TRUER WORDS

Draft of February 27, 2002

Preface to this Draft

Right now this manuscript lacks a little focus, because it is trying to serve two roles at once. First, it is meant to be the reading notes for a seminar on vagueness I am running in 'spring' semester 2002 at Brown University. Secondly, it is meant to be a draft of a book on vagueness. The two purposes are somewhat at odds. And I, foolishly, tried to draft something that would fit these purposes under ridiculous time pressure. And it's a first draft, at least of the book. (Given the time constraints, it will necessarily be the last draft of the teaching notes.)

There are a few sections that don't seem ideally suitable for a book. The little primer on intuitionist logic at the end of chapter 1 is the most obvious - I assume most readers will either be familiar with this or can find out about it somewhere else, but it does seem useful for students. And there are a few sections that are pretty irrelevant to a course. The appendix to chapter five where I tidy up some loose ends Field left in his recent paper is the most obvious, but I suppose more will become apparent as we go through it. Of course, these sections that are not suitable for both purposes are merely a minor annoyance. The real problem with the two goals is not in the choice of topics, but rather in the way they are tackled.

Having said all that, the manuscript does at least have a broad structure, and this structure even seems suitable for both roles. It divides into four parts. The first part (Chapter 1) is a very quick survey of the issues that arise in producing a theory of vagueness. The second part (Chapters 2 to 7) is a more detailed survey of the various approaches that have been tried as means of solving these problems. None of these work, mostly for reasons that are by now familiar. The third part (Chapters 8 and 9) sets out my solution. And the last part (Chapters 10 and 11) looks at the issues that arise when we try and account for languages with slightly more complexity than the proposition calculus. (By that, I mean languages with as much complexity as the predicate calculus, though maybe in later drafts we will look at even more complex issues, such as how modal operators, and conditionals with modal force, can be drawn into the loop.)

Naturally enough, the areas that provide the most interest for cognoscenti will be chapters 8 and 9. Most of parts 1 and 2 will be fairly familiar. There are some prominent exceptions to this. As with the negative parts of most books, most of these chapters have one or two, relatively small, *new* points to make, in the midst of my distillations of the strongest arguments against the view in question. Relatively small, that is, relative to the bombshells in chapter 8 and 9! The big exception is the argument against Field in chapter 5. This generalises to a sound argument against all approaches for representing uncertainty numerically but not probabilistically. The search for such an argument constitutes the single biggest issue in the philosophy of probability of the last century, so my completion of the quest here is of some importance. (All of this is conditional, of course, on my argument working. Given the percentage of philosophy arguments that do work, I would understand dear reader if you are somewhat sceptical of my inflated claims here. But that doesn't mean I am about to retract them!) The discussion of Schiffer's argument against supervaluationism in chapter 11 also leads to some interesting new points in ways you might not have expected. But the main thing I want to note here is that if the first four chapters feel fairly slow going, rest assured that the story will presently get much more interesting.

I want to flag one problem with the text that will be corrected in subsequent drafts. My official view is that both *truer* and, derivatively, *determinately*, come in both predicate and operator variants, just like *true*. So there is a relational predicate *is truer than*, that takes two *names of sentences* and produces a sentence out of them, just like any other relational predicate such as *is smarter than*. And there is an operator, helpfully also spelled *is truer than*, that takes two sentences (not names or descriptions thereof) and produces another sentence. This is quite unlike *is smarter than*, and much more like *and*. Anyway, at various parts of the text, this isn't made altogether clear. There are several points where the text needs improvement (this is a first draft after all), but this feels like the most prominent. I don't think there are many points where I get this wrong, but there are several points where I stumble around in confusion because I haven't made the position clear. The next draft won't have this paragraph, and will have at least fewer places where I stumble around blindly.

