Content Carving: Some Ways

Stephen Yablo, MIT

1. The problem

Here is Frege in Foundations of Arithmetic, paragraph 64:

> The judgment 'Line a is parallel to line b', in symbols: a || 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 || 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 usa new concept" ({Frege, 1997 #2059}, 110-1).

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

(1) 'a || b' and 'dir(a) = dir(b)' have the same content,

and

(2) reflecting on that can lead one to the concept of direction.

Why doesn't (2) contradict (1)? (2) has us grasping the content of 'a || b' before we grasp that of 'dir(a) = dir(b).' How in that case can the contents be identical, as (1) assures us they are? Leibniz's Law would seem to forbid it. If we grasp content **X** at 11.14am, and content **X** = content **Y**, then by gum we grasp content **Y** at 11.14am.

I know of only one good way of getting around this. The neophyte does grasp the content of 'dir(a) = dir(b)'; she just fails to know it as the content of an identity-sentence. She won'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.

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 sentence.

Remember, the neophyte has to grasp the content of 'dir(a) = dir(b)' before possessing the concept of direction. So if content is sense, she must be able to grasp the sense of 'dir(a) = dir(b)' before possessing 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 'dir(a) = dir(b)'? That is, each of the three indicated achievements presupposes the one "before" it:

Grasp content of 'dir(a) = dir(b)'

⇓ ⇑

Acquire concept of direction ⇒ Grasp sense of 'dir(…)'

Since Frege's strategy does not appear to work, if content is sense, it must be something other than 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 right-to-left arrow is compulsory too,

for grasping the sense of ‘direction-of’ is appreciating that it expresses that 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 way to arrange that is to make content coarser-grained than sense, though presumably still finer-grained than reference. Then everything hangs together just right:

(1) 'dir(a) = dir(b)' shares something with 'a || 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.

2. 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 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 pile 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 have got 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.

So, 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 run roughshod over grammatical appearances.

Now 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 structure, 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 rather than one structured like so 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 in situations where, as Quine later puts it, greater theoretical profundity is professed.

Consider in this connection 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 #2059}, 106-7). Yes, we do use the attributive form (that's the empiricism talking), but it is possible, and he suggests more perspicuous, to convey the information with an identity statement. This occurs shortly before para. 64 on content carving and foreshadows his principal application of the notion

The motivation for carving is normative, or ameliorative: it would be better in some important respect if the job had been given to S* instead of S. This not to say that that the notion of carving is itself normative. Rather the normative claim has a factual presupposition, and carving pertains to that presupposition. S* can't do the job better 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. It turns out, though, that the factual notion becomes less mysterious if we remember the ameliorative motivation.

3. Roads to Content

So, what exactly 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 and a second route down from sense.

Up from Truth-Value

Call a theory materially adequate if it assigns semantic values and semantic structures that together yield the correct truth values. At first it seems our project 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 in principle be a matter of luck that it succeeds in making all the truths come out true and the falsehoods false; if different sentences had been true the thing would not have been possible. It is hard to imagine Frege saying, "I sure hope it snows tonight, for if not, then truth values will be distributed in a way I am powerless to understand, given the structures I have assigned."

Suppose we mistakenly treated "someone" as a name. Then there would be no way to make "someone is F" and "someone is not-F" both come out true; whatever value you assigned to the "name," if Fs is true then not—Fs is 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 need Fs and its negation to both come out true; if someone is F, everyone is F. By pure luck our theory would escape refutation.

A good theory should not depend for its material adequacy on lucky accidents. That is essentially to say that we need a policy of semantic value assignment that, no matter how the world turns out to be, gives you values that plugged into the structures you have assigned takes you to the correct destination, true or false. The policy should say 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, first we ask what truth value a sentence S deserves on the hypothesis that we are in

situation W. Then we ask what truth value the sentence receives if basic expressions are assigned values according to the policy. A successful assignment policy has these always coming out the same.

So far so good, but where do contents come in? They are already in. An assignment policy is no different from a bunch of functions (one per expression E) taking circumstances W to semantic values SV. A truth profile for sentence S is no different from a function taking circumstances W to S's truth-value in W. These functions are my candidates for the content of expression E and 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 his 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 that aspect of the thought potentially relevant to truth.

