DIGITAL ACCESS TO
SCHOLARSHIP AT HARVARD

Reply to MacFarlane, Scharp, Shapiro, and Wright

(Article begins on next page)

Reply to MacFarlane, Scharp, Shapiro, and Wright

Many thanks to MacFarlane, Scharp, Shapiro, and Wright for their trenchant and very useful comments. It's an honor to be so vigorously and well criticized. What follows takes up four topics: the idea that the central notion in semantics and logic is that of an apt or appropriate commitment; a defense of what *When Truth Gives Out* (*WT*, henceforth) says about vagueness, along with remarks on higher order vagueness; a discussion of how invoking speech acts beyond assertion adds to our understanding of the semantic paradoxes; disagreement and relativism.

1. Commitments and Semantics. Chapters 2 and 3 of WT make three points. (1) One kind of expressivist semantics has a straightforward response to the Frege-Geach problem; (2) this semantics helps explain how we can correctly say that a sentence or claim is neither true nor false, and thus affords some insight into vagueness and semantic paradox; (3) it also provides a response to worries that expressivist and emotivist accounts of evaluative language can't give a coherent account of negation, disjunction, and the like.

The semantics is based on three assumptions. (1) Simple sentences are conventional means for expressing mental attitudes (like belief, valuing, and being (morally) tolerant) and performing speech acts (like assertion and denial). (2) These mental states and speech acts involve *commitments*. Belief and assertion commit one to the truth of what is believed; valuing X commits one to X's being valuable. These commitments, in turn, have *aptness* conditions: commitment to p's truth is apt provided p is true; valuing x is apt provided x is valuable; toleration is apt provided what is tolerated is tolerable. This means that associated with the simplest sentences of our language are a mental attitude, an associated commitment, and the aptness conditions of the commitment. (3) For semantic purposes, we may blur the distinction between a commitment and its aptness conditions. An expressivist semantics ought compositionally assign commitments *represented in terms of their aptness conditions* to sentences; in particular, it should treat sentence compounding devices as expressions whose meanings map commitments-represented-via-their-aptness-conditions to

commitments. In doing this, the semantics makes clear how sentences are related to speech acts and mental states: a sentence is a conventional means for expressing a mental state S or performing act A provided that the characteristic commitment of S or A is the semantic value of the sentence.[1]

MacFarlane wonders if I have adequately explained what commitments are. In the case of belief he says that in order to be entitled to speak of commitments I need to first explain what it would be to violate or fulfill a commitment to (say) the truth of the claim that snow is white.

I respond as follows. The expressivist I am defending holds that the conventional role of logically simple sentences is to express mental states. It's clear that there are necessary and sufficient conditions for its being apt for an individual to be in the relevant states. Sometimes –in the case of belief –it is possible to understand these conditions in terms of truth: a belief is apt iff its object is true. The expressivist about valuings, as I understand him, holds that valuing something is not well understood as believing something about it; valuing is having a certain complex of emotional and volitional relations that (typically) are backed up by reasons. But the expressivist can surely agree that we can ask whether someone's valuing something is apt and –prescinding from vagueness and such –expect a more or less determinate answer.

The commitments WT invokes involve, and for semantic purposes are represented as, commitments to aptness conditions being in place. As such, the natures of these commitments are no more but no less clear than the conditions under which it is appropriate to believe p, to value o, or to feel one's self required to perform an act. Certainly the nature of the aptness conditions of these commitments is clear enough for us to *represent them for semantic purposes.* Truth conditions can be represented using sets of worlds; o's being valuable for x can *be represented in a model* simply by designating a set 'Valuable things for x' and putting o in it. It's clear enough what the *shape* of a commitment's aptness conditions are for us to avail ourselves of them when we're doing semantics.

MacFarlane claims that picking out a commitment via its aptness conditions is "a very indirect way of picking out the commitment", and that different sorts of

---

[1] I expand on what WT says about expressivism and semantics in Richard, ms. A.

commitments can have the same aptness conditions. He goes on to suggest that this is a problem. He notes, for example, that my semantics assigns identical aptness conditions to logically equivalent sentences. Thus –given that the semantics is in the business of assigning commitments to sentences –the semantics identifies distinct commitments – e.g., it identifies a commitment to the truth of the claim expressed by

(P1) $\neg$ (P v Q)

with a commitment to the truth of the claim expressed by

(P2) $\neg$P & $\neg$Q.[2]

I plead guilty to identifying commitments with their aptness conditions, for the purposes of giving a manageable semantics. Doing so is of a piece with identifying propositions in modal logic with their possible worlds truth conditions. If for semantic purposes we don't need to distinguish amongst commitments with the same aptness conditions, what's the problem?

Note, however, that aptness conditions are constructed in the semantics out of propositions.[3] The semantic value of the sentence (P1) is constructed from the proposition (P1) expresses ; the semantic value of logically equivalent (P2) is constructed from the proposition *it* expresses. Given that the sentences express different propositions, they receive different semantic values that determine the same aptness conditions. But the difference in the semantic values is such that when we get around to doing the

---

[2] In this paper I use propositional logic connectives to represent uses of connectives with a 'Strong Kleene' interpretation. I combine them with English expressions to indicate a particular understanding of English. For example, I write 'Jo is $\neg$ bald' when I want to discuss the use of 'Jo is not bald' when 'not' is used to express truth functional (as I see it, strong Kleene) negation. In displayed sentences, I use 'not', 'or', 'iff', and so on for occurrences of connectives that are to be understood as signaling force.

[3] This is true only of the aptness conditions of the sentences of the languages discussed in Chapter 2 and Appendix I of WT, which contain only proposition expressing sentences and the results of compounding them with force connectives. But the point I am making about aptness conditions generalizes to languages in which there are atomics that, for example, expressing valuing or approval.

semantics for attitude verbs, it will not imply that whoever believes that ¬P & ¬Q thereby believes that ¬(PvQ).