1. Introductions

This is a book on the philosophical issues surrounding vagueness. Vagueness is an interesting topic for a fairly superficial reason, and a fairly deep reason. The superficial reason is that vagueness generates certain paradoxes, and philosophy made it its business a whiles ago to solve these, so we should do our bit in the project. The deep reason is that getting clear on the issues surrounding vagueness *might*, depending on how they are resolved, have many implications for philosophy of language, for philosophical logic, and even for metaphysics and epistemology.

Why stop there? A recent book, Timothy Endicott's *Vagueness in Law* discusses the implications of vagueness in jurisprudence, and draws several dramatic conclusions concerning the possibility of judicial objectivity. And another book, John Coates's *The Claims of Common Sense*, argues that a particular theory of vagueness was doing a lot of the work in Keynes's *General Theory*. Since *The General Theory* is the most influential academic book since, I guess, Darwin's *Origin of Species*, this might make vagueness a very important topic indeed. Well, Coates is all wrong about the interpretative issues, so this perhaps isn't the best example, but it indicates just how seriously vagueness can be taken if you have the right, or Coates's case the wrong, frame of mind. We will come back to the role of vagueness in the *General Theory* from time to time, because I think it is a good example of how people reason about vagueness when not corrupted by too much philosophical theorising.

It might be a mild disappointment, but I'm not particularly convinced by any of these grand claims for the importance of vagueness. I have a reason of moderate depth for being interested in vagueness: getting clear on the right way to think about vague terms prevents us from slipping into some of these bits of grand speculation. I think puzzles concerning vagueness are more like Fermat's Last Theorem than the Continuum Hypothesis. It is nice to know how they turn out, and we should be moderately grateful to those who show us how they do turn out, but they don't have a huge number of implications for other fields. Hopefully the analogy holds in another dimension, and puzzles concerning vagueness have actual solutions, ones that can be and will be found. (If there were really no solutions to the vagueness paradoxes, like the Continuum hypothesis, that would be a *very* deep result, one that we'd have to spend even more time getting clear on than we should spend on vagueness as it is.)

1.1. What is Vagueness

Paul Grice says that "[t]o say that an expression is vague (in a broad sense of vague) is presumably, roughly speaking, to say that there are cases (actual or possible) in which one does not know whether to apply the expression or to withhold it, and one's not knowing is not just due to ignorance of the facts." This will do as a rough first approximation, but it needs clarification in a few ways.

Most of the literature on vagueness concentrates on vague predicates, and indeed Grice's definition seems to be most immediately applicable to this case. So it is vague, for example, whether 'tall' applies to an adult American male who is 179cm tall is tall, or whether 'vehicle' applies to skateboards. But singular terms, even directly referential singular terms, seem to be vague. We may not know just which object is being referred to by *Mount Kosciuszko*, or *Saul Kripke*, given some plausible metaphysical assumptions about what kinds of objects there are. And, naturally, relational predicates also suffer from vagueness. Some might think it is vague whether Churchill was a better leader than Roosevelt, or just which cities count as being between New York and Boston. So the ignorance characteristic of vagueness might not just be about whether a particular objec t satisfies a particular predicate, but whether that object is referred to by a particular term, or whether an *n*-tuple of objects satisfies a particular *n*-place predicate.

Just how there could be this kind of ignorance is puzzling enough. Presumably Grice's 'one' here is meant to refer to an arbitrary person, so these are meant to be cases in which no one could know whether the term applied or did not apply. Now that there might be facts which not only do no humans know, but that in fact no humans could know, is not surprising. Even what we can perceive of the universe reveals that it is too big for us to know more than a fraction of its secrets, and if there's more to the world than the perceivable universe, then there are even more things we cannot know. Intuitively, though, linguistic facts like whether a particular predicate is satisfied by objects of a certain kind, should not be the kinds of facts of which we are ignorant. Linguistic facts are not part of the world we (collectively) discover, they are part of the world we (collectively) create. How these facts could be hidden from their creators is a minor mystery. This is not, however, the main mystery concerning vagueness, though it does become particularly pressing on one prominent account of vagueness. Rather, vagueness gets its philosophical importance from challenges that arise when considering the following puzzles.

