grammar textbooks. I will be worrying only about the top of that tree.

Now clearly, there is no backward road from a set to its history. Sets can be constructed in millions of ways, limited only by the ingenuity of the constructor. It helps a little to restrict the modes of construction to intersection, complement, and other functions expressed by logical devices present in natural language. But it doesn't help very much. Just as every number is a sum, difference, product, and so on, many times over, every set of cases is a union, intersection, complement, and so on, many times over.

Someone might say, what's wrong with that? Let a hundred flowers bloom. Maybe the resistance is just aesthetic and can be overcome.

But it is not just aesthetic. The resistance has to do with the role content carving is supposed to play in the introduction (or revelation) of objects. Initially, one is suspicious of so-and-so's and reluctant to accept them as real. Then it is pointed out that they are discernible in recarvings of contents one already accepts. This is supposed to be reassuring. The objects were there lying in wait; they spring into view as soon as we set our logical microscope to the correct power.

This does sound reassuring. But not if it turns out that there are no controls on the operation – that objects of practically whatever type

you like can be discerned in contents of practically whatever type you like. And this is a very real danger, if all it takes to discern a type of object in a content is to work it into the calculation by which the content is obtained.

Do we really need Hume's Principle to exhibit arithmetic as already implicitly there in the contents of sentences we accept? If incorporating numbers into a constructional history is enough, then it can be done a lot more easily. Start with any sentence you like, say, 'I reckon upon a speedy dissolution.'⁵ One route to its content is to look for the cases where Hume reckoned on a speedy dissolution. Another is to look for the worlds where he reckoned on a speedy dissolution and Peano's Axioms hold. You get the same worlds either way, arguably. But we don't want to say that numbers can be discerned in the content of 'I reckoned upon a speedy dissolution.'

It might be held that numbers *can* harmlessly be worked into any old content once we've got them – once we've obtained a guarantee of their existence. But to obtain that guarantee, you need a recarving with the right sort of epistemological backing. This is what Hume's Principle was supposed to provide, and abstraction principles more generally. To the extent that these can be regarded as merely concept-introducing – as teaching us what a direction or number is supposed to *be* – they seem well positioned to give us the required guarantee.

But problems also arise with recarvings backed by abstraction principles. I am not thinking here of the Bad Company Objection, which points to superficially similar principles – Frege's Basic Law (V) or Boolos's Parity Principle – that, singly or together, generate contradictions. Suppose that the inconsistency-threatening principles can be cordoned off somehow. We are still left with what is after all a more common problem with Company: that of being Uninvited and Unwelcome.⁶ The Uninvited Company Objection, as I will call it, goes like this: Principles superficially similar to Hume's can be used to introduce perfectly consistent objects which have, however, nothing to recommend them. If the case for numbers is no better than for some of these other objects (see below), then it is hard to see why

5 This was Hume's comment on learning that his condition was 'mortal and incurable' ('My Own Life').

6 The paper's working title was 'Visiting relatives can be boring.'

anyone but the extreme ontological maximalist should take numbers seriously (Eklund 2006). Richard Heck puts the problem as follows:

Let xQy be an equivalence relation, chosen completely at random: It might, for example, have as one of its equivalence classes the set containing each of my shoes, my daughter Isobel, the blackboard in Emerson 104, and some other things. We can now introduce names purporting to stand for objects of a certain sort, call them *duds*, just as we introduced names of shapes and [directions]:

$dud(a) = dud(b)$ iff aQb .

But are we really to believe that there are such abstract objects as *duds*? To take a less random case, consider the equivalence relation: person x has the same parents as person y . In terms of this relation, we can introduce names of what I shall call *daps*. But are there such abstract objects as *daps*? In so far as one has an intuition that there are no such objects ... there ought to be a corresponding doubt whether the neo-Fregean explanation of names of [shapes and directions] explains *these* names in such a way that they must denote abstract objects.... Nothing in the neo-Fregean story distinguishes these cases. (Heck 2000, 145)

Hale in a very subtle discussion gives some less artificial examples:

can we really believe that our world contains, alongside our PM, that lady's *whereabouts*, and that in addition to Smith's murderer, there is another object, *his identity*, and, besides the claimant, his or her *marital status*? (Hale 1988, 22).

It would be silly to suppose that numbers had no better claim on our attention than this lot; one doesn't want to throw the baby out with the *bathwater*. But, and this is the worry, one doesn't want to take the bathwater in with the baby either.