I take it that MacFarlane wants an account of what it would be to violate or fulfill a commitment to something's truth because he thinks of having a commitment as a species of the broader notion of having an obligation, and takes the notion of obligation as a notion that is tied to action in a broad sense. Here MacFarlane and I seem to be on different pages. I don't find it odd at all to speak of someone who believes or asserts p as being committed to p's truth. I don't think we can give a terribly informative account of how someone with that commitment should or should not behave. To give such an account would pretty much be giving an account of how someone who believes p should behave. But there is no such account. At least there is no such account save an infinitely disjunctive one that prescribes how one should behave when one believes p *given* information about one's other beliefs and desires.

To have a commitment to p's truth is not exactly to believe p, for one can have the commitment in virtue of putting oneself forward, by asserting p, as having the belief. Or one can have a number of beliefs that taken together induce commitment to p's truth, though one does not (having not performed the deduction) believe p. I take the notion of a commitment to the truth of a claim as primitive. But it doesn't strike me as a particularly obscure primitive.

I would say similar things about other commitments. For example, there is a distinctive commitment one incurs by denying a claim. If you understand what it is to deny a claim (as opposed to asserting its negation), you have a handle on this commitment; if you don't, you don't. If you don't understand what it is to deny a claim, I can try to explain it to you by trying to get you do understand that a claim may be true, false, *or* neither true nor false, and that denying a claim is the thing to do when the claim is not true. But if you didn't understand denial to begin with, this probably won't help.

Normative expressivists usually say that the moral approval expressed by 'charity is good' is a fundamentally different sort of state than belief. One might then wonder what this expressivist has in mind when she speaks of the commitment expressed by 'charity is good' or speaks of that commitment's being apt. The expressivist's response, as

I understand her, is that to say that charity is good is to express a commitment…to charity's being good; that commitment is apt provided that….charity is good.

Of course in identifying the commitment as a commitment to charity's being good, the expressivist is not saying that it is a commitment to the *truth* of the claim that charity is good; in giving aptness conditions she is not saying that the commitment is apt iff it is *true* that charity is good. If pressed as to what the commitment *is*, the expressivist, as I understand her, will explain by saying that we understand very well what it is to *have* the commitment, for we understand what it is to approve of something – it's what we are doing when we say that it is good. Likewise, we understand what it is for the commitment to be appropriate, for we understand what someone is doing, when they approve of charity; but when someone approves of charity, by saying 'charity is good', they are *inter alia* conveying that the commitment is appropriate.

Another way to put the point here is this: in giving the semantics for her use of normative language, the expressivist will, like everyone else, say things like

'charity is good' is an appropriate thing to say iff charity is good.

But the expressivist will *also* point out that the right hand side of this sentence is not making a truth evaluable claim, but is functioning in the way it normally functions. So long as the expressivist has a response to Frege and Geach's query, about how to understand the application of a connective to a sentence that is not a device of assertion, this seems a perfectly coherent view. It is part of the burden of Chapters 2 and 3 of WT to give such a response.

I think the view I here ascribe to the normative expressivist is coherent. I frankly don't know if I endorse it or not. What I was trying to show in the third chapter of WT was not that expressivism about the normative is correct, but that it is perfectly coherent, and that there is a coherent semantical story for the expressivist to tell. This story is a natural extension of what I take to be the obviously correct story to tell about our talk about truth valueless claims, like those made by paradoxical sentences or applications of vague predicates to their borderline cases.

2. Vagueness. Wright presents what he calls the Fregean conception of vagueness, on which to be a vague predicate is simply to have a partial definition. He conjectures that such a view is 'congenial' to me, and makes two objections: (a) the phenomenological objection that the Fregean view makes "bad predictions about our characteristic responses to…and responses to responses to" borderline cases; (b) the objection that it cannot account for higher order vagueness (HOV). In the latter case, he also specifically targets my own view as being committed to an incoherent account of HOV. Scharp and Shapiro (S&S, henceforth) worry that though I "resolve" one version of the sorities, there is a version of the sorities involving denial I cannot account for.

Let's begin by considering Wright's "Fregean account" of vagueness. On this view, 'pearl' is vague if defined as follows:

If x has chemical composition C and arises in an oyster, x is a pearl;

If x is a pearl, x has chemical composition C.

I don't think that 'pearl' so defined is vague, and I don't think that vagueness is simply a result of partial definition. I think that vagueness *involves* truth value gaps –if a predicate is vague, there are[4] objects of which it is neither true nor false. But I don't think that any predicate that suffers from truth value gaps is vague. 'pearl' above is an example. There's nothing vague about this predicate at all: assuming that 'arises in an oyster' and 'has chemical composition C' are precise, the notion of a pearl so defined is perfectly precise, and so not vague.

If vagueness is not simply a matter of a predicate having a partial definition, what is it? As I see it, the extensions of our predicates, in so far as they are determined, are determined by various facts about use. Such facts about usage give –or would give, to someone who knew all them –reasons for or reasons against applying a predicate to an object. Often they give both. Sometimes there is a way of balancing such reasons when they conflict so that there is a single verdict to the question 'does this predicate apply to

---

[4] or could be –but let's ignore this sort of modal qualification henceforth

that object?' But not always. Vagueness, as I see it, arises when there are conflicting, unbalanceable reasons for applying a predicate to an object.[5]

In addition to this, vague predicates are typically *left open* by use, in the sense that it is common knowledge that such predicates have a range of cases "open to accommodation". Object x is such a case when a predicate's use does not determine, once and for all, whether it applies to x, and conversational factors (of the sort Lewis points to in "Scorekeeping in a Language Game") may make the predicate true or false of x "for the purposes of conversation." I take this to be a natural concomitant of the sort of incommensurability that generates vagueness, though my instincts are that it is a different phenomenon.

All this said, there is something about Wright's "Fregean picture" that I accept: when a predicate is vague, it sorts its domain into three sets –those of which it is true, those of which it is false, and the rest. Call this, for future reference, *Trisection*. Does this make me liable to either Wright's phenomenological objection, or to an incoherent view of HOV?