This is not so different from what Frege himself does when he abstracts away from tone and color 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 since that makes no difference to truth — Alfred's turning out not to be expected would not make the sentence false — he leaves that aspect of meaning out of the thought. The idea here is to continue Frege's project of purging whatever is truth-irrelevant and focussing on what is left. One natural stopping point is the thought, but it is possible to go further.

Take 'Today is sunny' (uttered today) and 'August 16 2002 is sunny' (this is an example of Michael Beaney's). What do they share? Not sense or thought, 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 says 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 #2059}, 34-5; see 52-3, Begriffschrift 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 Aug 16 cannot differ in truth-value from the the one expressed by 'Aug 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 this indexicality-based example can hardly serve as a model of the relation between "these lines are parallel" and "their directions are identical." An example of Davidson's comes a bit closer ({Davidson, 1979 #2182}). Suppose that everything has exactly one shadow, and vice versa. 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 iff 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. Regardless of how matters in fact stand, S is true iff $S^{\#}$ 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 because 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 then

we are sunk, because all truths will wind up sharing a content.

The question is, is the substitution of S for T 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 potential relevance of an epistemic sort -- what might for all we know change the truth value -– with the intended semantic notion -- what is in a position to change truth value, even if we happen to know that it does not. The slingshot argument gives no reason for rejecting content as we are (slowly) beginning to conceive it here.

4. What Content Is

Two routes to content have been described. First was a route up from truth value, where the upward pressure is exerted by the hopefully non-accidental character of our success at predicting truth values. Second was a route down to content from sense, where the downward pressure comes from the 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 they have the same truth value no matter what. And one could ask that it be apriori knowable independently of any intelligence one might possess about what the sentences' truth values actually are.[1] But the more elaborate approach, although it gets us a same-content relation on sentences, does 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 my approach calls for contents, I cannot afford to be so fancy. In this paper, "same

[1] This is to allow for differences in content between sentences both of which are knowable a priori, eg., the Prime Number Theorem and Fermat's Last Theorem.

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 #2074}, 168).

But to say that Frege had no use for metaphysical modality is not to say that he rejects modality altogether. Frege accepts apriority and an epistemicized sort of analyticity. And some sort of modality seems to be implicit in the notion of sense. Frege explains sense in more than one way, but one explanation is mode of determination ({Frege, 1997 #2059}, 22-23).

People sometimes object to the mode of determination account that sense cannot determine reference all by itself, or 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 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 body visible in the evening 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 set of 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.

So, to go with the usual example, 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.

Now clearly the mode of evaluation relevant to sense is verification; to say that thought determines truth value is to say that whether actually true or actually false depends on what (actually) happens. But then given that content is a coarsening of sense, the mode of evaluation bearing on content is verification too.[2] I find it is easier to keep the verification aspect clearly in mind if we speak not of worlds but cases. ("Have you heard? The morning-visible planet turns out

[2] "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).

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 is non-metaphysical, since 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 — it could have been Julius's mother that invented the zip — but conceptually — it could not turn out that the inventor wasn't Julius.

5. Conflation

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

One is that there are not enough contents to go around, so that sentences that ought intuitively to be assigned different contents will be forced to share a content. I will call this the Conflation Problem. One common version of this has been answered, viz, that Kripkean aposteriori necessities will share a content — the necessary content — despite conveying very different empirical information. Our answer to this is that Kripkean aposteriori necessities aren't true in all cases, or in the same cases. There are plenty of things we could learn that would lead us to say, "if that is really the case, then this lectern is made of ice."

Be all that as it may, the same content is assigned to sentences true in all cases: sentences as intuitively inequivalent as “Julius (if he existed) invented the zip,” “Sisters are siblings,” and “If f is a continuous function on the reals, then every y between f(a) and f(b) is f(x) for some x between a and b.” One doesn't feel that "Julius invented the zip" recarves the content of the Intermediate Value Theorem.

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

> 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 involed ({Hale, 1997 #2220}, 95).

Potter and 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, 2001 #2235} 328).

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 <u>tolerating</u> the recarving S* provides, and its <u>inviting</u> the recarving S* provides. Smiley and Potter 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 read as directed at the second. And then they seem entirely sensible. S <u>invites</u> 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.

6. 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 in where we put the emphasis. Conflation puts it on <u>one content</u>. Each of the carvings may be in its own way legitimate; but they shouldn't all be <u>of the same content</u>. Proliferation puts the emphasis on <u>too many thoughts</u>. There is no problem about these thoughts' carving the same content, supposing them to be otherwise admissible; but lots of them <u>aren't</u> otherwise admissible. This worry arises in

a particularly sharp form on the conception of content proposed earlier.