8. Bad Carving and Material Falsity

I spoke earlier of Frege's *rationalism*. Tyler Burge shows (Burge 1984; 1992; 1998) that this rationalism runs very deep. Frege believes in a *natural order of thoughts* to which human cogitation is naturally drawn.

These thoughts are grasped obscurely to begin with, but more and more clearly as inquiry progresses. When Cauchy and Weierstrass gave their epsilon-delta definition of a limit, they did not replace one lot of calculus thoughts with another, so much as clarify the thoughts that people had already had.

Frege also holds that there are *objective laws of truth* charting the relations between thoughts. And he arguably also holds that the more a thought's entailment relations are subsumable under laws of truth, the better the thought is. The reason, or one reason, that epsilon-delta thoughts are so good is that they turn what would otherwise be analytic entailments (trading on special features of limits, or of infinitesimals) into maximally general logical entailments.

Now I want to sketch a different rationalist theme whose role in Frege's thinking has not been much discussed. It's a theme that goes back to Descartes. Descartes gave our ideas two kinds of representational task. First is the task of *standing for whatever it is that they stand for*. Second is the task of *giving a non-misleading impression of that something*. Success at the first task by no means ensures success at the second; indeed, failure at the second task presupposes success at the first. An idea has to reference something, before it can count as giving a wrong impression of that something.

Our idea of pain does a fine job, he thinks, where referring to pain is concerned. But it gives a confused or misleading impression of what pain is. Our ideas of heat and cold leave it unclear whether heat is the 'real and positive' partner, and cold merely the absence of heat, or the other way around. Redness looks to be an intrinsic property of the red object, but it is really something else, perhaps a disposition to cause reddish sensations. (Locke complains about secondary quality ideas quite generally that they fail to resemble the properties they are ideas of.) Ideas that misrepresent, in the sense of giving a false or misleading impression of, their objects are called (by Descartes) *materially false*.

Now, Frege is not very interested in ideas. His preferred representational vehicle is the sentence. But a similar distinction can be made with respect to them. A sentence might succeed in expressing a certain content, while giving a wrong impression of that content.

The analogy may seem strained, to begin with. What sort of impression do sentences give, after all, of their contents? Well, a sentence

containing a name of Socrates might give the impression of expressing a content in which Socrates figures. A sentence with a certain kind of logical structure might give the impression of expressing a content that is structured the same way.

Here the analogy breaks down, one might think. How pain and colour can match up with our ideas of them is clear enough. (The idea of redness presents it as intrinsic, and this impression is either right or wrong, depending on what redness is really like.) But it is not initially clear how contents – sets of cases – could have things (Socrates) figuring in them, or how they could possess a certain logical structure.

The proper name case is easier. ‘Superman isn’t real’ gives the impression of being true in cases where there is an individual, *Superman*, with the property of not being real. A thing’s figuring in a content is its existing in the cases that make up that content.

But how can a sentence be misleading about the logical structure of its content? Wouldn’t that require contents to be structured, and aren’t sets precisely unstructured? They are certainly not explicitly structured. It could be, however, that sets of cases ‘lend themselves’ to a certain style of decomposition, as the set of muskrats and bees lends itself to decomposition into the set of muskrats and the set of bees.

An analysis of this ‘lends itself’ talk is suggested by David Lewis. A set of particulars is disjunctive, he says, if it is the union of two sets each of which is much more natural than it is (Lewis and Langton 2001). The same should hold for sets of cases or (as I’ll now say) worlds. The worlds with lots of neutrinos or lots of dragons in them decompose naturally into the lots-of-neutrino worlds and the lots-of-dragons ones. This, Lewis says, is because (i) the first set is the union of the second and the third, and (ii) the first set is much less natural than the second and third are. A similar analysis suggests itself of negative contents and perhaps also conjunctive ones.

Even if contents are sets, then, it is not out of the question that sentence *S* should give a ‘materially false’ account of its content. *S* performs wonderfully at its primary task of *expressing* the relevant content. But it gives a misleading impression of that content, because the content is disjunctive and *S* is of the form *A* & *B*.

Suppose we revisit the Conflation and Proliferation problems with these notions in mind. Regarding Conflation, Hale suggests that sentences with the same content could nevertheless fail to count

as alternative carvings of that content. Potter and Smiley find this incomprehensible. I said that both sides can be right, once we distinguish *S tolerating* recarving by S^* , and its *inviting* that kind of recarving.