Wright observes that (1) in the case of a predicate defined as 'pearl' above, we would expect speakers to have definite, sharp responses as to whether the predicate applies, but; (2) for paradigm vague predicates like 'bald' and 'red' we find 'weak…conflicting…unstable opinions.' But why should the mere fact that a predicate trisects its domain –or the fact that competent speakers are generally cognizant of that fact –lead us to predict that they won't have weak, conflicting, unstable opinions? Such speakers can be expected to be more or less aware of the currents and eddies of use that bear on the predicate's application. They will thus see reason, in borderline cases, to apply the predicate, reason to apply its negation, and fail to see a conclusive reason to go either way. Some will incline to use the predicate one way, some another, some will shrug; most will not have strong opinions, and opinions will tend to be unstable. I don't see *any* reason to think that Trisection makes the phenomenology of our use of vague predicates at all unlikely.

---

[5] John Broome 1988 sees incommensurability as a kind of vagueness; I see vagueness as a kind of incommensurability. I don't see that these views are incompatible.

If vague predicates trisect their domains, we can expect that in a sorites series for 'bald' there will be an object –n, call it –such that n is bald, while the next object in the series, n+1, is neither bald nor ¬ bald. S&S observe that this is a consequence of the semantics I give in WT; they think this to be an objection to the whole enterprise: "…the sorites issue comes up…in the language with force connectives, and we are given no solution to it."

I think S&S misdescribe their objection. They have in mind a version of the sorites in which object 1 is definitely bald, object 30,000 is definitely not, and the problem is to explain what is wrong with the argument

S.   Object 1 is bald.
        For any n: if object n is bald, then object n+1 is bald.
        So, object 30,000 is bald,

where the quantifier and conditional in the major premises are given force interpretations. To *that* problem, I give the solution: deny the major premises. It is the consequence of that denial, that

C.   There is an object n such that n is bald but it's not the case that n+1 is bald,

with quantifier and connectives given force readings, that S&S don't like.

Unfortunately, S&S don't say *why* they think that C is problematic. Surely it isn't *simply* because C implies that there is some sort of "boundary" in the relevant sorites series. Everyone should be willing to grant that there is *some* sort of boundary in such a series. For everyone should grant that each object in the series falls under just one of the following: Bald; ¬Bald; something else.[6] Somewhere in the series there is a transition from baldness; somewhere in the series there is a transition to ¬baldness.

---

[6] The dialethiest may say that some objects fall under more than one of these categories. But she had better acknowledge that *some* of the objects in the series fall under only the first category, some fall under only the second. And then the same sort of consideration will show that she acknowledges the existence of a boundary.

Perhaps it will be said that C implies that there is a "sharp boundary" in the sorites series, but there can't be such a boundary. But why should we think that C represents a "sharp boundary"? Object n is (definitely) bald, while there's just no saying whether object n+1 is bald. It is *indeterminate* whether n+1 is bald, since (this being a sorites series) object n+1 is sort of bald, sort of not bald; there's no saying either way, as there's no fact of the matter. There are two different ways that something of the form

Object n+1 is not bald

can be apt, when 'not' expresses denial: there can be a "hard fact" about n+1's baldness (n+1 is definitely not bald), or there can simply be an absence of fact about the matter. In the latter case, though there is a sort of a "boundary" between n and n+1 –there is a transition from fact to absence of fact --it is hardly a sharp one.

S&S remark, quite rightly, that C raises issues similar to the issue of higher order vagueness. WT doesn't address the latter issue, though I have discussed it elsewhere.[7] I will summarize that discussion, and then respond to Wright's claim that if I try to address HOV using the categories *true*, $\neg$*true*, and *not true and not* $\neg$*true*, a "general form of aporia…ensnares" me.

To say there is HOV is to say, for example, that an object might be on the borderline between being an F and being a borderline F. This is presumably equivalent to saying that it might be indeterminate whether an object is F or indeterminately F. Thus, one way to approach the issue of higher order vagueness is by asking what it means to say that it is indeterminate that such and so.

Talk of indeterminancy, it seems to me, is essentially contrastive: when we say 'it is indeterminate whether S', there is a T such that we are saying that it is indeterminate whether S or T. Abbreviate *it's indeterminate whether S or T* with ∇(S, T). Typically, though not invariably, when we say it's indeterminate whether S, for relatively simple S (e.g., where S =' J is bald') what's meant is that it's indeterminate whether S or $\neg$S. 'It's

---

Those who say that vagueness involves a sort of "boundarylessness" sometimes seem to deny what is said in the text. But it is never clear what view of sorites series they offer as alternative to the sort of trisection just mentioned in the text.
[7] Richard 2010. The next two pages repeat material from the fourth section of this paper.

indeterminate whether it's indeterminate whether S' would naturally, I think, be understood as saying $\nabla(\,S, \nabla(S, \neg S))$.

When are such claims apt? Well, if it's indeterminate whether S or T, then it's not settled whether S or whether T. And if it's not settled whether S or T, there must be some way of "tightening up" S and T's meanings, without affecting the non-semantic facts, on which it comes out apt to say S, unapt to say T.[8] Likewise, there must be some way of tightening up these meanings so that it comes out being apt to say that T, unapt to say S. For present purposes, we can think of such tightening up as a matter of extending a given assignment of extensions and anti-extensions to predicates, by putting some things for which a predicate is undefined in the predicate's extension or anti-extension.

Allow me the notion of an *adequate extension* of the semantics of a language. Such extensions will not subtract objects from the extensions or anti-extensions of predicates, but may add objects to these. Using that notion, we may say that

$\nabla(S,T)$ is appropriate in a language L if and only if there are adequate extensions f and g of the semantics of L such that:

in f, S is apt, T is not;

in g, T is apt, but S is not.