1.2. The Sorites Paradox

We're going to start with four puzzles, many of which will be familiar to most of you. The first of these is the Sorites paradox. Sorites puzzles come in many varieties, but it'll be helpful to start with a specific instance, and then generalise. So consider argument (1)

(1) P1. Light of wavelength 700nm is red
P2. If light of wavelength 700nm is red, then light of wavelength 699.9nm is red
P3. If light of wavelength 699.9nm is red, then light of wavelength 699.8nm is red
…
P3001. If light of wavelength 400.1nm is red, then light of wavelength 400nm is red
C Light of wavelength 400nm is red

The first premise is true: 700nm is hardly perfect red, but it is red, and that is enough for our purposes. If you disagree, pick a wavelength that you agree is red. (If you think that light only has a colour relative to background conditions, perhaps because you're impressed by the data on colour constancy, make all the claims here relative to a suitable background.) The other premises, the so-called Sorites premises all seem pretty plausible. We can't detect, it seems, colour differences in lights whose wavelengths differ by 0.1nm. So it can hardly be the case that one of these lights is red and the other is not. Hence all the premises are true. And it seems the argument form, iterated modus ponens, is truth-preserving. Even those who quibble about modus ponens in general would accept it here. But the conclusion is clearly false. Light of wavelength 400nm is not at all red, it is blue. So we have an apparently truth-preserving argument form, with apparently true premises, and a false conclusion. That seems to count as a paradox.

There are several related arguments to this one. (1q) uses a single quantified premise rather than a string of conditionals. (2) uses negated conjunctions rather than conditionals, and (3) uses disjunctions rather than either conditionals or negated conjunctions. There are, of course, versions of each of these arguments that use a single quantified premise rather than the string of premises, but I won't write them out. And (4), which I guess is my favourite Sorites argument, appeals directly to a theorem of classical mathematics to do all the dirty work.

(1q) P1. Light of wavelength 700nm is red
P2. For all *n*, if light of wavelength *n* nm is red, then light of wavelength *n*-0.1 nm is red
C Light of wavelength 400nm is red

(2) P1. Light of wavelength 700nm is red
P2. It is not the case that light of wavelength 700nm is red, and light of wavelength 699.9nm is not.
P3. It is not the case that light of wavelength 699.9nm is red, and light of wavelength 699.8nm is not.
…
P3001. It is not the case that light of wavelength 400.1nm is red, and light of wavelength 400nm is not.
C Light of wavelength 400nm is red

(3) P1. Light of wavelength 700nm is red
P2. Either light of wavelength 700nm is not red, or light of wavelength 699.9nm is red
P3. Either light of wavelength 699.9nm is not red, or light of wavelength 699.8nm is red
…
P3001. Either light of wavelength 400.1nm is red, or light of wavelength 400nm is red
C Light of wavelength 400nm is red

(4) P1. Light of wavelength 700nm is red
P2. If light of wavelength 700nm is red, and light of wavelength 400nm is not, then there is an *n* such that $400 \leq n \leq 700$, and light of wavelength *n* nm is not red, and light of wavelength (*n*+0.1) nm is red.
P3. There is no *n* such that $400 \leq n \leq 700$, and light of wavelength *n* nm is not red, and light of wavelength (*n*+0.1) nm is red.
C. Light of wavelength 400nm is red.

In (4), P2 is a consequence of the least number theorem, applied now to the colour of light of various wavelengths, and P3 is a quantified Sorites premise. The argument here goes by modus tollens, which is less persuasive than modus ponens, but still pretty reliable in the circumstances. It is worthwhile reflecting for a little bit on how persuasive you find each of these arguments, and which of them you find persuasive because of their connection to other arguments. In particular, think for all three of the non-quantified versions whether the premises hold up when we get to cases around 650nm, or wherever it starts getting blurry that we've left the red zone, and are into orange. I have my feelings about the relative persuasiveness of the arguments, but I'll not reveal them just yet.