One key element in *S* inviting recarving by S^* is that S^* does better at what we have called its secondary representational task; it exhibits the relevant content as put together in a way that is truer to that content's nature.⁷

Our answer to the Conflation problem is this. When sentences share a content, each is indeed amenable to recarving in the style of the other. But these recarvings will normally be uninvited and unilluminating. There is nothing objectionable about a content's tolerating lots of recarvings, so long as it doesn't invite all of them.

Something similar applies to Proliferation. Unwanted objects may be *discernible* in lots of contents – but that will be because the carving was uninvited. Our policy should be to recognize only the entities that cry out to be recognized because their contents lend themselves to quantificational carving. Of course, I haven't yet said how the quantificational case is supposed to go, and what I do say might be found unconvincing. But this should not distract from the key distinction: objects revealed when a content *cries out* for quantificational recarving vs. objects 'revealed' when a content *tolerates* quantificational recarving.

9. Respecting a Content's Internal Nature

What I would give you now, if I had it, is a general analysis of what is involved in a content's being implicitly disjunctive, or negative, or quantificational, and so on, through all the logical forms. No such analysis is known to me. On one conception of logical form, I doubt it is even possible. This is the conception whereby a content is disjunctive, say, pure and simple – disjunctive *as opposed to* negative, or existential. There is no reason why some contents shouldn't lend themselves to more than one sort of decomposition.

⁷ A tertiary representational task will be brought in later. S^* outperforms *S* at the tertiary task if its way of carving *S*'s content does better justice to that content's external nature – its entailment relations with other contents.

Because the labels 'disjunctive,' 'negative,' and so on are apt to sound exclusive ('which is it?'), I will use a slightly altered terminology. Instead of calling a content disjunctive, I will say it has *disjunctivitis*, on the understanding that a content can in principle have two or more -itis'es at the same time. (The content of '*p* is an electron or positron' is disjunctive with respect to charge – positive or negative – but conjunctive with respect to charge and mass.)

A different reason for preferring the '-itis' labels is that they suggest a cluster of related conditions, not all of which need be present on every occasion. Disjunctivitis (say) might be defined by a largish list of such conditions. Today I will not be trying to finish these lists; it will be enough if we can get them started. ('...' in what follows is short for 'along with other conditions to be named later.')

C has *negativitis* iff
it is the complement of a more natural content ...

C has *disjunctivitis* iff
it is the union of a finite number of contents each more natural than it ...

C has *conjunctivitis* iff
it is the complement of a content with disjunctivitis ...

The hard part, of course, is the quantifiers. I will state the proposal, explain it, defend it, and finally apply it.

C has *existentialitis* iff
it is the union of a congruent, complete bunch of contents ...⁸

C has *universalitis* iff
it is the intersection of a congruent, complete bunch of contents ...⁹

Congruence and completeness are explained as follows. Let the C_x s be a bunch of contents.

8 Another possible requirement is that finite sub-unions have disjunctivitis.

9 Another possible requirement is finite sub-intersections have conjunctivitis.

The C_k s are *congruent* iff the intersection of some of them with one other is more natural than the intersection of some of them with the complement of that other.

The C_k s are *complete* iff the intersection of some but not all of them is less natural than the intersection of all of them.¹⁰

A sentence P mirrors C 's structure, or, better, mirrors a structure of C , if

C has negativitis with respect to some content and P is the negation of a sentence expressing that content, or

C has disjunctivitis with respect to some contents and P is a disjunction of sentences expressing those contents, or

C has conjunctivitis with respect to some contents and P is a conjunction of sentences expressing those contents, or

C has existentialitis with respect to some contents and P is an existential generalization whose instances express those contents, or

C has universalitis with respect to some contents and P is a universal generalization whose instances express those contents.

10 Universal quantifiers should exhaust their domains. How can exhaustiveness be captured without a 'that's all' clause? I have no good answer to this, but here are a couple of thoughts. Sometimes the objects in a domain are 'of their own nature' such as to exhaust that domain. The omega-rule testifies to this in the case of numbers; from $\varphi(1), \varphi(2), \dots \varphi(k) \dots$, it directly follows that $(n)\varphi(n)$. That 1, 2, 3, ... are all the numbers is internal to them, obviating the need for a separate clause. Another example might be the universal domain (the domain of all possible objects). Perhaps we should think of ourselves as defining universalitis just for this particular case – the case of objects that by nature exhaust their kind. A quite different approach would be to amend the definition to say that contents are complete if the intersection of some of them is less natural, not than the intersection of all of them, but than a content 'just stronger' than that intersection – in the sense, perhaps, that it properly entails the intersection and all and only its entailments.

Now the main definition.