Here are some consequences of thinking of indeterminacy in this way.[9]

(1) $\nabla(S, \neg S)$ implies that neither S nor $\neg S$ has a truth value. (2) Suppose that S is the application of a garden variety vague predicate to a borderline instance, and so truth valueless. Then $\nabla(S, \neg S)$ is apt. (3) Suppose $\nabla(S, \neg S)$ is apt. Then so are

(i) $\nabla(S, \nabla(S, \neg S))$

_____

[8] To say this is to suggest that all vagueness is a result of "semantic indecision." The approach I'm sketching here does not, in fact require that we think this. (At least it doesn't on some ways of explaining the notion of an adequate extension of a model introduced in the next paragraph.) It would take us far too far afield to discuss this here.
[9] The proofs are elementary and in Richard 2010. (2) and (5) require us to assume that when x is, in interpretation I, a borderline case of the predicate F, there are adequate extensions I' and I'' of I such that x is the extension of F in I' and x is in the anti-extension of F in I''.

(ii) $\nabla(S, \nabla(S, \nabla(S, \neg S)))$

and so on up. (i) gives the form of the most likely regimentation of 'it's indeterminate whether Smith is bald or borderline bald'; (ii) gives the form of the most likely regimentation of 'it's indeterminate whether Smith is bald or on the borderline between being bald and being on the borderline between being bald and being borderline bald'. Thus, if there's first order vagueness, there is higher order vagueness. (4) A unary indeterminacy operator can be defined as follows:

> ∇S is apt in a language L if and only if there are adequate extensions f and g of the semantics of L such that:
> > in f, S is apt;
> > in g S is not apt.

If *S* is a sentence that is truth apt –it is, say, the application of an "ordinary" vague predicate such as 'bald' to an object –the sentence *not (S and ∇S)* is valid. (For if S is apt, it is true, and thus there will not be an adequate extension of the language's semantics in which S is untrue and ∇S is apt.)

> (5) When S is truth apt, the sentence

> (iii)   If ∇S, then $\nabla\nabla$S

is also valid. (For if ∇S is apt in I, then (i) S is apt and thus true in some adequate extension of I, and so ∇(S) is not apt in some adequate extension of I, and; (ii) ∇(S) is apt in the null extension of I.)

Let us now turn to Wright's complaints about how my view might deal with HOV. Wright first argues that

> …the Fregean conception [of vagueness, on which it is partial definition] can offer no coherent description of what a higher-order borderline case *is*. The borderline cases of *pearl* and *not a pearl* are cases where there is a mandate to

bring them under neither concept. Borderline cases of borderline cases…must thus be cases where there is neither mandate to apply the concept, nor to apply its opposite, *nor mandate to regard them as borderline of that distinction*. They are cases, then, which, for example, there is no mandate to affirm as red, no mandate to affirm as not red, but also no mandate to characterize as cases which there is neither mandate to affirm as red nor mandate to affirm as not red! That's incoherent.

As I understand him, when Wright says that borderline pearls for the Fregean are objects "where there is a mandate to bring them under neither" *pearl* nor ¬*pearl*, he means that the Fregean is committed to not applying *pearl* and not applying ¬*pearl* to them. I would regiment this commitment, where p is a borderline pearl, with

F1.   not (p is a pearl) and not (¬p is a pearl).

If p is a borderline perarl, I accept F1. What does Wright mean, when he says that for the Fregean a borderline case of a borderline case is one where there is not a "mandate to regard [it] as" a borderline case? I conjecture what he means is that the Fregean would have to say that if p is a borderline borderline pearl, then we cannot say that p is a borderline pearl. But if we cannot say that p is a borderline pearl, we have to deny F1. But –and here is the incoherence –a borderline borderline case of a pearl is surely not a pearl and not a ¬pearl. So it looks to be the case that the Fregean has to both put F1 forward for borderline borderline pearls and deny it, and that is indeed incoherent.

I am not altogether sure why the Fregean must deny that a borderline borderline pearl is a borderline pearl, and I am going to leave it to the Fregean to defend himself. I am pretty sure that *I* don't have to say this. To say that p is a borderline borderline pearl is, I take it, to say that it is indeterminate whether it is indeterminate that p is a pearl. That is, using the above definitions, it is to say ∇∇(p is a pearl). This is perfectly consistent with the claim that ∇(p is a pearl); both the later and the former will be apt when F1 is apt.

I expect Wright to object to the account of HOV above. One objection he will doubtless make is that to say that ∇(p is a pearl) and $\nabla\nabla$ (p is a pearl) is to say something incoherent --how can it be right to say anything of the form *p and it's indeterminate that p*? Now, it *is* wrong to say anything of this form if p is something that might be true or false –if p, for example, is something like 'Jo is bald'. If it is true (or false) that p, it can't also be indeterminate that p. And, as noted above, on the current account of indeterminacy, *not (p and ∇p)* is valid when p is a sentence like 'Jo is bald'.

But suppose p is a claim of indeterminacy –let say it's the claim that it's indeterminate that Jo is bald. In this case, p itself may be the apt thing to say, but it won't be *determinate* that this is so. After all, if it is indeterminate whether Jo is bald, it *is* possible to tighten up the meaning of 'bald' so that it is true, not indeterminate, that Jo is bald. And so it is possible to tighten up the meaning of 'bald' so that 'it is indeterminate that Jo is bald' is the wrong thing to say, as tightening up the meaning in the requisite way renders 'Jo is bald' determinately true. And so it is not *determinate* (or true) that it's indeterminate that Jo is bald.[10] Genuine indeterminacy must itself be indeterminate.

3. Semantic Paradox. 1. Wright complains that I do not make a case for rejecting the strengthened liar

L. $\neg$ (L is true).

His argument is intricate, but its nub, I think, is this. I claim that we can prove that the liar does not say anything true or false. Wright conjectures correctly that when I speak of proof, I have in mind something like the classical reasoning, from an instance of the T-Schema for L

---

[10] That it is never *true* that it's indeterminate whether S will sound odd. But recall that, given the above, 'it's indeterminate whether S' is apt iff (not S and not $\neg$S). The right hand side of this biconditional isn't the sort of thing that can be true or false. It can be *apt*, though, just as the claim that S is indeterminate can be.