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 and B, and T the disjunction of C and D. A & B carves its content conjunctively by exhibiting that content as arrived at by taking the intersection of two other contents, those of A and B respectively. C ∨ D carves that same content disjunctively because it represents it as obtained by taking the union of C's content with D'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 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 certain objects and reluctant to accept them as real. Then it is pointed out that they are quantified over in recarvings of contents one already accepts. This is supposed to be reassuring. The objects were in some sense already there lying in wait; they spring into view as soon as we set our logical microscope to the correct power.

This does I admit 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 that

type of object into the calculation by which the content is obtained.

So, to be a bit silly about it, why do we need Hume's Principle to exhibit arithmetic as already implicitly there in 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, "George W. Bush is from Texas." One route to its content is to look for the cases where Bush is from Texas. Another is to look for the worlds where Bush is from Texas and Peano's Axioms hold. You get the same worlds either way. Surely though we don't want to say that numbers can be discerned in the content of "Bush is from Texas."

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 threaten contradiction. Suppose that inconsistency-threatening principles can be cordoned off somehow. We are still left with what is after all a more common problem with Company: that they are Boring and Useless.[3] The Boring and Useless Company Objection notes that principles superficially similar to Hume's can be used to introduce perfectly consistent objects which have, however, nothing to recommend them.

Examples of bogus-seeming abstracta are not hard to imagine, but here are a few to get the ball rolling. Are there really such things as rents {Heck Jr, 2000 #807}, defined by the principle that

x's rent = y's rent iff x has the same parents as y?

What about gregrets, introduced as follows:

[3] I was thinking of calling this paper "Visiting relatives can be boring."

> x's gregret = y's gregret iff both most regret getting a full body tattoo or both most regret missing the 1999 Harmonic Convergence or …?

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 #1149}, 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.

7. Bad Carving and Material Falsity

I spoke earlier of Frege's rationalism. Tyler Burge has argued ({Burge, 1998 #2266}, {Burge, 1984 #2278}, {Burge, 1992 #2271}) that this rationalism runs very deep. Frege holds that there is a natural order of

thoughts to which human thinking 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 call attention to a different rationalist theme whose role in Frege's thinking has not been much discussed.

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 entails success at the

second; indeed failure at the second task presupposes success at the first. An idea has to reference some phenomenon before it can count as giving a wrong impression of that phenomenon.

Descartes thinks our idea of pain does just fine as far as referring to pain is concerned, but gives a confused or misleading impression of what pain is. He and others say similar things about our ideas of color; 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 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 (not by Locke, but some others) materially false.

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

How would that work? 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 we seem to run into a disanalogy, for pain and color can really be, or fail to be, as represented by our ideas. As already mentioned, the idea of redness might give the impression that it is an intrinsic property, and this impression is either right or wrong depending on what redness is really like. But it's not initially clear how contents — sets of cases — could have things (Socrates) figuring in them or possess this or that logical structure.

That having been said, it takes only a little imagination to see how the extension might go. "Superman isn't real" is misleading because it gives the impression of being true in worlds where there is an individual Superman with the property of not being real. X figuring in a content would be X occurring in the worlds that make up that content.

But what about the structure of the sentence giving a potentially false impression of the structure of the 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 "lend themselves" to a certain style of decomposition. An example would be the set of muskrats and bees lending itself to decomposition into the set of muskrats and the set of bees.

An analysis of this "lend itself" talk is suggested by David Lewis. A set is disjunctive, he says, if it is the union of two sets each of which is much more natural than their union ({Lewis, 2001 #1663}). The same could apply to sets of cases or (as I'll now say) worlds. Consider the set of worlds in which there are trillions of neutrinos and Bush is Texan. Surely this lends itself to decomposition into the trillion-neutrino worlds and the Bush-is-Texan ones. Lewis would explain this by saying that it is the union of two other sets of worlds both of which are much more natural than it. A similar analysis suggests itself of negative contents and perhaps also conjunctive ones.

Even if contents are sets, then, it is not out o the question that sentence S could give a "materially false" account of its content. The sentence performs wonderfully at its primary task of expressing the relevant content. But it gives a misleading impression of the content it expresses, because that 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 might be right, once we distinguish S tolerating recarving by S*, and its inviting that kind of recarving.