S does better justice to C's structure than S does iff S* mirrors a structure of C that S fails to mirror.*¹¹

Of course, *S* could in principle return the favour; it could mirror a structure that *S** fails to mirror. *S** does *strictly* better justice than *S* to *C*'s structure if the relation holds in one direction only.

10. 'Enumerative' Induction¹²

Imagine that we are physical scientists who have never thought of quantifying over numbers. Electrons and protons we have lumped together as 'trons.' Whether an atom is electrically charged, we notice, is not always predictable from how many trons it has. Atoms with two trons, or four, or six, are sometimes charged and sometimes neutral. A four-tron atom is neutral if two of its trons are protons and two are electrons; otherwise it is charged. Atoms with one tron, however, or three, or five, are always charged. Further testing reveals that the same holds of atoms with seven trons, or nine. At this point, we stop to review our findings.

[One] Atoms with one tron are charged.

-
- 11 'Doing better justice to *C*'s structure' is doing better justice to *C*'s internal nature (its implicit logical form). Ideally one would like to assign some weight also to doing better justice to *C*'s external nature (its implication relations).
- 12 This section was inspired by an example in Field (1994 §11). Should I believe Charley, when he says there was a foot of snow on the ground in Mobile, Alabama, one day in 1936? I should, for he has made many surprising claims in the past, and they all turned out to be true. The inference here appears to be an induction on truth: Charley spoke the truth when he said parts of Virginia are north of parts of New Jersey, Charley spoke the truth when he said the Soviets secretly supplied arms to Chiang Kai-Shek, etc., so most likely Charley is speaking the truth about the snow in Mobile. How are deflationists supposed to understand this inference, rejecting as they do a projectible property of truth? Premises of the form 'Charley said "*Ak*" and *Ak*' would not seem to lend inductive support to a conclusion of the form 'If Charley said "*B*" then *B*.' I adapt Field's example to the case of numbers. (I should say that Field is less impressed by it than I am, or was when I wrote the paper.)

[Two] Atoms with three trons are charged.

...

[Nine] Atoms with seventeen trons are charged.

The data seem to suggest some larger hypothesis. How to express this hypothesis is not clear, until someone has the idea of introducing the device of infinitary conjunction.

The data suggest that

$[\infty]$ [One] & [Two] & [Three] & ...

Not only are [One]-[Nine] evidence for $[\infty]$, they confirm it in the way that a lawful generalization is confirmed by its instances, with examined cases supporting unexamined cases. They confirm it in the way a black raven confirms *All ravens are black*, as opposed to the way a fair coin's coming up heads confirms *This coin comes up heads every time*.

Statements do not normally provide this kind of support – *inductive* support – to their conjunction. Why now? It is true that this particular conjunction has an especially natural content. But it is unclear why this would make a confirmational difference.

A familiar if simple-minded picture of confirmation runs as follows. Confirmation is the converse of explanation. Smoke on the mountain indicates the presence of fire there because, and to the extent that, the fire hypothesis explains why there would have been smoke.

(CON) A given body of data D confirms hypothesis H iff

- (a) H explains D , or
- (b) H follows from an $H+$ that explains D .

Say our data D is various bits of copper all conducting electricity. D confirms *Copper (as a rule) conducts electricity* because, and to the extent that, *Copper conducts electricity* explains why the tested bits were found to conduct electricity. It confirms *Other, so far unexamined, bits of copper also conduct electricity* because this follows from the explanation we gave of the tested bits' conducting electricity.

To this theory of confirmation, let us now add a simple-minded (broadly Humean) theory of explanation:

(EXP) K explains D iff

- (a) K entails¹ D
- (b) K is a highly natural hypothesis¹³
- (c) K is no stronger than (a) and (b) require.¹⁴

Copper conducts electricity is a highly natural data-entailing hypothesis that is weaker than other such hypotheses; it is weaker, e.g., than *Everything conducts electricity*. According to (EXP), then, it is copper's conducting electricity that explains why the observed bits of copper were found to conduct electricity.

One further principle will be needed; it is suggested by (CON) and (EXP), and plausible on its face:

(NAT) if D confirms H , then $D \ \& \ H$ is more natural than $D \ \& \ \sim H$.

Why do I say this is suggested by (CON) and (EXP)? Suppose that D confirms H . By (CON), H either (a) explains D itself, or (b) follows from an $H+$ that explains D . Suppose first that H explains D itself. Then H is the weakest highly natural hypothesis that entails D (by (EXP)). But then $D \ \& \ \sim H$, which also entails D , cannot be highly natural; if it were, H would not be weakest among highly natural D -entailing hypotheses. $D \ \& \ H$ is highly natural, though, since $D \ \& \ H = H$ (H entails D by (EXP) (a)). So $D \ \& \ \sim H$ is not as natural as $D \ \& \ H$.