LT. '¬ (L is true)' is true $\leftrightarrow$ ¬ (L is true)[11]

to a contradiction. He thinks that I will also depend on that reasoning to show that there is something amiss with endorsing LT.   Again, Wright has me aright. But he doubts that I can do both:

> …Richard cannot both treat the original paradoxical reasoning as showing that there is something wrong with the T-biconditional in full generality under its non-forced interpretation *and* treat that reasoning as showing that L says nothing either true or false. For in order for the reasoning to show the latter, it has to be *valid*. If it's valid, then its use of the normal T-schema is legitimate, we still have the paradox, and Richard's invocation of his forced interpretation [of the biconditional] is pointless at best. But if it is not valid, then it makes no case for regarding L as saying nothing either true or false, so no case for rejecting L.

I don't see the problem. If L has a truth value, then LT itself is true: there is nothing wrong with an instance of the schema

T. 'S' is true $\leftrightarrow$ S

when *S* is truth valued. But then, if L is truth valued, there can be no objection to reasoning from LT classically, including reasoning that involves assuming L or its negation for conditional proof: the classical rules of proof are undeniably sound when applied only to premises and assumptions that are truth valued. Such reasoning, of course, leads to a contradiction. Contradictions cannot be true; if L is truth valued, then a contradiction is true. So, we must reject L and its negation.

What of LT itself? Well, I would have thought that once we knew that L was neither true nor false, that settled the matter. LT, after all, is a *material* biconditional. So since it is not true (or false), it must be rejected.

I think that if a sentence S can be used to make an assertion, then there must be *some* sense in which it is correct to say that S is true if and only if what it is says is so. But this doesn't mean that *LT* is true. The forced biconditional

LTF. '¬ (L is true)' is true iff ¬L is true

gives the right thing to say about L. LFT is apt just in case it's apt to assert both what L says and that what L says is true, or apt to deny both. Since the latter is apt, so is LFT.

2. Wright wonders if I can give a unified account of the semantic paradoxes. In particular, he wonders whether I can motivate dismissing Curry's paradox by counseling that we reject its instances. Curry's paradox, recall, goes something like this. Let C be the sentence

C is true → there are as many sets of finite ordinals as real numbers.

Using classical conditional proof and an instance of T built on C, C is provable. I say that it would be precipitate to reject C. If the continuum hypothesis (C's consequent) *is* true, then C is too.[12]

What is surely, *obviously* wrong is the thought that simply by writing down a sentence of the form

X is true → T

and then christening the sentence so written 'X', we are guaranteed to have manufactured a truth. What is puzzling about Curry's paradox is that (a) we are inclined to accept the relevant instance of the T-schema, (b) we want to endorse the reasoning it involves, but (c) we don't want to say that we are thereby committed to the truth of the continuum hypothesis.

---

[12] The conditional in C is material. So, provided that C says something to begin with, it says something whose truth is guaranteed by the truth of its consequent. But C does say something, at least once the reference of 'C' as it occurs therein is fixed.

An advantage of the account I offer is that it *does* seem to help with this puzzle. The version of the T-biconditional for C that is valid whether or not C itself is true is the forced version

   TC. C is true iff (C is true → there are as many sets of finite ordinals as real numbers).

The "forced analog" of conditional proof involving forced conditionals,

   FCP: If T follows from a set of premises X and the sentence S, the (forced) conditional *if S, then T* follows from X,

is valid given the semantics in Chapter 2 of WT.[13] Thus, we can validly reason

   1.Suppose C is true.
   2.C is true iff (C is true → there are as many sets of finite ordinals as real numbers). (Logical validity)
   So, 3. C is true → there are as many sets of finite ordinals as real numbers. (1,2)
   So, 4. there are as many sets of finite ordinals as real numbers. (1, 3)
   Thus, 5. If C is true, then there are as many sets of finite ordinals as real numbers. (1-4, FCP)

This reasoning is perfectly good. However, it does not show that C is true, for the conclusion is not the material conditional C, but a sentence whose utterance performs a particular speech act s, one apt iff (C is not true, or its consequent is). The fact that performing s is apt does not imply C is true. So we can't get from what the reasoning proves to the continuum hypothesis's truth.

   Wright objects that I do not give a recipe for determining whether an arbitrary 'Curry sentence' (roughly, one of the form *if this sentence is true, then T*) is to be rejected.

---

[13] There is a discussion of the proof theory for a language with truth functional and force connectives in Richard, ms. B.

To give this one would have to give an account that decided whether and when an arbitrary 'ungrounded' ascription of truth could be true. I don't give an account of this. But I don't see that we need such an account in order to be justified in thinking that many paradoxical sentences say something neither true nor false and so are to be rejected. Nor do we need such in order to explain what goes wrong in the Curry paradox.

3. Wright, Scharp, and Shapiro press two related worries[14]:

(a) The languages I discuss are ones in which there are paradoxical sentences which involve either speech acts like denial or semantic notions (like appropriateness) that govern such acts; thus, I have not given an adequate account of the semantic paradoxes.

(b) Postulating a hierarchy of speech acts (asserting a negation; denying a claim; rejecting a denial; etc.) along with a hierarchy of semantic properties (truth; appropriateness; one that stands to appropriateness as appropriateness does to truth; etc.) as I do is for one or another reason objectionable. In particular, this leads to one or another sort of ineffability.

I think that (a) is simply wrong. It might be correct the approach I sketch implies that there is an ineffable property of "absolute rejection". But it is far from clear that this is so, and even less clear that this means that there is something wrong with the approach.

---

[14] Terminology: In WT, I use 'denial' for a *sui* generis speech act that is appropriate just when what is denied is false or truth valueless. In most of the book, I use 'denial' and 'rejection' interchangeably, but in the book's first appendix I reserve 'rejection' for a speech act that is appropriate when directed at speech acts that, as I put it, suffer from 'appropriateness gaps'. Wright departs from this terminology. He uses 'rejection' for the act I call denial. He uses 'denial' for 'a form of illocution of P appropriate when P is false in the same way and to just the extent that the assertion of P is appropriate when P is true. (Wright, note 1). That is, I take it, he uses 'denial of p' as I use 'assertion of the (truth functional) negation of p' in WT.

