The Liar and Sorites Paradoxes: Toward a Unified Treatment

Jamie Tappenden

The Journal of Philosophy, Volume 90, Issue 11 (Nov., 1993), 551-577.

THE JOURNAL OF PHILOSOPHY

VOLUME XC, NO. 11, NOVEMBER 1993

THE LIAR AND SORITES PARADOXES: TOWARD A UNIFIED TREATMENT*

This paper develops one aspect of a program aimed at studying normative constraints on the evolution of language. The theme to be developed is that vague predicates may usefully be seen as analogous to the truth predicate.[1] In this paper, one respect of similarity will take center stage: both the truth predicate and vague predicates support speech acts which are initially very puzzling and which admit of a uniform explanation.

The first section presents two perplexing examples on which the paper will focus. In the second section to the fourth, the behavior of vague predicates under increases in precision is examined, with the upshot that our use of vague predicates requires us to recognize an important class of sentences whose distinguishing feature is that they are not acceptably called false (though they need not be true). In the fifth, the way in which the "nonfalseness" of such sentences may be exploited by speakers to perform speech acts is explored. In the final section, these reflections are carried over to account for some counterintuitive consequences of an otherwise appealing theory of the truth predicate. To display the relevant phenomena clearly requires simplifying assumptions that limit the extent to which this investigation may claim to be the ultimate resolution of liar and sorties paradoxes. The point here, however, is not to provide definitive solutions but rather to uncover and render compre-

* I am indebted to audiences at Princeton University, Sidney University, University of Michigan, and University of Pittsburgh for helpful discussion. Special debts for comments and conversation are owed to Paul Benacerraf, Bob Brandom, Anil Gupta, Saul Kripke, Gideon Rosen, Scott Soames, Rich Thomason, Paul Thomson, and Bas van Fraassen. I also owe an overwhelming debt to the writings of Kit Fine.

[1] These ideas were explored also by Vann McGee, *Truth, Vagueness, and Paradox* (Indianapolis: Hackett, 1991).

0022-362X/93/9011/551-77

hensible shared patterns of use so as to illuminate the nature of both vague predicates and the truth predicate.

I. FOR ORIENTATION

To focus attention, consider this variant of the liar paradox:

(1*) For all terms a and sentences S, if a refers to S then (a is true iff S).
(2*) 'k' refers to "k is not true" (by stipulation).

Therefore, (3*) k is true iff k is not true.

(3*) follows by modus ponens and two applications of universal instantiation. But (3*) is self-contradictory; whatever may be said of (3*), one should not call it true.

It might be suggested that we regard (1*) as attractive because we have made a simple error of overgeneralizing from the ordinary run of cases. The proponent of this line might suggest that when we realize that (3*) is a consequence of (1*) we should see the error of our ways and restrict the quantifiers in (1*) to untroublesome instances. But this suggestion neglects how attractive (3*) is. (3*) is self-contradictory and hence cannot correctly be said to be true, but it also seems to capture one feature of the sentence k in virtue of which k is paradoxical! Does this not sum up nicely what is funny about k: "k is true if and only if k is not true"?

This variant of the liar paradox is especially puzzling because of the naturalness of the utterance of the biconditional (3*). Imagine that you are explaining the liar to someone who does not catch on right away. You might well say: "Here is what is funny about k. If k is true then it is not true, and if k is not true then it is true." You have apparently just contradicted yourself by uttering a variation on (3*). But (3*) seems to be the right thing to say in the situation. And does it not get across precisely what is odd about the liar? (3*) apparently can be used to make assertions in a way that requires the truth of the sentence; yet the sentence cannot be true. The usual liar sentences have for most of us no particular intuitive pull toward affirmation or denial, while here we do have intuitions that instruct us as to how the sentence should be assessed. But the intuitions conflict.

A similar concern arises with some sorites paradoxes. Recall:

(1) A man with no money at all is poor.
(2) For all n, if a man with exactly n cents is poor, then a man with $n + 1$ cents is poor.

Therefore, (3) however much money a man may have, he is poor.

The conclusion follows by mathematical induction, or repeated applications of universal instantiation and modus ponens.

Although the details differ, most treatments of the sorites paradoxes point to (2) as the culprit. But a problem arises here as well. Say that someone consistently demonstrated a misunderstanding of the proper use of 'looks red to me' or 'observationally indistinguishable', saying that one color sample looked red to him and another did not, while maintaining that the two samples in question were observationally indistinguishable to him. You might say: "*Look, if any two samples are observationally indistinguishable to you then one looks red to you if and only if if the other looks red to you.*" The remark has been put in italics to induce the reader to look at the sentence with assessing eyes. It sounds like a natural thing to say in the imagined circumstances, and it would appear to be innocuous when hazarded in everyday conversation, but it appears to commit the speaker to the crucial conditional premise of the sorites paradox. In conjunction with some indisputable facts about the existence of settled cases and sorites chains of indiscernible color samples, it yields a contradiction. Yet the italicized sentence seems like just the thing to say given that one trying to correct the other person's mistakes. And the utterance will probably succeed; that is, it probably will have the desired effect of getting the hearer to stop talking in the proscribed manner. But if we are attempting to say something true, we are failing.

I want to explain not only what is wrong, but also what is right, about the intuitions that would produce the envisioned speech episodes. It is not enough to explain away these intuitions by showing that they do indeed mislead us, and by detailing how we come to be misled. Because we have these intuitions, and we may presume that others have them as well, we can successfully accomplish communicative ends. Someone who lacked the inclination to regard the boldface sentence or (3*) as in some way or other distinguished would be missing something of significance.

I want more than just something or other to say about the sorites and liar paradoxes which would block the unwanted inferences. I am looking for an account which is independently motivated and which tells a coherent story about the speech acts performed in the imagined situations. In the next section, I shall expand on these brief remarks and propose a set of goals for a treatment of vague predicates generally.

II. DESIDERATA FOR AN ACCOUNT OF VAGUE PREDICATES

The account developed here focuses on a ubiquitous, but neglected, feature of our use of sentences containing vague predicates: the way that the extension of a given vague predicate varies according to circumstances. We augment and relax the precision of vague predicates. Sometimes more things, and sometimes fewer things are counted as instances. This happens all the time: if a certain predicate is appropriate for some linguistic task, but insufficiently finely calibrated, we may lay down a more precise delineation by fiat. We might just say: "Let's count these among the heavy ones." (Said of some unclear cases.) The extent of the indeterminacy a predicate exhibits may be reduced or increased in a given context. Here I shall concentrate on cases where the amount of indeterminacy has been reduced: in such a case I shall say that the predicate has undergone an increase in precision. Such increases may be temporary or permanent; it will be helpful to consider some examples.

For cases of permanent increases in precision, one looks naturally to the history of units of measurement. A straightforward and instructive example is 'second'. In 1967, the standard was changed from a certain fraction of a fixed tropical solar year to one that turned upon a cesium 133 atomic clock. The 1967 agreement changed the extension of 'is an interval of one second', and the change was in a certain respect quite extreme—the definition turned upon a newfangled atomic device instead of the motion of the earth. But there is also a respect in which the change was a conservative one. It is a criterion of success for a stipulation of this sort that almost any interval that counted as an interval of more than one second before the change counts as an interval of more than one second after it.

Susceptibility to the sort of change illustrated by this case marks vague expressions off from precise ones. It might be objected: "True, 'second' underwent a change in meaning, but so too (relatively) precise expressions can undergo a change in meaning, and frequently some do." Did not 'prove' once mean test, as we are reminded each time we hear 'That's an exception which proves the rule'? Well, there are changes and then there are changes; in the case just considered, the change in meaning is importantly different from that undergone by 'prove'. Although the meaning of 'second' differs before and after, the extension is for the most part preserved: previously undecided cases are decided one way or another, but almost all of the cases that had been decided one way or another remain classified as they had been. Furthermore, this conservativeness is a condition of acceptability placed upon the new definition of 'second': virtually all of the intervals that counted as instances of 'is an interval of one second' before the change should count as instances of that predicate after it.