One element in S inviting recarving by S* is that S* does better at its secondary representational task; it exhibits its content as carved in a way that is truer to that content's internal nature. This is the topic of the next section. Another element, to be discussed in the section after that, is that S* does better at (what we might call) its tertiary representational task by carving the contnt in a way that does better justice to its external nature, by which I mean its entailment relations with other contents.

Our answer to Conflation was that if two sentences share a content, then each permits recarving in the style of the other. But this recarving will normally be uninvited and unilluminating.

Something similar applies to Proliferation. Boring and useless 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 claim that we (properly) recognize objects when a content we accept cries out for quantificational recarving.

8. Doing better justice to the 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, however. On one conception of logical form, I doubt it is even possible. This is the conception whereby a content is disjunctive pure and simple, as opposed to negative, or existential. There is no reason why some contents shouldn't lend themselves more than one sort of decomposition.

Because the labels "disjunctive", "negative," and so on sound in this way exclusive ("which is it?"), I will use a slightly altered terminology. Instead of calling a content disjunctive I will say it has disjunctivitis. It is understood that a content can have two or more —itises at the same time.

A different reason for preferring the "—itis" labels is that they suggest not a single defining property but 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 these conditions. Today I will not be trying to finish these lists; it will be enough if we can get them started. With these qualifications in mind, I propose that

> C has disjunctivitis iff
> it's the union of contents much more natural than it, along with other conditions to be named later.
>
> C has conjunctivitis iff
> it's the intersection of much more natural contents, plus other conditions to be named later.
>
> C has negativitis iff
> it's the complement of a much more natural content, plus etc.

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

> C has existentialitis iff it is the union of a congruent, complete bunch of contents.
>
> C has universalitis iff it is the intersection of a congruent, complete bunch of contents.

A bunch of contents are congruent if whenever the Xs are some of them and C is another of them, intersecting the Xs with C's complement makes for much less natural results than intersecting them with C. A bunch of contents are complete if whenever the Xs are all but one of them, the intersection of the Xs with each other is much less natural than their intersection with the one remaining content.

There are several ways for a sentence P to mirror the, or better mirror a, structure of content C, corresponding to the structure-types just distinguished. Either

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

- C has conjunctivitis with respect to two contents and P is a conjunction of sentences expressing those two contents, or
- C has negativitis with respect to a content and P is the negation of a sentence expressing that content, or
- C has existentialitis with respect to a bunch of contents and P is an existential generalization whose instances express those contents, or
- C has universalitis with respect to a bunch of contents and P is a universal generalization whose instances express those contents.

Now finally we can say what it is for S* to do better justice than S to C's structure.

> S* does better justice than S to C's structure iff S* mirrors some structure of C that S doesn't mirror.

"Doing-better-justice-than" sounds like it ought to be antisymmetric and transitive, but as defined here it is neither. It is not antisymmetric because each of two sentences might mirror a structure that the other misses. Transitivity would mean that each of the two sentences mirrored a structure that it didn't mirror, which is absurd. S* does strictly better justice than S to C's structure if the better-than relation obtains asymmetrically. Doing-strictly-better-justice than is

antisymmetric and transitive, although transitivity doesn't amount to much since if S* does strictly better than S, S** does at most weakly better than S*. (It is only when we drop the restriction to surface structure that transitivity begins to bite.)

Suppose that we are atomic scientists who have never thought of quantifying over numbers. Electrons, protons, and neutrons are in our terminology "trons." We notice that whether an atom is charged can often be predicted from how many trons it has.

[1] An atom with one tron is charged.
[3] An atom with three trons is charged.
…
[79] An atom with seventy-nine trons is charged.
…

And so on. Whenever we determine that all atoms with so many trons are charged, we find that all atoms with two more trons than that are charged too. We sense a pattern here and seek a hypothesis that would synthesize and perhaps even explain our data. But so far we lack even the means to express such a hypothesis. Someone has the bright idea of introducing infinitary conjunction; now we say

[∞] $\prod_{k=\text{one, two, three,}\ldots}$An atom with $2k+1$ trons is charged.