I stick with my terminology in these responses, in order to make the discussion continuous with WT. In this section (and only this section) I use 'rejection' only as it is used in Appendix 1 of WT.

17

The hierarchy WT gestures at arises as speakers reflect on the semantics of their language, add semantic vocabulary, and, having done this, acquire the ability to perform new acts with their language. We can describe its levels by describing what non-semantic and semantic vocabulary is available at a level, as well as what speech acts are routinely performed by users of the language.[15] WT's hierarchy is one not simply of languages *cum* interpreted abstract objects, but also of patterns of use, successive stages adding new semantic vocabulary and novel patterns of use.

At the base –call it level *Zero* –is a language with non-bivalent non-semantic predicates –words like 'big', 'heavy', and 'hairy' –and strong Kleene connectives and quantifiers. There is no semantic vocabulary, and the only (relevant) speech act performed by speakers is assertion.

The next level, *First*, represents what happens when speakers reflect on Zero, theorizing about truth and about the gappiness of truth in Zero. First's language has the same non-semantic vocabulary as Zero's. It also contains its own (Kripkean) truth predicate (and whatever semantic ideology –e.g., 'x refers in L to y' –is necessary to state a theory that characterizes the predicate). In First, we find speakers ascribing truth and falsity to sentences.[16] We find them disposed to assent to arbitrary instances of the sentential T-schema for sentences in Zero. But we also find that they don't accept bivalence: when a is a borderline case of the vague predicate F they say "a's not really F, and a's not really not F either"; they are inclined to say that the liar can't be true or false. This shows that they think they can say something correct about truth valueless claims. So in First speakers not only use the sentences of their language to make assertions, they also deny claims, as well as perform the speech acts associated with the force connectives *and*, *or*, *iff*, etc.

Note that in neither Zero nor First do we find (semantically) paradoxical sentences. Zero contains no semantic vocabulary. In First, speakers are able to frame liar sentences like

---

[15] More precisely: what speech acts are routinely associated with the use of the language's connectives, as denial is associated with negation.

[16] It would complicate the discussion in irrelevant ways to drag in propositions here.

L.   ¬L is true.

But this sentence isn't paradoxical.   The sentence

LT.   '¬ (L is true)' is true ↔ ¬ (L is true)

is not itself true, so we can't soundly reason from it to a contradiction. The force T-sentence for L

FT.   '¬ (L is true)' is true iff ¬ (L is true)

is indeed apt, but all we can derive from it is that L is not (denial) true and neither (denial) is L's negation.   This is not contradictory –it's appropriate; it's just the thing to say about L.   Though in First speakers are denying claims, First does not have the ideology to evaluate denials as appropriate or otherwise. So in First 'self-denial' does not occur (just as in Zero the strengthened liar is nowhere to be seen). There are no sentences or speech acts in First that lead to trouble.

      Furthermore, in First we are apparently able to say all there is to be said about truth-in-the-language-we-speak with that language. We can say of any sentence S whether it is true (with ' 'S' is true'), false ('-('S'is true'), or neither (by denying that S is true and denying that it is false). Using FT we can say for each sentence of the language under what conditions it is true or false.[17] And we can soundly theorize about truth –for example, we can soundly say that

> For any sentences R, S, and T: if T is the conjunction of S and R, then T is true iff (S is true and R is true),

---

[17] In saying this, I am thinking of the *language* of First as the result of adding the Kripkean truth predicate to the language of Zero. So that language does not include force connectives like 'not' and 'iff'. Rather, these connectives are part of a meta-language (like the one developed in Appendix A of WT) that *regiments* the *behavior* of speakers in First. Speakers *use* sentences of their language in several ways. It no more follows that those sentences are ambiguous than it follows from the fact that 'can you pass the salt?' is used to request the salt that this sentence is ambiguous.

the connectives and quantifiers having their force interpretation. We can reason using what are, formally, the same rules as we use in classical logic –so long as the uses of connectives are understood as force connectives and not as (the three valued) truth functional connectives.[18] At this stage in the hierarchy, we can apparently say all there is to say about the semantic and properties and relations our language enables us to talk about. And we can do so without paradox.

Of course reflecting on this situation, one will want to say more. When we reflect on First, we will eventually notice that we are *doing* more things with the sentences of the language than simply asserting claims. When we see this, we will want to theorize about the speech act of denial and the other speech acts associated with the forced uses of the connectives. We will thus add to the language a predicate true of speech acts when they are appropriate. At this point, we move to the next level in the hierarchy, *Second*. Here the non-semantic vocabulary is just that of First. The semantic vocabulary will be that of First augmented with an appropriateness predicate, defined in the way explained in Appendix I of WT. Once this predicate is part of the language, one can, by uttering

N. This sentence is not appropriate

in denial, produce something analogous to the liar. I suggested in WT that, just as we can see that there is something wrong with the liar, and say so by denying it, so we can see that there's something wrong with N. And we can say so –we would naturally use the idiom of negation, saying '*that*'s not right'. As I see it, when we do this, we are performing an act –rejection –that stands to denial somewhat as denial stands to asserting a negation. Denial is an act that we can only perform once we have reflected enough on assertion and truth to recognize that for some claims p, neither they nor their negation is assertible. Rejection is an act that we can perform only once we have reflected enough on denial and appropriateness to recognize that there are some acts whose appropriateness we can neither maintain nor deny. The speech acts performed in Second include (all those performed in First and) rejection, along with conjunctive and

---

[18] See Richard, ms. B, for a proof.

disjunctive acts that involve rejection. Of course, once we find ourselves in Second, we will reflect, and may add vocabulary that will allow us to evaluate rejection. This brings new expressive power and triggers the ability to perform acts that stand to rejection as it does to denial. Here is the next level in a (presumably) never ending series.