Permanent increases in determinateness are less common than temporary, context-specific ones. Consider a phrase like 'too heavy to be lifted safely'. Someone drafting occupational health and safety regulations might specify precise boundaries by writing something like: "All objects too heavy to be lifted safely (i.e., over 90 kg.) must be moved with a forklift." This does not commit the author of the regulations to speak with such precision all the time. Rather, it indicates that, in the context of the discussion and application of the regulations, the predicate is to be understood as having the precise boundaries specified. In a different context, the predicate may have the usual range of indeterminate cases.

The choice to stipulate a boundary at 90 kg. rather than 95 kg. may well be arbitrary. The situation is one in which some sharp boundary is called for, though which boundary among those within some range should be chosen is not determined by the information given. Certainly some boundaries may be preferred as more salient or convenient than others. There may be reason to choose boundaries closer to the clear cases than to countercases or conversely. But even after all of the contextual influences have been taken into account, some arbitrariness in the choice of stipulation remains.

Increases in precision need not resolve all the vagueness in a predicate: sometimes it is convenient to stipulate away one or a handful of borderline cases while leaving many others undecided. Examples of this sort appear frequently in legal contexts. A question may face a judge, in which the proper answer is not determined by the existing body of written law and precedent. There may be no decision that is warranted by existing positive law. Bearing in mind that hard cases make bad law the judge may rule narrowly on the question, rather than deliver a broad principle to resolve not only the narrow question but also many others that had been similarly undetermined. Sometimes such judicial resolution of indeterminacy may be seen as the stipulation of an increase in precision for a vague predicate. American court decisions in which the powers of the various branches of government are delineated provide useful examples. The proper application of the phrase 'is a legitimate function of the legislative branch' may be taken to be fixed by the articles of the U.S. Constitution, acts of Congress, and relevant judicial decisions. So understood, the predicate is vague at least to the extent that there are questions concerning separation of powers which are left unsettled by existing law.

For an example, suppose that when the legislative veto was brought before the Supreme Court, the restrictions upon the Court's decisions did not determine whether 'is a legitimate function of the U.S. Congress' applied or failed to apply. If so, the court simply had to decree that Congress would possess, or that it would lack, the power claimed. To put it in the formal mode: the court had to lay down by fiat that the legislative veto was, or was not, in the extension of 'is a legitimate exercise of congressional powers'. Prior to the court's ruling, there was no fact of the matter. After the ruling, 'is a legitimate exercise of congressional powers' remains vague, though slightly less vague than before.

To say, as here, that the Court had some discretion is not to say that it was completely unfettered. It could not award the power both to Congress and the judiciary. It could not award it to the state of West Virginia, or the parliament of the Netherlands, or David Lewis. There are constraints upon the increases in precision that 'is a legitimate exercise of congressional powers' can admit.

These examples display the aspect of language I want to clarify: the mix of license and restriction which governs increases in precision, the way such sharpenings can be arbitrary but nonetheless within constraints. One restriction on arbitrariness is given by settled instances and counterinstances. For each predicate P there is a (possibly empty) set of objects to which the predicate clearly applies and another to which it clearly fails to apply. Call these the extension and anti-extension of the predicate; the remaining objects will be called borderline cases. The general question may be reassessed: In what ways is it acceptable to add objects from the borderline range to the extension and anti-extension? I shall use the phrase '*admissible extension*' to describe a semantic structure that results from another by effecting an acceptable increase in precision. That is, the admissible extensions are those which result from the acceptable ways of drawing sharper boundaries.

An artificial example (inspired by one devised by Scott Soames and Nathan Salmon) will be helpful. Say we introduce into English a predicate 'tung' whose use is governed just by the following rules: (a) 'tung' applies to anything of mass greater than 200 kg; (b) 'tung' does not apply to anything of mass less than 100 kg. The applicability of 'tung' to objects of mass between 100 and 200 kg. inclusive is left unsettled. We could introduce a predicate of this type into the language and it could come to be used just as any new predicate could be. 'Tung' would behave in certain respects like 'heavy', though it would differ in one crucial way:[2] if our understanding of

[2] Of course, the respect that is noted in the text is not the only one. Certainly 'heavy' is higher-order vague while 'tung' is not, and 'heavy' is context-sensitive in ways that 'tung' might not be.

'tung' is completely given by the above rules, then 'tung' allows a type of increase in precision which 'heavy' does not. Say that *a* and *b* have counted neither as tung nor as not tung, but they must now be taken to count one way or another, and say that *a* is heavier than *b*. The rules for 'tung' give no reason why we should not stipulate that *a*—the heavier one—counts as nontung while the lighter *b* gets counted as tung. In this easily overlooked respect, 'tung' differs from the otherwise similar 'heavy': when increasing precision for a collection of borderline cases one must respect the order of comparative weight in the case of 'heavy' while there is no such constraint in the case of 'tung'.

Rather than speak abstractly of constraints on increases in precision, we can talk of restrictions on assignments of truth values to sentences. 'Heavy' differs from 'tung' in the following crucial respect: anyone who understands the meaning of 'heavy' and who knows that *a* is heavier than *b* knows that new boundaries may not be stipulated in such a way that 'If *b* is heavy then *a* is heavy' becomes false, while it is acceptable to draw boundaries in such a way that 'If *b* is tung then *a* is tung' becomes false.

The above discussion suggests one collection of constraints on increases in precision: those which may be formulated, "Never make words $w_1, \ldots, w_n$ more precise in such a way that sentence *S* becomes false." Sentences that constrain increases in precision as *S* does in this example will be called pre-analytic: a sentence *S* is *pre-analytic* if anyone who understands *S* knows not to draw more precise boundaries to any of the expressions in *S* in such a way that *S* would be false in any circumstances. Intuitively, pre-analytic sentences are like analytic sentences, except that analytic sentences are always true, while pre-analytic sentences are never false. All analytic sentences are pre-analytic, but the converse may not hold. An example of a pre-analytic sentence is 'If *a* is heavier than *b* and *b* is heavy then *a* is heavy'. Even if *a* and *b* are borderline cases of heavy things, so that it is acceptable to specify new boundaries on which '*a* is heavy' comes out true or on which it comes out false, and similarly for '*b* is heavy', there is a restriction on how one may resolve both at once. When sharpening the predicate to resolve application to both *a* and *b*, boundary drawing is constrained at least in that 'If *a* is heavier than *b* and *b* is heavy then *a* is heavy' must not come out false. Note that if a pre-analytic sentence has any truth value at all, it must be true.

Call a sentence *indeterminate* if there is an admissible extension that counts it true and one that counts it false. It will be assumed for simplicity that, if *a* is a borderline case of *P*, then '*a* is *P*' is indeterminate. Note that it is *not* stated that all sentences that are neither true nor false (i.e., have no classical truth value) are indeterminate (i.e., may come to have either classical truth value as a result of an acceptable increase in precision). Also, I adopt a principle I shall not defend here: indeterminate sentences lack truth value, so that an indeterminate sentence is neither correctly called true nor correctly called false.