Our earlier statements seem to confirm this conjunction in the way a generalization is confirmed by its instances, with examined cases supporting unexamined cases. Conjuncts do not normally provide that sort of support to conjunctions in which they figure. Why now? It is true that this particular conjunction has an unusually natural content. But it is unclear why this would make a confirmational difference. Two assumptions will be needed:

> (I) Data D confirm a hypothesis H that is <u>independent</u> of D when the result of conjoining D with H is more natural (content-wise) than the result of conjoining D with not-H.
>
> (E) Data D confirm a hypothesis H that <u>entails</u> D when H is more natural (content-wise) than D and the same cannot be said of weaker D-entailing hypotheses.

Give these assumptions, what does it take for [∞] to bear the evidential relations to D = (say) [1], [3],..., [83] that a generalization bears to its instances; that is, for D to confirm <u>all</u> the conjuncts

> (I′) $\prod_{\underline{k}=\text{one, two, three,}\ldots}$ D confirms [<u>k</u>]

as well as their conjunction

(E′) D confirms $\prod_{\underline{k}=\text{one, two, three,}} [\underline{k}]$.

It follows from (I) that (I′) holds if

(I″) D's conjunction not-[$\underline{k}$] is less natural than its conjunction with [$\underline{k}$]

It follows from (E) that (E′) holds if

(E″) [∞] is more natural than D and the same cannot be said of any weaker D-entailing hypothesis

Note that (I″) is what above we called completeness, and (E″) is coherence. This suggests that D-like data support [∞] in the usual lawlike way just when the content of [∞] has universalitis with respect to the contents of [1],[3],[5], etc. [1], [3], [5], etc. bearing a familiar sort of evidential relation to their conjunction means that their conjunction invites recarving as a universal generalization, presumably "For all $\underline{n}$, an atom with $2\underline{n}$+1 trons is charged."

9. Doing better justice to the content's *external* nature

There is a problem. S* doing better (or strictly better) justice than S to their shared content is supposed to have the result that S invites recarving by S*. This makes the invites-recarving-by relation too indiscriminate. Suppose for argument's sake that the necessary content N is well carved by the Fundamental Theorem of Algebra, and badly carved by "Julius invented the zip if any one person did"; N has structure mirrored by the first sentence but none mirrored by the second. Do we really want to say that the Fundamental Theorem of Algebra invites recarving by the Julius sentence?

I see only one way to fix this. Contents are not only structured, they stand in implication relations; one content entails another iff 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 iff 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 rocks in my shoe" and Q might be "there is a 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 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 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 if it analytically entails some sentence with that content, and similarly for being analytically entailed by a content. Then S takes responsibility for the contents it analytically entails and those that analytically entail it. Once that is seen it is pretty clear what S* must do to fulfill that responsibility better.

> S* does better justice than S to **C**'s implication relations iff
>
> (1) S* analytically entails every content analytically entailed by S, and is analytically entailed by every content that analytically entails S;
>
> (2) S* logically entails some contents that S only analytically entails, and/or is logically entailed by some contents that only analytically entail S.

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.

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 Q 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 1st order language with variables ranging over concreta. Numerical quantifiers $\exists_5 \underline{x}\ F\underline{x}$ (there are five Fs) are defined the usual way. A binary quantifier $\exists_= \underline{x}\ [F\underline{x}, G\underline{x}]$ ("there are exactly as many Fs as Gs") is introduced, its meaning specified by rules like the following:

$\exists_n \underline{x}\ F\underline{x}, \exists_n \underline{x}\ G\underline{x}$
$\therefore \exists_= \underline{x}\ [F\underline{x}, G\underline{x}]$

$\exists_n \underline{x}\ F\underline{x}, \sim\exists_n \underline{x}\ G\underline{x}$
$\therefore \sim\exists_= \underline{x}\ [F\underline{x}, G\underline{x}]$

$\exists_= \underline{x}\ [F\underline{x}, G\underline{x}], \exists_n \underline{x}\ F\underline{x}$
$\therefore \exists_n \underline{x}\ G\underline{x}$

$\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$], ~$\exists_{n}\underline{x}$ F$\underline{x}$
∴ ~$\exists_{n}\underline{x}$ G$\underline{x}$

~$\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$], $\exists_{n}\underline{x}$ F$\underline{x}$
∴ ~$\exists_{n}\underline{x}$ G$\underline{x}$.[4]

Number terms are introduced by a version of Hume's Principle: num(F)=num(G) ≡ $\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$]. Numerals are introduced by what I will call Frege's Principle: num(F)=$\underline{n}$ ≡ $\exists_{n}\underline{x}$ F$\underline{x}$.