I didn't try to model Second in WT. What I suggested there –but didn't try to prove, or even rigorously formulate –is that in Second we would be able to convey in the language of Second, using rejection, other new speech acts whose performance accompanies the use of rejection, and the truth and appropriateness predicates, all there is to be said about truth and appropriateness of uses of the sentences in Second that makes assertions, denials, and perform the other speech acts regimented using the force connectives.   If this is correct, then in Second, it seems, we can say all there is to be said about what was going on, semantically, in the situation First. We can apparently also say, speaking Second's language, all there is to say about the semantic and properties and relations that language enables us to talk about. So far as I can see, no sentence in Second's language is paradoxical.

Because of all this, I conjecture that the hierarchy discussed in WT contains no paradoxical sentences. Certainly Scharp, Shapiro and Wright give no such reason to think otherwise. I think that objection (a) is just wrong.

The hierarchy just gestured at ought make us wonder whether the paradoxes indicate a serious philosophical problem about semantic notions such as the notion of sentential truth. If we think of the introduction of semantic predicates such as 'true' and 'appropriate' in the way just sketched, their introduction does not lead to paradox. Nor does it lead to our being unable to say something about the properties those predicates express. Once we introduce a truth predicate, we have a language in which we can say in our language –by asserting, denying, and performing other speech acts –all there is to be said about the distribution in our language of the property the truth predicate expresses. I conjecture that once we go on to introduce a predicate for evaluating denials – 'appropriate' –we will be in a position to say in our language all there is to be said about the distribution in our language of truth and the property the appropriateness predicate expresses. And so on up. If so, there don't appear to be a semantic property of sentences or acts performed by uttering sentences about which we are unable to adequately theorize

in the language of those sentences. Neither does there seem to be a semantic property whose ascription leads to paradox. As it says on the T-shirt, life is good.

What of objection (b), that (as Wright puts it) there is a perfectly intelligible property, "utter rejection", common to asserting the negation, denying a claim, rejecting a speech act, etc., that is ineffable on my account?   The charge is that there is some property that all and only the acts in the series

S: asserting a negation, denying a claim, rejecting a denial,…

(or all and only the members of some transfinite projection of this series) have in common.   I would be a liar, were I to say I don't feel the tug of the intuition that there is something all the acts in S have in common, something in virtue of which they are all importantly similar. I also feel the tug of the intuition that if there is a determinate property, being an ordinal, then there must be a collection of the things that have the property. The latter intuition, I take it, is wrong. I don't know what to make of the former one. I am not sure that if the account I give is correct so far as it goes, then there must be the sort of property Wright alludes to. In part this is because I don't know just what I would be agreeing to, if I agreed that there was such a property. If properties are as common as sets, then there might be such a property. We couldn't define it, but I'm not sure why that is problematic. If the sort of property Wright has in mind is a "sparse" property, I'm not sure what to say. Truth and appropriateness seem very different properties, as are falsity and lack of appropriateness. So asserting a negation and denying a claim are, I think, very different properties. I'm not a big fan of the notion of a 'natural property', but I think the disjunction of asserting a negation and denial is less natural than either of them. So I'm not sure that there is particularly "natural" property corresponding to sequence S. If "sparse properties" are natural ones, then probably isn't a sparse property corresponding to S.

I do deny that there is a natural property corresponding to S that we would ever be a position to grasp. As I see it, one can't grasp the relation of denial until one has meditated a bit on the fact that truth is gappy –and thus not until one has meditated a bit on the notion of truth. Likewise, one can't grasp the relation of rejection until one has

meditated some on the fact that appropriateness is gappy –and so not until one *has* that notion. And so on up.

Suppose that the story I am telling demands, for some reason, that there be a natural property corresponding to S. Does that tend to show that the account I suggest in WT is inadequate, wrong, or unsatisfying? Perhaps it does, but I don't see that Wright, Scharp, or Sharpiro have given any reason to think this. Let S be any set of semantic properties applicable to the acts we might perform in a natural language we speak, all of whose members we might at some one time grasp. If WT's account is on track, we can theorize about the distribution of all the members of S to the sentences and acts performed by uttering the sentences of a language L in that very language in an adequate way without the danger of contradiction. Such an S is finite and might be expanded to a larger set of semantic properties about the distribution of which (in an extension of L) we can adequately theorize (in that extension of L) without encountering contradiction. Why should we think this situation unsatisfactory? What theoretical or philosophical purpose might we have that is in principle unsatisfiable, if all this is so, but the semantic properties of languages we could never speak elude our grasp?[19]

---

[19]Two final comments. First, I want to point out that objections (a) and (b) are pretty orthogonal to the primary point of suggesting that everyday reasoning involves speech acts like denial that aren't reducible to assertion. The primary point of that suggestion is that when we understand our reasoning in this way, it can be understood as sound *without* assuming bivalence (or invoking supervaluations): reasoning we **take** to be sound that is not sound when the connectives are given a truth functional interpretation **is** sound when they are given a forced interpretation. (Again, see Richard, ms. B.) Furthermore, given the systematic correlation between our sentences and the range of speech acts discussed in WT, our reactions to paradoxical sentences –e.g., that the strengthened liar is *not* true, even though it says that it is not true and that, quite generally, sentences are true iff what they say is so –these reactions are sound. Surely this is good reason to accept the idea that there is more to saying than assertion, even if it turns out that a complete account of the semantic paradoxes demands more than the notions of assertion and denial and (some extension of) Kripke's strategy for defining truth in a language.

Secondly: I should address Wright's claim that the notion of denial is 'unintelligible' because if we say that denying L "states the fact that L says nothing true or false, we are asserting that L says nothing true or false." I think there is a perfectly good sense of 'say' in which we can report what someone does when they assert or deny a claim by uttering *he said that S*, where S (intuitively) translates their utterance into our idiom. So if John denies that sentence L is true by saying 'L is not true', I can report him

4. Faultless Disagreement. Relativists often motivate relativism by claiming that the view best explains "faultless disagreement" –situations in which people have incompatible beliefs but in which neither is in any epistemic way at fault. The standard example of such disagreement is a difference in taste, as when I think coriander tasty and you think it not tasty. The relativist says that if it can be true for me that coriander is tasty while it's false for you, we can have contrary opinions about coriander while neither of us is in an epistemically defective state.