You might also notice that I appealed to the premises in (2) to justify the premises in (1). This will become relevant to some of the theories we will discuss later. I might be mistaken here, but I doubt one could give an *argument* for the soundness of (1) that did not appeal, at a level pretty close to the surface, to something like the premises in (2). This is important because it goes directly to what would count as a solution to the Sorites paradoxes.

Since the conclusions of the arguments are unacceptable, we have to isolate what has gone wrong. Well, there's essentially three options. Either the Sorites premises are not (entirely) true, or the non-Sorites premises are not (entirely) true, or the forms of argument involved do not preserve (entire) truth. Now just picking one of these options does not amount to a solution to the problem. We have to also explain why we found the arguments attractive in the first place. Just saying, for instance, "Oh, there is an *n* such that light of wavelength *n* nm is not red, and light of wavelength (*n*+0.1) nm is red," isn't particularly helpful, we have to explain why we thought that in the first place.

Such explanations are, it seems to me, subject to two constraints. First, they must explain all the data. If I have a theory that explains why (1) seems like a good argument, but does nothing to explain why (2) seems like a good argument, then I can hardly have claimed to have solved the Sorites paradox. This is why getting clear on which of the arguments provide foundations for other arguments is important. If in all cases, the reason we accept the Sorites premises is because we accept the premises in (2), and we implicitly accept that the premises in (2) entail the premises in the other arguments, and we can explain why we accept the premises in (2), and this apparent inference from those premises to the other ones, we're done! Secondly, the explanation must not over-generate. If our intuitive reaction to one of these Sorites arguments is to deny that all its premises are (entirely) true, our explanation for the apparent acceptability of the other arguments had better not also 'explain' our
that argument. This seems like an important point to me, since roughly none of the theories on the market is particularly proficient at drawing any distinctions between these forms of argument, and they do strike me as being quite different in persuasive force.

As well as Sorites arguments, there are also 'forced march' versions of the Sorites. In these, you are meant to imagine facing an interlocutor, who, to make it specific, shows you light of various wavelengths, and asks you "Red?" Perhaps she also tells you what its wavelength is, just to remind you how small the margins are in this case. And she lets you know before you start she'll be quizzing you on your answers, asking you to justify your behaviour if it is too erratic. So you can see what comes next. The dialogue will go something like this, with A as the questioner and B as you:

A: (Shows light of 700nm) That's 700nm. Red?
B: Yes
A: (Shows light of 699.9nm) That's 699.9nm. Red?
B: Yes
…

Now as the dialogue continues, one of two things will happen. Either (a) you will stop saying 'Yes', or (b) you will say 'Yes' to a question where the answer is obviously 'No'. If (a), then there will be a last question to which you answer 'Yes', and the quizzer will be entitled to be somewhat bemused as to why you gave different answers to questions concerning apparently indiscriminable lights. Note that I didn't assume here that there would be adjacent questions where you will say 'Yes' to one and 'No' to the other. I just assumed that you would stop saying 'Yes', perhaps by refusing to answer, or perhaps by saying 'It seems like it is', or something, at some time. This point will be worth bearing in mind as we continue through the discussion. Because of this concern about shrugs and qualifiers, it isn't immediately obvious that the forced march Sorites and its attendant puzzles raise exactly the same issues as the Sorites arguments, and their attendant puzzles. I hope the issues they raise are similar enough that there can be a common solution to them, but for now we just note that these *might* be separate problems.

1.3. Logical Anomalies

Assume that it is vague whether Louis is bald. That is, assume that Louis is one of those cases that Grice discussed, where it is impossible to tell whether he is bald, and this is not just due to ignorance about the status (or otherwise) of Louis's hair. Which of the following sentences strike you as being (uncontroversially) true?