Imagine now that H does not itself explain D ; D is explained by an $H+$ that entails H . That $H+$ explains D means that $H+$ is the weakest highly natural hypothesis that entails D (again by (EXP)). There can be no equally natural hypothesis entailing $D \ \& \ \sim H$, or $H+$ would not be weakest among highly natural hypotheses entailing D . Hence $D \ \& \ H$ is implied by a more natural hypothesis (viz. $H+$, recall that $H+$ implies H) than any implying $D \ \& \ \sim H$. Hypotheses are more or less natural, other things being equal, according to the naturalness of what implies them; so $D \ \& \ H$, being implied by $H+$, is more natural than $D \ \& \ \sim H$. This completes the argument that D combines more naturally with an H that it confirms than with that H 's negation.

13 Meaning, at least, more natural than the data (not necessarily limited to D) that K is called on to explain.

14 Meaning, K is weakest among highly natural D -entailing hypotheses.

Now, as we saw, $[\infty]$ bears the same inductive relations to [One], [Two], ..., and [Nine] as a lawlike generalization bears to its observed instances. How, according to our simple-minded theory of confirmation, must [One]-[Nine] ($= D$) and $[\infty]$ ($= H$) be related for this to be so? A lawlike generalization's observed instances are supposed to confirm

- (i) the generalization as a whole,

and

- (ii) the generalization's unobserved instances.¹⁵

The question, then, is this: How must D and H be related for D to confirm

- (i') H as a whole, that is, the conjunction of [One], [Two], [Three], ..., [Ten], [Eleven], [Twelve], etc.

and also

- (ii') H 's unobserved instances, that is, [Ten], [Eleven], [Twelve], etc.?

By (NAT), D confirms H 's unobserved instances only if

- (ii'') D 's conjunction with [Ten] ([Eleven], [Twelve], ...) is more natural than its conjunction with not-[Ten] (not-[Eleven], not-[Twelve], etc.)

By (CON) and (EXP), D confirms H as a whole only if

- (i'') H is weakest among highly natural D -entailing hypotheses.

This is interesting, because (ii'') is a special case of what we above called Congruence.¹⁶ H 's conjuncts are congruent provided that (ii'') continues to hold no matter which $[k]$'s go in for [One]-[Nine] in D .

15 A fair coin's coming up heads confirms (makes it likelier) that the coin always comes up heads. But it doesn't inductively confirm that the coin always comes up heads, because it doesn't make it any likelier that it will come up heads on the next toss.

16 The more natural hypothesis, I'm assuming, is the one with the more natural content.

Similarly (i''), assuming it too continues to hold for other choices of D , implies what we above called Completeness. For suppose the conjunction of only *some* of [One], [Two], [Three], ... , [Nine], [Ten], [Eleven], ... was as natural as the conjunction of all of them. Then, letting that sub-conjunction be D , H would not be weakest among highly natural D -entailing hypotheses; D would itself be a weaker such hypothesis. H 's conjuncts are complete provided that (i'') continues to hold no matter which $\{k\}$'s go in for [One]-[Nine] in D .

Where does this leave us? H 's conjuncts inductively confirm H only if they are complete and congruent. For them to be complete and congruent is for H 's content to have universalitis with respect to its conjuncts' contents. So,

Statements like *Atoms with three trons are charged* inductively confirm *Atoms with one or three or five or ... trons are charged* only if the latter's content has universalitis with respect to the contents of statements like the former.¹⁷

Contents with universalitis are contents that invite recarving as universal generalizations. To think of $[\infty]$ as inductively supported by its conjuncts, then, is to conceive $[\infty]$'s content as deeply, underlyingly, *general*, whatever grammatical face it has been presenting to the world up to now.

By now the neo-platonic drift of all this should be clear. A generalization needs a domain to generalize over. That domain would seem, in this case, to be the natural numbers. The content-respecting way to express $[\infty]$ is:

$(\alpha)(n)$ (if the number of α 's trons is $2n - 1$, then α is charged).

This gives new meaning to the term 'enumerative induction.' Numbers worm their way into empirical contents, not just on representational grounds, but also evidential ones. Induction needs an axis to operate along. Numbers provide that axis.¹⁸

17 H 's content is, as ever, the conjunction of the contents of [One], [Two], It thus admits conjunctive recarving. The present issue is what sort of recarving H 's content invites.