If expressivist views are coherent –and part of the burden of the the first half of WT is to argue that they are –faultless disagreement can't by itself motivate relativism, since an expressivist can presumably also tell a tale on which we may make incompatible taste judgments while neither of us is epistemically blameworthy. If an expressivist account is to be genuinely different from a relativist account, the expressivist must say that such judgments are not ones whose objects are true or false, and so she will have to give some account of incompatibility that does not explain it in terms of the impossibility of joint truth. Chapter 5 of WT sketches what might be called an expressivist account of judgments of taste, one on which there may be faultless difference over incompatible taste judgments. I think the positive account given in that chapter is tenable; it might even be correct.[20]

However, there is a false step in the chapter. In it, I give a simple argument that faultless *disagreement* in sense the relativist has in mind is impossible. The argument was this: The relativist says that to *disagree* is to have beliefs (or make assertions) that can't both be true. Suppose that I believe p (which is true in my situation), and you believe ¬p (which is true in yours). For a disagreement to be faultless, it must be that we have incompatible beliefs that are both true. But there is no situation relative to which both our beliefs are true, and so no situation relative to which our disagreement is

---

by uttering 'John said that L is not true'. And I would say, using 'say' in this way, that to deny a claim p is to say p is not true. I appreciate that one might want to see a detailed semantics for this use of 'say'; it doesn't appear to me to be difficult to provide, but perhaps I am missing something.
[20] I am, however, less enamored with expressivism about taste than I was.

faultless.[21] MacFarlane rightly criticizes this argument. It doesn't work, and faultless disagreement as the relativist conceives of it is indeed possible. I have discussed my argument and the topic of disagreement elsewhere, and I don't propose to repeat that discussion here.[22] I'll instead limit myself to two brief comments.

First off: One of the conclusions I drew in my discussion of claims about taste is that it might vary from perspective to perspective as to whether judgments about taste are judgments about a truth apt claim or are "merely" expressions of attitudes of approving of or valuing an object. This is a sort of "interpretivism" about such judgments, of a piece with (but somewhat more radical than) the sort of interpretivism that holds that the (truth evaluable) content of a judgment is at least in part determined by he who interprets the judgment. I think the latter sort of interpretivism about humdrum judgments is quite plausible. It is, for example, correct for to say that the Greeks shared our belief that water is wet. But the correctness of this is not settled simply by semantic properties ancient Greek and modern English words (or concepts) have independently of any interpretive context. In some contexts it is correct to observe that the ancient Greeks and we do not refer to precisely the same thing with 'water' and Ὕδωρ'. That it is correct in other contexts to ascribe the Greeks water beliefs is in part the result of a *decision* on our part, to interpret them in a particular way. I see no reason to think, if this sort of interpretivism is correct, that the more radical interpretivism might not also be correct.[23] I retract the argument that MacFarlane criticizes, but continue to think that the view I used it to motivate is coherent and quite possibly correct.

Secondly: Even if we adopt a relativist account of judgments of taste, we should, I think, be chary of characterizing most differences in taste as *disagreements*. As I see it, there is genuine disagreement between people with incompatible beliefs about a claim p only when there is reason for them to have the same attitude about p. That this is

---

[21] I take relativism to be the view that the truth of some claims varies within a world as something that depends on human interests varies. I call the thing this latter thing a perspective. For a relativist, claims are true or false relative to a world and a perspective. I use *situation* as shorthand for these combinations of a world and a perspective.

[22] Richard, ms. C. That paper, in turn, revises some of the things that I said about disagreement in Richard 2011.

[23] There is an argument for a kindred sort of interpretivism, about what's said by slurring speech, in the last few sections of Chapter 1 of WT.

necessary for genuine disagreement becomes clear, I think, when we mediate on an example derived from MacFarlane.[24] : you see John standing and think that John is standing; your near twin in a near possible world sees John sitting and thinks that John is not standing. Do you disagree? Of course not. Why not? Well, obviously, because there is no reason for you two to agree. You are in different possible worlds. Genuine disagreement is a matter not simply of incompatible beliefs, but of the believers being such that there is reason for them to agree.

When we say of people with incompatible beliefs that they disagree, we invariably presuppose that there is (all things considered) good reason for the two to have the same opinion about the matter at hand. But the standard examples of "faultless disagreement" –I think okra tasty, you think it not –are naturally understood as cases that *aren't* disagreements in this sense. There is generally no reason we should agree about food's taste; it is this, I say, that makes us want to say that differences over okra and coriander aren't really disagreements.

It is not that we can't motivate relativism by appeal to the claim that it best explains our intuitions about disagreement. We can. But if we are to do this, we should present cases that are clearly cases of genuine disagreement, cases in which those whose beliefs are at odds and those who observe them have a strong intuition that *someone* is epistemically at fault because the disputants *ought* to agree on the issue over which they differ. I try to present just such cases in Chapter 4 of WT.

Philosophy Department
Harvard University
Richard4@fas.harvard.edu

Broome, John. 1998. Is Incommensurability Vagueness?. In Chang, R., ed. *Incommensurability, Incomparability,and Practical Reason*. Harvard University Press.

---

[24] This occurs in MacFarlane 2007.

MacFarlane, John. 2007. Relativism and Disagreement. *Philosophical Studies 132*.

Richard, Mark. 2010. Indeterminacy and Truth Value Gaps. In Dietz, R. and Moruzzi, S., eds. *Cuts and Clouds. Vagueness, Its Nature, and Its Logic*. Oxford University Press.

Richard, Mark. 2011. Relativistic Content and Disagreement. *Philosophical Studies*.

Richard, Mark. ms. A. What Would an Expressivist Semantics Be? Posted on addURL

Richard, Mark. ms. B. Classical Logic with Truth Value Gaps and without Supervaluations. Posted on add URL.

Richard, Mark. ms. C. What is Disagreement? Posted on add URL