Logical compounds of indeterminate sentences serve up a dilemma. In the following examples, *a* and *b* are borderline tall and borderline heavy, respectively. In most cases, it seems correct to count as indeterminate compounds like 'If *p* then *q*' or '*p* and *q*' of indeterminate sentences *p* and *q*. Given that '*a* is tall' and '*b* is heavy' are indeterminate, it appears that '*a* is tall and *b* is heavy' and 'If *a* is tall then *b* is heavy' are indeterminate as well. But there are some delicate exceptions, like 'If *b* is heavy then *c* is heavy' where *c*—like *b*—is borderline heavy, but *c* is heavier than *b*. There are two significant things to note about this sentence. First, given the relative weights of *b* and *c*, it is inadmissible to stipulate sharper boundaries in such a way that the sentence comes out false. So the sentence is not indeterminate. But what is more striking is that the sentence appears to be true, even though the constituents are indeterminate. This brings out the problem: 'If *b* is heavy then *c* is heavy' does indeed appear true, but it does not appear "to get to be true" in the right way or for the right sort of reason. We have two conflicting intuitions about the truth value—if any—that a sentence like 'If *b* is heavy then *c* is heavy' should be taken to have.

First consider what will be called the *truth-functional intuition*, which decrees that sentential connectives 'and', 'if . . then', 'or', and 'not' should correspond to truth functions. That is, assignments of truth values to sentences containing connectives should be uniform functions of the truth values of the constituents and conversely no truth value should be assigned to a sentence containing a connective unless the same value is assigned to every sentence with the same connective joining constituents with the same values. More loosely speaking: if sentences S_1 and S_2 have the same form then they have the same truth value if and only if their sentential constituents have the same truth values.

I shall use the phrase 'truth-functional intuition' for the impulse to construe the connectives as truth functions, though in some cases—particularly with conditionals—there is no strong untutored intuitive pull toward a truth-functional treatment. The tendency toward truth functionality is an amalgam of several different motivations. We have unrefined intuitions that pull us toward a truth function in many particular cases (especially when the principal connective is 'or' or 'and'), as well as philosophical reflections on substitution, compositionality, extensionality, simplicity, and whatever else leads philosophers to have a preference for truth-functional accounts of the connectives. Since my concern here is not to defend the truth-functional intuition but rather to use it to help illustrate tensions in our goals for accounts of vague expressions, this combination of intuitions and theoretical considerations is sufficient motivation.

Examples illustrating the truth-functional intuition are easy to find. One might feel that, because '*a* is tall and *b* is heavy' is neither true nor false, every instance of '*p* and *q*' must lack truth value when the sentences substituted for *p* and *q* lack truth value. Hence, given a systematic assignment of the classical truth values incorporating the truth-functional intuition (i–iii) jointly ensure 'Harry is bald and Melvin is sleepy' lacks truth value: (i) 'Harry is bald' lacks a value (ii) 'Melvin is sleepy' lacks a value (iii) when '*a* is tall' lacks a value and '*b* is tall' lacks a value, then '*a* is heavy and *b* is tall' lacks a value.

Now, say that we accept the truth-functional intuition, and we accept that at least some instances of 'If *p* then *q*' are neither true nor false, when the sentences substituted for *p* and *q* are indeterminate. It follows that we accept that all (material) conditionals with indeterminate components are neither true nor false. Then some pre-analytic sentences, like 'If *a* is heavy then *c* is heavy' (with *a* and *c* borderline heavy, and *c* heavier than *a* as above) will be assigned no value. But now a problem confronts us: in most contexts, when *b* is heavier than *a*, it seems that 'If *a* is heavy then *b* is heavy' is true, even if *a* and *b* are both borderline heavy.[3] Relative to some assignment respecting the truth-functional intuition, pre-analytic sentences which are assigned neither true nor false but which we are strongly inclined to regard as true will be called (using terminology due to Fine) *penumbral sentences.* The example just considered conflicts with the truth-functional intuition insofar as it shows that there are some penumbral sentences. There are pre-analytic sentences which we are strongly inclined to regard as true but which would receive no truth value if truth assignments conform to the truth-functional intuition. I shall call the intuition that some penumbral sentences are true the *penumbral intuition.* A central leitmotif in

[3] Strictly speaking, this sentence is not pre-analytic, but to keep the example simple it is useful to treat it as one. The example would be strictly correct, though more convoluted, if 'If *c* is heavier than *a* and *a* is heavy then *b* is heavy' were used, with the appropriate changes elsewhere in the example.

what follows will be the development and resolution of the conflict of the penumbral and truth-functional intuitions.

The notion of penumbral sentence will be crucial, so it will pay to consider a few more cases. It is not difficult to find apparently indeterminate sentences of the form '$\neg(p\&\neg q)$' with p and q indeterminate. For instance, with Harry and Melvin as before, both 'Harry is bald and Melvin is not sleepy' and 'It is not the case that Harry is bald and Melvin is sleepy' are indeterminate, and so they lack truth value. The truth-functional intuition gives the conclusion that all sentences of the form $\neg(p\&\neg q)$ and $(p\&\neg q)$ should lack truth value, if the constituents lack truth value. Conflicting with this, instances of '$\neg(p\&\neg p)$' typically appear true and instances of $(p\&\neg p)$ appear false even if the constituents are indeterminate. Given that 'Harry is bald' and 'Melvin is sleepy' are indeterminate, and so lack truth value, it appears that we have to give up one of three things: (a) the secure intuition that it is acceptable to stipulate—at least temporarily—new boundaries for 'bald' and 'sleepy' so that 'It is not the case both that Harry is bald and that Melvin is not sleepy' may come out true or may come out false (so the conjunction is indeterminate and hence lacks a truth value); (b) (the penumbral intuition) the apparently equally secure intuition that 'It is not the case both that Harry is bald and that Harry is not bald' is true; or (c) (the truth-functional intuition) the view that assignments of truth values to sentences containing 'and' and 'not' are uniform.

The intuitive appeal of some penumbral sentences turns not only on logical form, but also on meaning relations among predicates. Say for the sake of the example that 'x is lanky iff x is tall and x is slender' defines 'lanky'. Say that Stretch is borderline lanky and borderline slender: even so, one would be inclined to call 'If Stretch is lanky then Stretch is slender' true, just because of what 'lanky' and 'slender' mean. This is not the judgment the truth-functional intuition delivers. One can find penumbral sentences that appear to turn upon relations of exclusion among predicates: one might feel that 'a is not both orange and red' is true even if a is borderline orange and borderline red. Sometimes ancillary information might seem to ensure the truth of a penumbral sentence. In a case where it is indeterminate whether F is a legitimate function of Congress, and indeterminate whether F is a legitimate function of the executive, one might hold that 'F is a legitimate function of Congress or F is a legitimate function of the executive' is true if it can be determined that F is a legitimate function of some body or other, and it can be shown not to be a legitimate function of any body but Congress or the executive.

There appears to be a connection between penumbral sentences and semantic constraints on increases in precision. One reason penumbral sentences appear true is that to imagine boundaries that would count the sentence false does violence to one's understanding of what some predicate means. The most straightforward way for a sentence never to be correctly called false is for it always to be correctly called true. And to be sure, many penumbral sentences seem to be the sort of sentence that has traditionally been called "analytic."

To sort through these questions, we need a more rigorous treatment of the underlying semantic structure which will allow us to frame clearer definitions of notions like penumbral and pre-analytic sentence. From the above, we may isolate four goals for the investigation:

(1) Shed some light on the sorites paradoxes.
(2) Develop a framework in which the relation between the meaning of a predicate and the acceptable ways to make the predicate precise can be clarified.
(3) Characterize penumbral sentences and assess their semantic status.
(4) Explain the puzzling speech behavior considered in the introduction.

The focus will be narrowed in two ways. There will be no attempt to provide a systematic theory of the context sensitivity of vague expressions and higher-order vagueness (vagueness in the range of clear cases and countercases) will not be addressed in any systematic way. It will be assumed, to simplify things, that there are sharp boundaries to the range of cases and countercases of a given vague predicate. This limits the scope of the investigation—note that (1) reads 'shed some light on' rather than 'provide a robust solution to'—but they allow the phenomena to be clearly presented.