18 'Does this mean that numbers are evidentially relevant to the course of physical events?' Imagine a missile evades our defences because it's surrounded by

11. Respecting a Content's External Nature

Contents are not only structured, they stand in implication relations; one content entails another just if it is a subset of the other. These two aspects of content seem connected, for we have generalizations like this: a conjunctive content necessitates its conjuncts. It seems a natural conjecture that *implicit structure and implication relations are two sides of the same coin*. That is, a content is conjunctive just if so construing it helps to explain why it has such and such implications and why thus and such contents imply it.

But now what does *that* mean? Frege thinks there are 'laws of truth.' They are maximally general both in being topic-neutral and in quantifying over absolutely everything. One sentence *logically entails* another if the corresponding conditional is a law of truth, or an instance of such a law. (I am treating the laws of truth as sentences rather than thoughts.)

Now suppose that content *C* implies content *D*; and suppose *C* and *D* are the contents of sentences *P* and *Q* such that it is a law of truth, or an instance of such a law, that if *P*, then *Q*. So, *P* might be 'There are at least two rocks in my shoe' and *Q* might be 'There is at least one rock in my shoe.' The fact that 'If *P*, then *Q*' is a maximally general truth explains why *C* implies *D* in the Humean sense of subsuming the entailment under a general law. *C* implies *D* because *C* and *D* are the contents of sentences *P* and *Q* where *P* logically entails *Q*. *Mutatis mutandis* for cases where it is *D* that implies *C*.

There is a second way of explaining implication relations. Suppose that we can find sentences with contents *C* and *D* whose corresponding conditional *becomes* subsumable under a law of truth once certain synonymous substitutions are made. The sentences might be 'Lucy is someone's sister' and 'Lucy is someone's sibling,' which become logically related when 'female sibling' is substituted for 'sister.' This again helps us to understand why *C* implies *D*. The explanation informally

a large number of decoys. Does this mean that numbers bear causally on the course of physical effects? No, the number serves here merely to mark the fact that there were a lot of decoys. Numbers' role in induction is admittedly deeper than this. But the general point is the same.

put is that a sentence with content *C* *analytically* (not logically) entails a sentence with content *D*.

Now, although there are these two explanations, the second is not quite as good because the corresponding conditional as written, that is, without the synonymous substitutions, is subsumable only under a *pretty* general law, viz. the law that whoever is someone's sister is their sibling. If we have up to now been explaining why *C* implies *D* by reference to analytically related sentences, it would be better – more explanatory – if we could substitute sentences that were *logically* related.

A given content stands in lots and lots of implication relations. The necessary content is implied by (the content of) 'Whoever has two apples has a prime number of apples' and also by (the content of) 'Julius invented the zip if any one person did.' Not all of these relations can be explained by means of a single sentential representation, for the sentence would have to follow analytically both from 'Whoever has two apples...' and 'Julius invented....'

If *C* is *S*'s content, explaining *C*'s implication relations is a *shared responsibility*, undertaken by a bunch of sentences of which *S* is one. *S*'s share of this responsibility is limited to the contents of sentences that it analytically entails, and conversely. Let's say that *S* analytically entails a content *D* if it analytically entails a sentence *T* with *D* as its content; and similarly for being analytically entailed by a content. Then the contents *S* takes responsibility for are the ones it analytically entails and those that analytically entail it. *S** shoulders this responsibility better than *S* to the extent that it turns analytic entailment relations into logical ones.

*S** does better justice than *S* to *C*'s implication relations iff

*S** logically entails some contents that *S* only analytically entails, or is logically entailed by some contents that only analytically entail *S*.¹⁹

Once again there is a stricter version, which adds that *S* does *not* logically entail any contents that *S** (only) analytically entails, and similarly for entailment-by.

19 I assume that *S** analytically entails whatever *S* does and is analytically entailed by whatever analytically entails *S*.

Now suppose we want to explain C 's implication of D using a sentence P that expresses C . Then we face two distinct challenges: first, the challenge of finding a sentence Q with content D ; second, the challenge of making it a Q that P entails logically and not just analytically. Corresponding to these two tasks, P can have two kinds of explanatory advantage over S . It can open the C - D implication up to explanation in the first place; this will occur when the S idiom is not up to the task of expressing D , while the P idiom is up to the task. Second, P can explain the C - D implication better by bringing it under a more general law. Let me first give an example of the second sort of advantage.