(5) Louis is bald.
(6) Louis is not bald.
(7) Louis is bald or Louis is not bald.
(8) Louis is bald or not bald.
(9) If Louis is bald then Louis is bald.
(10) Louis is bald if bald.

I'm not so sure that (10) is a well-formed English sentence, though it seems to me that it ought to be if it isn't. If classical logic applies to natural languages, then all of (7) through (10) are true. And if the semantics routinely associated with classical logic, Tarskian truth theory, is appropriate for natural languages, then one of (5) and (6) must be true too.

Still, many people have had the intuition that several of these sentences are not true. Frege, for instance, thought that none of them were true because no sentence containing a vague term can be true. And we will be looking presently at semantic models on which none of these are true, though some sentences containing vague terms are true. Even a conservative like Russell thought that (7) was not true, and he might have dissented from the others. (I know Russell wasn't a conservative on many points, but in logic he clearly was.)

I will eventually argue that (7) through (10) are all true, and that in a sense, one of (5) and (6) has to be true as well. But I do not think this is the most intuitive position. I think intuitively we think that none of (5) to (8) are true, though we do think that (9) and (10) are true. Admittedly not everyone has these intuitions, but I think they are the most common reactions to the sentences.

If one of (7) through (10) is not true, then something is wrong with classical logic, and we have to start shopping for another logic. This is, to put it mildly, not a decision that should be taken lightly. Various luminaries, Hartry Field, Vann McGee and Timothy Williamson, have made some pointed and well-directed criticisms at people who abandon logic without better motivation than this. Even Michael Dummett, no logical conservative (he favours abandoning classical logic because of metaphysical scruples), once described these considerations as providing 'superficial' reasons to leave the realm of classical logic. Dummett later changed his mind on this point, so maybe he's getting more superficial in his old age, or maybe there is a reason for leaving classical logic after all.

One rather substantial issue that arises when we change logic concerns *our* reasoning practices as philosophers. If classical logic is shown to be defective because it improperly classifies all of (7) through (10) as true, then classical logic is defective and we shouldn't keep on using it. Hence any reasoning above (or elsewhere) that relied on classical logic should be viewed with suspicion. But that's pretty much all the reasoning ever undertaken, and certainly all the reasoning here. We seem to be stuck in a morass.

Let's illustrate how deep the morass is. I said above that the only responses to the Sorites were to say that the Sorites premises were not true, or the non-Sorites premises were not true, or the argument was not valid. That's a very standard way to frame the problem, but it should look a bit unstable, if not simple-minded, now. If we accept classical logic, then we can't reject the validity of the arguments, so there are really only two options not three. If we reject classical logic, then we need not accept the inference from the falsity of the conclusion to the truth of one of these options. So either way, it seems unlikely that these really are our options. But wait! My inference here took as an unargued premise that we either accept classical logic or we reject it. And that seems like a pretty dubious principle in the situation - it's just an instance of the law of excluded middle, and all but one of the non-classical logics we shall consider reject that law. So maybe there are just those three options after all.

There are some ways out, ways that we will explore quite a bit over the coming chapters. First, we can agree on a basic logic that we all share. Perhaps we can all agree that modus ponens is valid, for example. Secondly, we can reach more substantial agreement about which logical principles apply in certain fields. Maybe, for example, we can agree that classical logic, or something much like it, applies to reasoning about logics. In practice, it's the latter that usually is practiced, and whatever the merits of the first approach, we will adopt this here as well. (Unless there are specific reasons to not do so, when reasoning *about* a logical theory, we will take any intuitionistically valid argument to be valid. This is not a particularly bold assumption, even most non-classical theorists seem prepared to use classical logic when reasoning about theories, so I am not begging too many questions here.)

As well as producing surprising intuitions concerning the truth of various sentences, vagueness can produce strange results concerning our intuitions about which sentences are false. For example, one might have thought that a contradiction would sound intuitively false, yet (11) sounds acceptable to some people, where the 'he' is known to refer to Louis.