III. ESTABLISHING A FRAMEWORK

The first step is to adjudicate the semantic status of indeterminate sentences. It will be assumed that it is incorrect to say that such a sentence is true and it is incorrect to say that it is false. Indeterminate atomic sentences are—so to speak—semantically *hors concours.* The basic intuition is that, if a sentence is true (false), the truth (falsity) of the sentence is irrevocably settled by the empirical facts and the meaning of the sentence. (Here, of course, I mean an 'eternal sentence' in W. V. Quine's sense.) Since indeterminate sentences admit of some admissible stipulations that result in their being deemed true and they admit of others that result in their being deemed false, it cannot be right to count them as true. Loosely, if some sentence is true it is settled as true; truth is settled truth.

With truth understood as settled truth, it is natural to take falsehood to be settled falsehood. Negation will be understood univocally as internal negation: it behaves classically on classical values and the negation of a sentence which exhibits a gap itself exhibits a gap. A sentence S will be called false if $\neg S$ is true. Although it is a bit misleading, I shall sometimes relax my use of language and say that indeterminate sentences are neither true nor false.

For orientation, it is helpful to begin with a semantic account that just misses doing everything: supervaluational semantics.[4] The idea is that a sentence counts as true (false) if it is true (false) on all ways of making the predicates in the language perfectly precise. No value is assigned to sentences that obtain different values on different ways of sharpening the predicates. To facilitate comparisons between the truth-functional and penumbral intuitions, the presentation of supervaluations will begin with a gap theory that elaborates the truth-functional intuition. Rather than dispute the details of the partial scheme chosen, the strong Kleene scheme will be fixed.[5] This affords a formal definition of a structure representing satisfaction on the basis of the truth-functional intuition. A *pre-assignment* M is a partial model—an ordered pair $\langle D_M, I_M \rangle$ such that D_M is a domain of objects and I_M is an interpretation function that assigns objects $I_M(a)$ and $I_M(x)$ from D_M to names a and variables x and an ordered pair consisting of two disjoint sets of n-tuples of objects $I_M(P) = (E_{MP}{}^+, E_{MP}{}^-)$ (the extension and anti-extension of P) to each n-ary predicate P. Satisfaction ($\models$) of formulae by the model may be defined in a straightforward way, but now there is a need also for a parallel account of falsification ($=\!\!|$). (The metalanguage connectives and quantifiers are to be interpreted classically.):

For P atomic, and with t_i a variable or a name for $1 \leq i \leq n$:

$M \models Pt_1, \ldots t_n$ iff $\langle I_M(t_1), \ldots I_M(t_n) \rangle \in E_{MP}{}^+$

$M =\!\!| \, Pt_1, \ldots t_n$ iff $\langle I_M(t_1), \ldots I_M(t_n) \rangle \in E_{MP}{}^-$

$M \models \neg\theta$ iff $M =\!\!| \, \theta$ $\quad M =\!\!| \, \neg\theta$ iff $M \models \theta$

$M \models (\theta \vee \Phi)$ iff $M \models \theta$ or $M \models \Phi$ $\quad M =\!\!| \, (\theta \vee \Phi)$ iff $M =\!\!| \, \theta$ and $M =\!\!| \, \Phi$

[4] The notion of a supervaluation was introduced in van Fraassen, "Singular Terms, Truth-Value Gaps and Free Logic," this JOURNAL, LXIII, 17 (September 15, 1966): 481–95. The notion was extended to the semantics of vague expressions by inter alia: H. Kamp, "Two Theories about Adjectives," in E. Keenan, ed., *Formal Semantics of Natural Languages* (New York: Cambridge, 1975), pp. 123–55; Fine, "Vagueness, Truth and Logic," *Synthèse,* XXX (1975): 265–300. The importance of penumbral sentences was first emphasized clearly in the Fine article.

[5] For details on the strong Kleene scheme, cf. S. Blamey, "Partial Logic," in M. Gabbay, and F. Guenther, eds., *Handbook of Philosophical Logic* (Dordrecht: Reidel, 1986), pp. 1–70.

$M \vDash (x)\Phi(x)$ iff for every I^* that differs from I_M either not at all or in the assignment to x, $\langle D_M, I^* \rangle \vDash \Phi(x)$
$M \Dashv (x)\Phi(x)$ iff for some I^* that differs from I_M either not at all or in the assignment to x, $\langle D_M, I^* \rangle \Dashv \Phi(x)$

The truth-functional intuition as embodied in the pre-assignment serves as a minimal basis for truth-value assignments, but it appears to be inadequate in two respects. First, the penumbral intuition reveals another way we establish truth and falsity: if a sentence strikes us as logically or analytically true (logically or analytically false) we typically believe it (reject it). In addition, the truth-functional intuition as embodied in the strong Kleene scheme fails to reflect how constraints on increases in precision figure in our understanding of predicates; the pre-assignment would have no resources to distinguish 'tung' and 'heavy' if they had the same extensions and anti-extensions. The idea behind supervaluational semantics is to do these two jobs at once: identify truth (*simpliciter*) with truth on all acceptable ways of eliminating all vagueness by increasing precision.

A partial model M is an *elaboration* of another partial model M' if their domains are the same, their interpretation functions assign the same objects to the constants and variables, and, for each P, each component of $I_M(P)$ contains the corresponding component of $I_{M'}(P)$. (Sometimes, when there is no danger of confusion, elaborations will be called *extensions*.) The operation of making a predicate more precise corresponds to the action of replacing a partial model with an elaboration of it. There are two natural ways to represent constraints on increases in precision. One may take the notion of admissible elaboration as basic; then the pre-analytic sentences turn out to be sentences true on all ways of increasing precision with an admissible elaboration. Alternatively, one may take the pre-analytic sentences as basic: these may be represented by a set of sentences of the language. Although the former is more "semantic" in spirit, the latter will be more convenient. For now I shall require that the set of sentences true in the pre-assignment are consistent with the set of penumbral sentences. (This assumption will be examined and abandoned soon enough.) A (*constrained*) *supervaluational truth assignment* is given by an ordered pair $\langle M, S \rangle$ where M is a pre-assignment and S is a set of sentences consistent with M such that no member of S is assigned a value in M. An elaboration M' is *admissible* if no member of S is assigned false in M'. An admissible elaboration is *complete* if it is a classical model. Let K be the set of complete admissible elaborations relative to $\langle M, S \rangle$. I define satisfaction ($\vDash_{SV}$) and falsification ($\Dashv_{SV}$) for the constrained supervaluational structure. (Φ

a formula):

$\langle M, S\rangle \vDash_{SV}\Phi \leftrightarrow$ For all $N \in K[N \vDash \Phi]$
$\langle M, S\rangle \dashv_{SV}\Phi \leftrightarrow$ For all $N \in K[N \dashv \Phi]$
Otherwise neither $\vDash_{SV}\Phi$ nor $\dashv_{SV}\Phi$

This is a clean and conceptually simple assignment of truth values which satisfies desiderata (1–3). Penumbral sentences receive the value they appear to have—true—because they are true on all the complete elaborations that respect the constraints they represent. (In the next few paragraphs, 'A man with exactly n cents is poor' will be abbreviated $P(n)$.) The answer to the sorites problem is straightforward: 'For all n, If $P(n)$ then $P(n + 1)$' comes out false on all complete admissible elaborations. Consequently, it receives the SV-value false, and that is that. This approach has a cost in that one must abandon the truth-functional intuition. But for those who feel the tug of the penumbral intuition more than that of the truth-functional intuition, the supervaluational account looks like a wonderful story.