Suppose we speak an ordinary first-order language with variables ranging over concreta. Numerical quantifiers $(\exists_5x)Fx$ ('there are five F s') are defined in the usual way. A binary quantifier $(\exists_x)[Fx, Gx]$ ('there are exactly as many F s as G s') is introduced, its meaning specified by rules like the following:

$$(\exists_nx)Fx, (\exists_nx)Gx \\ \therefore (\exists_x)[Fx, Gx]$$

$$(\exists_nx)Fx, \sim(\exists_nx)Gx \\ \therefore \sim(\exists_x)[Fx, Gx]$$

$$(\exists_x)[Fx, Gx], (\exists_nx)Fx \\ \therefore (\exists_nx)Gx$$

$$(\exists_x)[Fx, Gx], \sim(\exists_nx)Fx \\ \therefore \sim(\exists_nx)Gx$$

$$\sim(\exists_x)[Fx, Gx], (\exists_nx)Fx \\ \therefore \sim(\exists_nx)Gx^{20}$$

20 Rules like this leave $(\exists_x)[Fx, Gx]$'s meaning unspecified when both predicates have infinitely large extensions. (Thanks to John Burgess and Neil Tennant for pointing it out.) One possible solution is to stipulate that $(\exists_x)[Fx, Gx]$ holds true whenever its negation is not provable from the other rules. All infinite totalities will then wind up with the same number.

Number terms are introduced by Hume's Principle: $\text{num}(F) = \text{num}(G) \leftrightarrow (\exists_{_x})[Fx, Gx]$. Numerals are introduced by Frege's Principle: $\text{num}(F) = n \leftrightarrow (\exists_{_n}x)Fx$.

Given all this, a lot of contents can be formulated either with numerical determiners $(\exists_{_n}x)Fx$ or numerical terms $(n = \text{num}(F))$. When C can be formulated either way, which formulation does better justice to the content's implication relations?

It depends. Adjectives do better with some such relations. As is familiar, $(\exists_{_5}x)Fx$ logically entails $\sim(\exists_{_7}x)Fx$. But $\text{num}(F) = 5$ only *analytically* entails $\text{num}(F) \neq 7$; the inference runs essentially through Frege's Principle. Terms do better with other implication relations. The symmetry of identity is a logical truth, so $\text{num}(F) = \text{num}(G)$ logically entails $\text{num}(G) = \text{num}(F)$. But $(\exists_{_x})[Fx, Gx]$ entails $(\exists_{_x})[Gx, Fx]$ only analytically; the inference runs essentially through Hume's Principle. Each formulation logicalizes an aspect of the implication profile left as analytic by the other. You need both for the full picture.

But there is another respect in which the numerical-term formulation seems superior. The style of expression used in P can't explain C 's implication relations with D unless it enables the formulation of a Q that expresses D . The problem with the determiner formulation is that there are plenty of relevant D s that are *prima facie* beyond its expressive powers. In the tron example, it was important that the content D of 'Atoms with one, three, five, seven, or etc. trons are charged' implied the content C of, e.g., 'Atoms with five trons are charged.' The first content is *prima facie* inexpressible in a first-order language ranging over concreta, since infinite conjunctions are not allowed. It can be stated, however, in a language with numerical terms: for all atoms a , if $\text{num}(a$'s trons) is odd, then a is charged. The numerical-term formulation allows us to subsume D 's entailment of C under a logical law – the law of universal instantiation.

12. Morals If Any

The hypothesis of this paper is that S invites recarving by S^* iff S^* does better justice than S to C 's internal and external nature. The question should finally be raised of how far this hypothesis supports the program of Fregean platonism. I confess I don't know. One complication is that the Fregean platonist may well look askance at our ontology of

cases and contents – not because abstract ontology is objectionable per se, but because it is supposed to be a consequence of the project rather than a presupposition of it.

But let's suppose that that can somehow be dealt with. Does it help Fregean platonism if the content of 'There are exactly as many Fs as Gs' invites recarving as 'The number of Fs = the number of Gs'?

On the one hand, I want to say *Yes*. If a carving-based reason is to be given why we should (or can at no epistemological risk, or already implicitly do) countenance numbers,²¹

(◇) known truths – contents we know to be true – can be rearticulated in a numerical fashion

seems a good deal less convincing than

(□) known truths – contents we know to be true – demand to be rearticulated in numerical fashion.

As we saw, (◇) makes an exceedingly weak case for numbers, no better than could be made for duds, or whereabouts, or party affiliations. (□) makes a stronger case. Known truths do not cry out to be rearticulated in terms of duds.