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

---

[4] Rules like this leave $\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$]'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 $\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$] holds true whenever its negation is not provable from the other rules. All infinite totalities will then wind up with the same number.

It depends. Adjectives do better with some such relations. As is familiar, $\exists_{5}\underline{x}\ F\underline{x}$ logically entails $\neg\exists_{7}\underline{x}\ F\underline{x}$. But num(F)=5 only <u>analytically</u> entails num(F)≠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 num(F) = num(G) logically entails num(G) = num(F). But $\exists_{=}\underline{x}$ [F$\underline{x}$, G$\underline{x}$] entails $\exists_{=}\underline{x}$ [G$\underline{x}$, F$\underline{x}$] 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 full picture.

But there is another respect in which the numerical-term formulation seems superior. Remember that 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 <u>a</u>, if num(<u>a</u>'s

trons) is odd, then a is charged. The numerical-term formulation thus opens **C**'s implication by **D** up to explanation, and indeed subsumes it under a logical law, the law of universal instantiation.

X. 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 structure and/or its implication relations. 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 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. Rather than saying that we are countenancing numbers because contents we accept can be rearticulated in a numerical fashion, now

we can say that we're countenancing them because contents we accept cry out to be rearticulated in numerical fashion.

But a completely different reaction is possible. That contents we accept 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. Taken a little further, that contents we accept cry out for numerical rearticulation could be heard as a reason to suspect that they already have been postulated them regardless of whether they are there. This way lies fictionalism, or figuralism, or presuppositionalism, presumably of the hermeneutic variety.

I have no idea which of these -- fregean platonism or hermeneutic fictionalism/figuralism/presuppositionalism -- is in the end preferable. My point is directed at both equally. Both camps have a tendency to remain always on the defensive; they try to weaken our resistance to numbers by making them, or our dealings with them, appear harmless and unobjectionable. 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.

Bibliography

Burge, T. (1998). "Frege on Knowing the Foundation." Mind.

Burge, T. (1992). "Frege on Knowing the Third Realm." Mind.

Burge, T. (1984). "Frege on Extensions of Concepts, from 1884 to 1903." Philosophical Review.

Davidson, D. (1979). "The Inscrutability of Reference." Southwestern Journal of Philosophy.

Frege, G. and M. Beaney (1997). The Frege reader. Oxford ; Malden, MA, Blackwell Publishers.

Hale, B. (2001). "A Response to Potter and Smiley: Abstraction by Recarving." Proceedings of the Aristotelian Society **101**: 339-358.

Hale, B. and C. Wright (2001). The Reason's Proper Study: Essays towards a Neo-Fregean Philosophy of Mathematics. Oxford, Clarendon Oxford Pr.

Hale, B. (1988). Abstract Objects. NY, BLACKWELL.

Heck Jr, R. G. (2000). "Syntactic Reductionism." Philosophia Mathematica **8**(2): 124-149.

Levine, J. (1996). "Logic and Truth in Frege: II." Aristotelian Society **Supp**(70): 141-175.

Lewis, D. and R. Langton (2001). "Redefining 'Intrinsic'." Philosophy and Phenomenological Research **63**(2): 381-398.

Potter, M. and T. Smiley (2001). "Abstraction by Recarving." Proceedings of the Aristotelian Society 101: 327-338.

Potter, M. and T. Smiley (2002). "Recarving Content: Hale's Final Proposal." Proceedings of the Aristotelian Society 102: 351-354.

Rosen, G. and J. P. Burgess (2005). Nominalism Reconsidered. The Oxford Handbook of Philosophy of Mathematics and Logic, Shapiro, Stewart. Oxford, Oxford Univ Pr: 515-535.

Wright, C. (1983). Scots Philosophical Monographs: Frege's Conception of Numbers as Objects", Andrew Brennan and Others (Eds). ABERDEEN, ABERDEEN UNIV PR.

Yablo, S. “Non-Catastrophic Presupposition Failure,” Stalnaker festschrifft

Yablo, S. (2005). “The Myth of the Seven” in Kalderon, ed., Fictionalist Approaches to Metaphysics

Yablo, S. (2002). "Abstract Objects: A Case Study." Nous Supplement 12: 220-240.

Yablo, S. (2001). "Go Figure: A Path through Fictionalism." Midwest Studies in Philosophy 25: 72-102.