One catch is that opponents of supervaluations may reiterate the truth-functional intuition in a more compelling guise. Since the supervaluation deems the conditional premise of the sorites paradox false, it deems true the claim that there is an n such that $(P(n)\&\neg P(n + 1))$. Alas, there is no such number. The supervaluational assignment of values to existentially quantified sentences (and disjunctions) may turn upon what Fine calls "truth value shift": in each complete admissible extension there is a different n such that '$P(n)\&\neg P(n + 1)$' is true. Despite the ingenuity of the supervaluational story, truth-value shift does not fit very well with the way we normally understand disjunction and existential quantification. One might describe the objection as the "objection from upper-case letters," since it is hard to resist the temptation to type 'When I say that there is a number n such that . . . I mean that THERE IS a number n such that'.[6] It is not merely that our intuitions rebel at truth-value assignments that turn upon truth-value shift; as Saul Kripke has noted, the supervaluational assignment sits ill with the way we respond in practice to information conveyed by disjunctive and existentially quantified sentences. We are told that 'p or q' is true and we naturally ask, "Well, which one is it (if not both)?" So, too, if we were told that there was a number with an interesting property P we might think to ourselves, "Let's investigate further to find out which particular number it is."

[6] This sort of objection was put forward in D. Sanford, "Competing Semantics of Vagueness: Many Values vs. Super-Truth," *Synthèse*, XXXIII (1976): 195–210.

This objection exploits the fact that, however one feels about the truth-functional intuition generally, the pre-analytic sentences one is least inclined to call true are disjunctions and existential quantifications. (Some people would prefer to say that they have a strong inclination not to call these true.) Compare the laws of excluded middle and noncontradiction in a particular instance. It generally seems correct to say of a borderline bald man 'This man is not both bald and not bald', while our intuitions range from mixed to strongly negative on 'This man is bald or this man is not bald'. Given the logical interderivability of these two sentences, these differing intuitive responses are perplexing.

This tendency to differential treatment arises from a difference in the functions of the utterance of these sentences. The point may be brought out if we consider the sort of structural aspects of models that are (broadly speaking) responsible for the truth of such sentences, when they are true. Noncontradiction in these cases is a "no overlap" condition—it says that the set of objects of which P is true and the set of objects of which P fails to be true are to be disjoint. The same sort of disjointness feature of intended interpretations is what gives intuitive appeal to 'This is not both orange and red' said of a borderline orange-red thing: our understanding of orange and red as contraries leads us to expect that any proper interpretations of 'orange' and 'red' will have disjoint sets as the extensions of the predicates.

By contrast, excluded middle functions as a "sharp boundaries" condition. When I say 'This guy is either bald or he isn't', I convey either that there is some boundary that determines that he should be counted as one or the other, or that some boundary should be settled upon. When I say 'Everyone is either bald or not bald', I convey that there is a boundary sharp enough to count everyone one way or the other, or that some boundary should be established to do that, or that we should commit ourselves to stipulating away borderline cases of baldness by sharpening the predicate whenever borderline cases are presented to us.

An example helps to clarify this point. Say you have the job of sorting color samples on an assembly line. The samples come along the line in varying shades of red or orange. No other colors are sent rolling out. You are to drop the orange samples into one bin and the red ones into another. Every so often an indeterminate case comes along and you cannot make up your mind about it, so you set it aside. The foreman notices the growing pile of samples by your side and says, "Every one of these samples is either red or orange." Among the effects of this utterance is to set high standards for resolving indeterminacy: indeterminate cases are to be put into one bin or another, so that 'red' and 'orange' function as predicates with sharp boundaries because the contextual standards are such as to get you to make them behave as if they have sharp boundaries.

The "no overlap" condition always holds, although the "sharp boundaries" condition might not. So the structural condition that in some intuitive sense is responsible for the truth of noncontradiction is present in every context, while the condition that is intuitively responsible for the truth of excluded middle is not. Those who utter pre-analytic sentences of the "no overlap" type cannot be, in any straightforward way, induced to withdraw the utterance, while those who utter sentences of the "sharp boundaries" type can. If I utter a sentence of the "Nothing satisfies both this description and that one" type, the way one typically would get me to withdraw the utterance (the "canonical way" so to speak) would be to produce a counterexample fitting both descriptions. ('No one is both Danish and a chess grandmaster? What about Larsen?') To get someone to withdraw a sentence of the 'Everything is one of these or one of those' type we also tend to produce a counterexample—but in this case it will be an instance that fits into neither category. ('Every Congressman is either Democrat or Republican? What about Saunders?') This may amount to the production of a *hard case*. And though overlap never occurs in the cases considered here, hard cases abound.

The upshot is that instances of noncontradiction and similar sentences can be stable in the sense that once one has uttered an instance, one will not be in the position where one would be forced by a counterexample to withdraw it. But an instance of excluded middle is always at risk: having uttered it, one is committed, at least in principle, to stipulating sharper boundaries to resolve indeterminacies whenever they are produced or arise on their own. Such stipulations may seem pointless, given the standards of precision in a given context.

This explains the low-level pragmatic phenomenon of the differing intuitions directed toward noncontradiction and excluded middle. To utter and adhere to an instance of excluded middle in a context is to set contextual parameters that require a high standard of determinacy, and to withdraw or refuse to accept such an utterance is to weaken the standards of determinacy in that context. Sometimes the level of precision we take instances of excluded middle to require is onerous and inappropriate. Hence our reluctance to embrace instances of excluded middle wholeheartedly.

IV. PENUMBRAL VERSUS TRUTH-FUNCTIONAL INTUITIONS: ROUND II

It might seem as if the supervaluationalist could respond that the objection from upper-case letters is just a variant of the truth-functional intuition, and we have known from the start that supervaluations would face *that* conflict. But the objection from upper-case letters cuts more deeply. For some predicates it seems incoherent to specify any sharp boundaries at all; a proper understanding of the predicates appears to require that one not do so.[7] This is a difficult blow for the supervaluational approach to absorb because it suggests that the complete extensions that are used to define supervaluational truth may all violate constraints on increases in precision. The set K of complete admissible extensions might be empty. That is, the pre-analytic sentences need not be consistent with the set of sentences in the pre-assignment or with each other.

This seems true, for example, for predicates that are relativized to observers in such a way that the predicates inherit the observers' perceptual limitations. Some examples of this: 'x looks red to most people', 'y appears straight', and 'z seems to me to be the same size as w'. Also, the modifier 'roughly' seems to have precisely the effect of making predicates vague in such a way as to exclude too much precision. It seems contrary to our understanding of what 'roughly' means to imagine drawing a sharp boundary for a predicate like 'x is roughly a handful of sand'. In cases like this it is not merely lethargy that prevents the drawing of new, completely sharp boundaries; such boundaries appear to be ruled out by the very principles governing the use of the predicates.

Say that a pre-analytic sentence is a *local consistency rule* if it is of the form 'If Qab then (Pa iff Pb)'. Pre-analytic sentences of this type ensure that, if some condition holds, boundaries will not be drawn in such a way as to separate a and b by assigning different values to Pa and Pb. So, for example, 'If a and b are one millimeter apart then (a is roughly within walking distance of Cleveland iff b is roughly within walking distance of Cleveland)'. It is contrary to our understanding of the predicate 'x is roughly within walking distance of Cleveland' to draw boundaries that would distinguish between two positions one millimeter apart, so this sentence is pre-analytic. As an additional example, note that if a and b are color samples 'If a and b are observationally indistinguishable to me in respect of color, then a looks red to me if and only if b looks red to me' is a local consistency rule.