The case (◇) makes for numbers is also un-Fregean. Numbers are supposed to be discernible in *cardinality* contents specifically – the kind found on the right-hand side of Hume's Principle. But all kinds of contents can be rearticulated so as to involve numbers. Cardinality contents have no particular advantage.

(□) avoids this problem. Numbers are not invited into the content of any old truth (e.g., that Hume reckons on a speedy dissolution), but only truths whose implicit logical structure is thereby illuminated.

(□) makes the better case. But the case is not irresistible. That known truths cry out for numerical rearticulation could be heard less as a theoretical argument for the objects' existence than a practical argument for postulating them quite regardless of whether they exist.

21 I say 'if a carving-based reason is to be given' because Crispin Wright offers a different sort of reason; the key point for him is that Hume's Principle canonically explains the concept of number.

(If numbers did not exist, it would be necessary to invent them.) It can even be heard as a reason to suspect that they *already have been* postulated regardless of whether they are there. This way lies fictionalism, or figuralism, or presuppositionalism, presumably of the hermeneutic variety.

I take no stand here on which of these views – Fregean platonism or hermeneutic fictionalism/figuralism/presuppositionalism – is in the end preferable. My point is directed at both equally. Both camps attempt to portray number-talk as harmless and unobjectionable. Our entitlement to arithmetic is secure because it flies beneath skeptical radar; there is nothing in it to arouse intellectual opposition.

I am suggesting that both camps would do better to take the offensive, arguing that the facts as we know them cry out for numerical treatment and we ought to heed their call. It can be left as a further question whether the heeding should take the form of believing in numbers or acting as if we believed in them.²²

References

- Burge, T. 1984. 'Frege on extensions of concepts, from 1884 to 1903.' *Philosophical Review* 93: 3–34.
- Burge, T. 1992. 'Frege on knowing the third realm.' *Mind* 101: 633–50.
- Burge, T. 1998. 'Frege on knowing the foundation.' *Mind* 107: 305–47.
- Davidson, D. 1979. 'The inscrutability of reference.' *Southwestern Journal of Philosophy* 10: 7–19.

22 Hans Herzberger offered a regular seminar on 'The Inexpressible' at the University of Toronto in the 1970s. I took it in 1979, my senior year. What an experience. I seem to remember every session: St. John of the Cross; Teresa of Avila; the tetralemma of the Buddha; the silence of the Buddha; Stace on mysticism and logic; Frege on the concept horse; Russell on self-reference and impredicativity; Wittgenstein on showing and saying; Ryle on heterologicality; Mirimanoff on grounding; Martin and Searle on the universality of natural language; Priest's logic of paradox; Carnap, Tarski, and Kripke on semantic closure. It was thrilling and unforgettable and a big part of why I became a philosopher. (Another big part was visiting Hans and Radhi in Rishi Valley the next year. He was working out his revision semantics for truth and allowed me a ringside seat.) The problems Hans showed me then are still the ones that obsess me today. Here in support of that is a paper on truth, logic, and expressive power.

- Eklund, M. 2006. 'Neo-Fregean ontology.' *Philosophical Perspectives* 20(1): 95–121.
- Field, H. 1994. 'Deflationist views of meaning and content.' *Mind* 103: 249–85.
- Frege, G. 1997. *The Frege Reader*. Oxford and Malden, MA: Blackwell.
- Hale, B. 1988. *Abstract Objects*. New York: Blackwell.
- Hale, B. 1997. 'Grundlagen §64.' *Proceedings of the Aristotelian Society* 97: 243–61.
- Hale, B. 2001. 'A response to Potter and Smiley: Abstraction by recarving.' *Proceedings of the Aristotelian Society* 101: 339–58.
- Heck Jr, R. G. 2000. 'Syntactic reductionism.' *Philosophia Mathematica* 8(2): 124–49.
- Levine, J. 1996. 'Logic and truth in Frege: II.' *Proceedings of the Aristotelian Society*, Supp. vol. 70: 141–75.
- Lewis, D., and R. Langton. 2001. 'Redefining "intrinsic".' *Philosophy and Phenomenological Research* 63(2): 381–98.
- Potter, M., and T. Smiley. 2001. 'Abstraction by recarving.' *Proceedings of the Aristotelian Society* 101: 327–38.
- Potter, M., and T. Smiley. 2002. 'Recarving content: Hale's final proposal.' *Proceedings of the Aristotelian Society* 102: 351–54.