Call a predicate P *essentially vague* if there is a sequence a_1, a_2, . . . , a_n, and a relation Q, such that a_1 is a clear case of P, a_n is a

[7] This is argued by Kamp, in the course of the apostasy, "The Paradox of the Heap," in U. Monnich, ed., *Aspects of Philosophical Logic* (Dordrecht: Reidel, 1981), pp. 225–77. The point is implicit in Crispin Wright's "The Coherence of Vague Predicates," *Synthèse*, XXX (1975): 325–65.

clear countercase, for each $i < n + 1$, Qa_ia_{i+1} is true, and each instance of 'If Qa_ia_{i+1} then (Pa_i if and only if Pa_{i+1})' is a local consistency rule. For example, a sequence of positions one millimeter apart, beginning a stride beyond the city limits of Cleveland, and ending in Calgary, witnesses to the essential vagueness of 'is within walking distance of Cleveland', where Qxy is 'x and y are one millimeter apart'. No admissible elaboration can make an essentially vague predicate completely precise, because any sharp boundary would violate one of the local consistency rules.

In addition to the semantic side of the problem, there is a syntactic side. The penumbral sentences of English cannot all be true for the hardest and most inescapable of reasons. Some of them are, collectively, inconsistent with such uncontroversial truths as 'There is a sequence of positions each one millimeter from its predecessor, with first member within walking distance of Cleveland and last member not'. But if we cannot make a principled distinction between the penumbral sentences that cannot be true, and those which can, this calls into question the intuitions that lead us to want to call any penumbral sentence true.

It would be too hasty to abandon the supervaluational approach in response to these concerns. It is one thing to say that we should not take SV-values to be truth values (which is correct) and quite another to say that some supervaluational structure should not be part of our semantic theory. Note that the supervaluation is performing two different tasks. It represents the collection of admissible increases in precision of the language, thus regimenting the notion of constraint on increases in precision. In addition—this task is distinct from the first—it provides an account of truth for a language with vague predicates by identifying SV-values with truth values. The motivation for taking the supervaluational structure to perform the second job is the penumbral intuition, which is indeed undercut by the objection from essentially vague predicates. But the motivation for a truth-gap theory that incorporates a notion of increase in precision, constraints, etc., has not gone away. What the objections show is that the supervaluational structure needs modification rather than elimination.

The modifications required are straightforward. First, drop the assumption that the set of penumbral sentences is consistent with the sentences assigned true in the pre-assignment: allow the set of penumbral sentences to be an arbitrary set of those sentences in the language which lack a truth value in the pre-assignment. Now proceed as above, except that a sentence is called pre-analytic if it is never assigned false on any elaboration.

It is granted to the truth-functional intuition that the identification of truth values and SV-values must go. Call a sentence true (false) if it is true (false) in the pre-assignment. What emerges is a representation that contains the two components needed: truth values given by the pre-assignment and what might be called invariance values given by the family of admissible extensions. This much is conceded to the penumbral intuition: pre-analytic sentences do have an important distinguished status. They are never appropriately called false. But contra the penumbral intuition, they are not always correctly called true.

The class of essentially vague predicates brings into relief the difference between two types of rules: constraints on increases in precision and rules of application. Consider a rule of application that may be formulated: "If *a* and *b* are, to me, observationally indistinguishable then (*a* looks red to me iff *b* looks red to me)." If our use of 'looks red to me' is taken to be governed by such a rule, then the existence of sorites chains of color samples generates several puzzles. One is that the rule would seem to make each member of a set of local consistency rules analytic, when the sorites argument shows the set to be inconsistent with certain incontrovertible facts. We might also think that the rules of use of 'looks red to me' are in some way incoherent, though it is not clear what this could mean.

In contrast, a constraint on increases in precision is a weaker condition. The puzzles just mentioned are not generated if we assume that *P* is governed by, among other conditions, a rule like "If *a* and *b* are indistinguishable to me in respect of color, never stipulate boundaries in such a way that '*a* looks red to me' counts as true and '*b* looks red to me' counts as false or vice versa." No contradiction is generated by the rule: never stipulate boundaries to 'looks red to me' in such a way that 'If *a* and *b* are indistinguishable to me in respect of color and *a* looks red to me then *b* looks red to me' comes out false. The constraint leaves open the option of suspending judgment on '*a* looks red to me' while calling '*b* looks red to me' true. The constraint requires that, if certain stipulations are to be made, they cannot be made in the proscribed fashion. It does not prescribe that any stipulations be made or even ensure that if any stipulations were made they would be acceptable.

V. PRE-ANALYTIC SENTENCES: A USER'S GUIDE

To explain what occurs in cases where pre-analytic sentences are uttered, we should turn attention to the apparently uninteresting, but crucial activities like correcting the linguistic mistakes of other people, teaching proper use, and the like. These are the oil changes and wheel realignments of a smoothly functioning language—the "routine maintenance" of the conventions of use. Coherent speech in a population is too complex a phenomenon not to degenerate into a confusion of tongues unless it is constantly serviced in these ways.

One way we maintain the stability of the conventions of English is by uttering analytic sentences. If someone misused the expression 'skinflint' (assume for the example that this expression is precise) one might well say 'Skinflints are tightwads'. The sentence uttered is true, but we are not asserting it in order to communicate the facts about skinflints to the listener. The speaker does not have a full grasp on the proper use of the words 'skinflint' and 'tightwad', and we want to set him straight. In a case like this, our goal is not to say something true, although we may accomplish the goal by saying something true, and we typically take ourselves to be saying something true. What we are principally attempting to do is to correct mistaken usage. The general pattern of this activity is: we utter a declarative sentence S in order to induce the withdrawal of a mistaken utterance of $\neg S$, or the withdrawal of other utterances that can only be true if $\neg S$ is true, or to ward off anticipated mistaken utterances of $\neg S$, by indicating that $\neg S$ is never correctly assertable.

What sort of speech act is performed when we correct mistakes or ward off anticipated ones in this way? The dissection of this activity is delicate: correcting mistakes does not fit comfortably into accepted categories. To help make clear why this is so, it is useful to first consider assertion. I shall look at assertion with reference to two features: the way in which truth values of sentences uttered are related to the conditions for correctness (broadly understood) and the intended effect on the hearer (i.e., the perlocutionary effect.)

It will here be held that a condition of correctness for a literal (nonironic, nonmetaphorical) assertion is that the sentence uttered must be true. That the truth of a sentence uttered is in this way internal to assertion is suggested by Michael Dummett:[8] "The connection between truth and assertion, . . . resides in the fact that the assertion of a sentence is its assertion *as true*, . . ." (*ibid.*, p. 460). For present purposes, though, nothing is served by disputes over the application of the term of art 'assertion'. The upcoming remarks isolate two features of assertion that may be taken to be basic to our understanding of what an assertion is. Those readers who balk may instead take them to isolate a subclass of assertions that are different

[8] *Frege: Philosophy of Language*, 1st ed. (London: Duckworth, 1973). This position represents something of a divide among those who discuss assertion. A different line of investigation holds assertions to be correct if some standards of justification are met—the truth of the sentence uttered is incidental to whether or not an assertion is correctly performed. A representative is Ernest Adams's treatment of indicative conditionals in *The Logic of Conditionals* (Dordrecht: Reidel, 1975).

from ordinary ones, in ways that are of crucial significance in this context.

The Dummett remark hints at an additional feature that rarely needs to be explicitly noted. The sentence is uttered for a *purpose*—communicating the information the sentence encodes—that requires the truth of the sentence in order that the purpose be attained correctly. For example, the purpose of the assertion of a sentence S may be to convince the hearer of what S states, or induce him to believe what S states. For the speaker to be correct in convincing the hearer of what S states, S must be true. But note that there is a distinction to be drawn between sincerely and literally uttering a sentence S to induce a hearer to believe what S states and to induce a hearer to withdraw an assertion of $\neg S$. In the first instance, it is a condition of correctness that S be true, while in the second, S must not be false.

Of course, there is rarely just one goal of uttering a sentence. The utterance of a declarative sentence can serve to convey the truth of a sentence and also serve some other purpose. 'You are out of uniform, soldier' can, under the appropriate circumstances both "state the facts" and issue an order. The utterance of 'Skinflints are tightwads' in settings like those above can achieve two distinct ends: "stating the facts" and correcting faulty use. For the former task, it must be that the sentence is true, and for the latter, it need only be that the sentence is not correctly called false. (Recall that we are taking S to be false exactly if $\neg S$ is true.) It is a happy coincidence that the behavior by which these two goals are attained is the same. 'Skinflints are tightwads' happens to be true, and most of the sentences of its type at least appear to be true. So we are never forced to recognize the double duty an utterance of the sentence may do, or to face the need for refinement of the notion of assertion.

To appreciate the intricacies involved in the utterance of pre-analytic sentences, however, such refinement is needed. In such cases, S may not be correctly called true, even though $\neg S$ is not correctly called true either. Care is needed to ascertain what aspects of the truth status of sentences are relevant to the intended purpose of the utterance, particularly for speech acts that are aimed at correcting attempts to stipulate inadmissible boundaries. To correct such faulty use properly, it is only necessary that the sentence uttered not be false. Unlike the skinflints/tightwads case, one does not correct a mistake by performing an act that is just like assertion but for a difference in the intended perlocutionary effect. There may also be a difference in the truth status necessary for the goal of the utterance to be correctly attained.

A sentence will be said to be *articulated* if it is uttered for the purpose of inducing someone else to withdraw, or refrain from asserting, some sentence. An example will clarify the distinction between articulation and assertion. Say you hear the following (said of two borderline tall men): "Joe is tall and John isn't." You know that John is taller than Joe, so you respond: "If Joe is tall then John is." You may not want to commit yourself to drawing boundaries sharp enough so that 'If Joe is tall then John is tall' would come out true. The point of uttering 'If Joe is tall then John is' is not to state the facts to someone who has not committed himself to an opinion, but rather to induce the withdrawal of the first sentence, the first speaker's assertion of which is a mistake. So in this example, 'If Joe is tall then John is' is articulated. The articulation of the sentence is correct if in fact the first utterance *was* a mistake. That is, the articulation of 'If Joe is tall then John is' is correct if it is incorrect to regard 'Joe is tall and John isn't' as true. The fact about the truth status of 'If Joe is tall then John is' which is relevant to the articulation being correct is that in this context the sentence is not correctly called false.

When inadmissible boundaries are proposed, a neglect of the difference between articulation and assertion leads us to speak in a way that is systematically misleading. We behave the same way when inadmissible boundaries are proposed as we do when the negation of an analytic sentence is asserted. If we hear someone saying $\neg S$, where $\neg S$ can be true only if inadmissible boundaries are drawn, we could say something like; "You can't correctly say $\neg S$," which would make it clear that commitment to S need not be correct either. But in fact we tend just to say S or something equivalent to it, which leads one to understand S to be endorsed as true.

The distinction between articulation and assertion takes on added significance in the case of essentially vague predicates. Here the sentences uttered may not be correctly called true in any reasonable context, so if we regard the utterance of the sentence as an attempt to make a correct assertion, we must conclude that it is a failed attempt. But it may well induce the withdrawal of the original utterance; if so, in this regard it is a success. If the speech act is judged by the ends we have in performing it, it is successful, but understood as an attempt to make a correct assertion, it cannot be successful. If someone consistently demonstrated a misunderstanding of the proper use of 'looks red to me' or 'observationally indistinguishable', one might well say to him: "Look, if any two samples are observationally indistinguishable to you then one looks red to you if and only if if the other looks red to you." In conjunction with some indisputable facts, this sentence yields a contradiction; the sentence is not correctly assessed as true. But the utterance will probably have the desired effect of getting the hearer to stop talking in the proscribed way.

Articulation differs in its underlying mechanisms from other cases —such as irony and metaphor—where one may succeed by sincerely uttering a sentence that cannot correctly be called true. Communication with metaphor, if effected by uttering sentences that are false, typically works only when the sentences uttered are recognized to be false. So, for example, on H. P. Grice's influential account it is typically crucial that, if the sentence uttered is false, it is recognized by the hearer to be false. The recognition of the falsity of a sentence uttered may cue an implicature—an alternative interpretation of what is to be communicated—in more or less rule-governed ways. In contrast, not only does it fail to be necessary that we recognize that the sentences we articulate should not be called true: we usually (mistakenly) judge that the truth-valueless pre-analytic sentences we utter and hear are true. This is what makes supervaluational semantics so seductive. But this judgment to the effect that S is true is not necessary in order to prevent the utterance of $\neg S$ by uttering S. Only the correct recognition that S cannot correctly be called false is needed to underwrite the message that $\neg S$ is not correctly assertible.

This helps explain why the penumbral intuition is so tenacious. We use pre-analytic sentences in a way that it seems could only be successful if these sentences were true. The natural confusion of asserting with articulating never gets one into trouble in one's everyday practice. We use pre-analytic sentences in ways that are clearly acceptable, and which apparently can only be appropriate if they are true, so we naturally come to think that these sentences must be true. But note that the semantic fact about the sentence which allows it to be properly used to correct mistakes is not that it is true but rather that it is not false.

What we learn from the existence of essentially vague predicates is that sometimes, when it is correct to say 'You can't correctly assert S', it is also correct to say 'You can't correctly assert $\neg S$'. Usually, if it is incorrect to assert S it is acceptable to assert $\neg S$ so we have fallen into a bad linguistic habit: rather than say something like 'You can't correctly assert S' we tend just to say '$\neg S$'. We must not interpret such an utterance of '$\neg S$', intended to perform the perlocutionary act of correcting a mistaken utterance of S, as committing the speaker to believe $\neg S$. For example, in articulating a sentence like 'If any two samples are observationally indistinguishable to you then one looks red to you if and only if if the other looks red to you', the speaker may be attempting to induce the hearer to realize that some other sentence should not be asserted, but this should not be taken to commit the speaker to asserting the sentence. In a more pressing formulation, say that the sentence S the speaker utters is 'It is not the case both that a looks red to you and b does not look red to you', where the hearer has just said that a and b are observationally indistinguishable to him. In saying S, one is articulating the sentence, but one had better not be asserting it. One does not want to be committed to the equivalent sentence 'a doesn't look red to you or b does'.

With this preparation completed, the account of the sorites implicit in the modified supervaluations may be made explicit: the inference is not correct because the second (conditional) premise lacks a truth value, though that premise typically has a distinguished status that leads to its being used as if it were true. This is clearest if the predicate that is taken to support the sorites inference is not essentially vague, so that it is acceptable to specify sharper boundaries to the predicate in such a way that the conditional premise comes out false. Consider, for example, the 'poor man' variant. As noted, it is acceptable, if the context warrants, to draw sharp boundaries for the predicate 'is poor'. When such sharpenings are effected, 'For all n, if a man with n cents is poor then a man with $n + 1$ cents is poor' comes out false. Given the principle that truth is settled truth, the second premise must be false or lack a truth value. Since the predicate 'is poor' in its ordinary use admits borderline cases, the second premise is not false, and so it must lack truth value. But we occasionally use sentences like the second premise to convey "no sharp boundaries" conditions, as detailed earlier; we come to regard the second premise as true because in most contexts we do not want to draw boundaries sharply enough for the sentence to count as false.

Essentially vague predicates yield a context-invariant version. If someone were to assert something that has as a consequence the negation of a local consistency rule associated with an essentially vague predicate, we could take it as indicating a failure to understand that predicate. If someone says 'Oh, yes, a looks red to me and b looks to me to be the same color as a, but b doesn't look red to me', we take this to betray a failure to understand either the 'looks red' idiom or the 'looks to me to be the same color as' idiom. We might respond by articulating the appropriate local consistency rule. But this pattern of use should not be taken to commit us to holding the sentences we articulate to be true.

VI. THE LIAR REVISITED

Consider the sentence 'k is not true iff k is true' in the circumstances imagined in the introduction. The sentence is the right one to utter, given what one is trying to accomplish by giving voice to it. It seems that this pattern of use serves the function that the articulation of a local consistency rule does: to bring the hearer to realize that certain sentences—liar sentences—cannot correctly be asserted. The purpose of this section is to draw out this parallel, with reference to a particular semantic account of the truth predicate—the Kripke theory of the minimum strong Kleene fixed point—and a widely voiced style of objection to it.[9] All we need to know of this account is that it is a theory that treats 'x is true' (written '$T(x)$') as a partially defined predicate of sentences with the strong Kleene evaluation scheme for complex sentences, that liar sentences lack a truth value, and that the truth predicate obeys the rules: T('S') $\vdash S$ and conversely, and $\neg S \vdash \neg T$('S') and conversely.

A common objection to the Kripke account is that many intuitively true sentences are assigned no value. Anil Gupta[10] puts the objection this way:

> By the [Kripke] definition various logical laws are sometimes paradoxical. For example, [the Kripke] definition entails that the law ["$(x)\neg(Tx\&\neg Tx)$"] is paradoxical when there is a liar-type sentence in the language—for now the law does not have a truth value in any of the Kripkean fixed points . . . Intuitively the law does not seem to be paradoxical. In fact even in the presence of paradoxical sentences, far from finding the law paradoxical, we are inclined to believe it (*ibid.*, p. 209).

Gupta is right that we are inclined to believe this sentence. But vague predicates and the truth predicate are sufficiently similar in relevant respects to lead us to be wary: in the case of vague predicates that it may be a mistake to yield to this inclination and take the sentence to be true. The possibility of augmentations of the extension of the truth predicate is implicit in our attitudes to sentences containing the truth predicate. Consider a truth-teller sentence like (17): '(17) is true'. One way in which we are inclined to distinguish the pathology (17) presents from that of a liar sentence is to note that we may resolve (17) either way. We could count it as true if we wanted to, or as false. In contrast, k cannot rightly be called true, nor can it rightly be called false. This suggests that some of the considerations that pertained to vague predicates may be carried over. We should expect to find identifiable constraints on such stipulations and indeed $\neg(T(17))\&\ \neg T(17)))$ appears to be one.

[9] Cf. Kripke, "Outline of a Theory of Truth," this JOURNAL, LXXII, 19 (November 6, 1975): 690–716. It should be noted that only a very special case of the general Kripke account is considered here.

[10] "Truth and Paradox," in R. Martin, ed., *Recent Essays on Truth and the Liar Paradox* (New York: Oxford, 1984), pp. 175–237.

It is illuminating to consider an inessentially vague predicate like 'is poor', and see how the truth predicate behaves. Even if the language is restricted so that liar-type paradoxes do not occur the sentence '$(x)\neg(Tx \& \neg Tx)$' should be taken to lack a truth value, if a partial predicate P is in the language. Say that Pa is indeterminate. Since it cannot be correct to assert 'Φ is true' if it is not correct to evaluate Φ as true, it is not correct to evaluate 'Pa is true' as true. Similarly, since $\neg Pa$ is not correctly called true, neither is $\neg(T(\text{'}Pa\text{'}))$. If neither $\neg T(\text{'}Pa\text{'})$ nor $T(\text{'}Pa\text{'})$ may correctly be evaluated as true, the strong Kleene rules give no value to $\neg(T(\text{'}Pa\text{'}) \& \neg T(\text{'}Pa\text{'})$ so $(x)\neg(Tx \& \neg Tx)$ lacks a truth value.

Matters are similar with $(x)\neg(Tx \& \neg Tx)$ when the language contains no partial predicates but the truth predicate, and the quantifiers range over liar sentences. The Gupta objection brings out a conflict between the penumbral and truth-functional intuitions reminiscent of vague predicates. With garden-variety vague predicates we discovered one powerful reason for resisting the penumbral intuition: in some cases, the penumbral intuition was directed toward sentences that were inconsistent with incontrovertible facts. This is even more stark with the truth predicate. The intuition we have when considering a local consistency rule for '*x* looks red to me' is of a piece with the intuition we have when considering Tarski biconditionals with liar instances. Both are sentences that appear, on the basis of our intuitions about meaning, as if they must be true, yet considerations of logical consistency ensure they cannot be.

These considerations undercut the inclination to regard $(x)\neg(Tx \& \neg Tx)$ as true. Perhaps the penumbral intuition—transplanted into this new setting—should be overruled by the truth-functional intuition. As embodied in the strong Kleene definition, the truth-functional intuition tells us that $(x)\neg(Tx \& \neg Tx)$ should lack a truth value. So in light of all this reconsider the remark, ". . . even in the presence of paradoxical sentences, far from finding the law paradoxical, we are inclined to believe it." Our response can mirror our attitude to the penumbral intuition for sentences with vague constituents: Yes, we want to believe it, but we may be making a mistake by giving in to the temptation.

As with vague predicates, we may explain away the penumbral intuition by noting the way the patterns of use of the relevant sentences lead us to take them unreflectively to be true. The Tarski biconditional with liar instances—'*k* is true if and only if *k* is not true'—might well be uttered in the course of an attempt to demonstrate why it is unacceptable to assert *k* and unacceptable to assert the negation of *k*. The imagined utterance of the Tarski biconditional is successful if the hearer recognizes that certain other sentences cannot be correctly asserted; so the account of articulation in v extends naturally to this expanded setting.

It need not be that every time the sentence $(x)\neg(Tx\&\neg Tx)$ has been pronounced, it has been, strictly speaking, a mistake. Often our reasoning is implicitly restricted by the context so when we utter $(x)\neg(Tx\&\neg Tx)$, the quantifier may be implicitly restricted to range over sentences with no semantic anomalies. With the quantifiers so restricted, the sentence $(x)\neg(Tx\&\neg Tx)$ is true and correctly assertable, just as it can be true to say at a faculty meeting 'Everyone is here'. The point is that one need not posit that such quantifier restriction *always* occurs; *sometimes* the context does not restrict quantifiers to sentences featuring no anomalies. In cases where this contextual restriction is absent, the sentence may be articulated. It was noted that sentences of the form $(x)(Rx\&\neg Rx)$ tend to function as "no overlap" conditions that indicate that a given context must not count it as false. There is a tendency to count such sentences true. The same point helps to explain the appeal of $(x)\neg(Tx\&\neg Tx)$: it can serve as a "no overlap" constraint as the analogous sentence would serve with 'looks red to me' in the place of 'T'. As with vague predicates, the intuition that this sentence is true may be explained.

This indicates what our attitude should be toward certain sentences which contain k as a constituent or which have quantifiers ranging over k. This tells a story about k indirectly, in that it contributes to a defense of the Kripke theory. But it also creates conceptual space for a positive account: here a gesture will be made toward such an account, with a more detailed treatment left for elsewhere.[11] What should our attitude be to the sentence 'k is not true'? Quite simply, we should neither assert it nor assert its negation. Rather, we should deny 'k is not true' and we should deny 'k is true'. The liar holds fewer terrors for us if we no longer take the denial of k to commit us to the assertion of 'It is not the case that k is not true' (which lands us in trouble). But such a notion of denial is waiting to be read off of the account of articulation in v): the articulation of S is a speech act that rejects $\neg S$ while allowing the option of rejecting S as well.

JAMIE TAPPENDEN

University of Pittsburgh

[11] Cf. my "Negation and Denial in Philosophical Logic," forthcoming in *Philosophical Topics.* On this point I am especially indebted to Soames, "Semantic Theories of Truth," unpublished manuscript (1984); and T. Parsons, "Assertion, Denial, and the Liar Paradox," *Journal of Philosophical Logic,* XIII (1984): 137–52.