(11) He is bald and he isn't bald.

And if (11) is not obviously false, then presumably (12) is not obviously true, even when it is specified that the initial negation takes widest possible scope.

(12) It is not the case that he is bald and he isn't bald.

Timothy Williamson has argued, convincingly to my mind, that cases like (11) are *sui generis*, in effect idioms, and that when we judge (11) to be acceptable, this is not because vagueness makes us feel more warmly towards contradictions. Williamson notes that if you merely use different, but co-referring, singular terms in the two conjuncts, (11) will sound false. So just replacing the second pronoun with a name gives you (13).

(13) He is bald and Louis is not bald.

And I think intuitively (13) does sound clearly mistaken. Similarly, once we specify that the negation takes widest scope, (14) seems intuitively true.

(14) It is not the case that he is bald and Louis is not bald.

This raises another problem for dealing with these anomalies by abandoning classical logic. Once we get clear on the status of (11), i.e. that it really is an idiom, it seems vagueness does not produce any surprising results in the and/if fragment of logic. So far it seems as if sentences containing those two connectives behave just as classical logic predicts. Hence any revision of logic that aims to handle the anomalies will have to stick fairly close to classical logic in how it handles conjunction and conditionals, and this puts a serious constraint on the theoretical space available.

1.4. Vague Objects

So far we have concentrated on semantic vagueness. Some of the most interesting puzzles concerning vagueness arise not in semantics but in metaphysics. These are often discussed under the somewhat misleading heading "Could there be vague objects?" There are several questions being run together parate them out.

(a) Could there be two objects such that it is indeterminate whether they are identical?
(b) Could there be two objects such that it is indeterminate whether they entirely overlap?
(c) Could there be an object and a property such that it is indeterminate whether the object has the property?
(d) Could there be an object such that it is indeterminate whether that object exists?

Positive answers to any of these four questions would constitute, I guess, a kind of metaphysical vagueness. It is rather easy to run (a) and (b) together, but it shouldn't be done. In fuzzy set theory, for example, one answers 'No' to (a) but 'Yes' to (b). This combination of answers might not be too plausible for set theory, where it might seem vagueness is excluded, but might be more plausible in mereology.

In regular mereology, for any two objects x and y, there is a determinate fact as to whether x is a part of y or x is not a part of y. In fuzzy mereology, it can be the case that x is a part of y to a certain degree. At a first approximation, these degrees are real numbers in $[0, 1]$, where 0 represents complete non-parthood, and 1 represents complete parthood. (To less of an approximation the numbers are generated out of sets just as in classical mathematics, but the underlying set theory is fuzzy not classical, where fuzzy set theory has degrees of membership just like fuzzy mereology.) Now let y be the object that has this desk and this chair as determinate members and nothing else (*the* fusion of the desk and the chair as we non-fuzzies would say), and z the object that has the desk as a member to degree 1, and the chair as a member to degree 0.6, and no other members. If there are such objects y and z, they constitute positive instances of (b), since it sure looks like it is indeterminate whether they entirely overlap. But, according to the fuzzy mereologists, they are not instances of (a). Since y has a property that z lacks, namely, determinately having the chair a part, they are not identical, indeed determinately not identical. In short, objects with vague boundaries look like instances of (b), but it is easy enough to generate theories on which they are not positive answers to (a). Now that we're all clear on *that* distinction, it is worth looking at a few putative instances of (a) to see whether they are instances of (a), or of (b), or of both, or of neither. The following (long) quote is from Terrence Parsons's recent book *Indeterminate Identity*, at the point where he is setting up the puzzles that he thinks a theory of metaphysical identity must solve:

> *The person/body*: Assuming that a person exists when and only when their body exists, is each person identical with his/her body?
> Data: If I am alone in a room, then there is exactly one person in the room, and there is exactly one human body in the room.
> Working assumption: There is no answer to the question whether the person in the room is the body in the room.
>
> *The ship*: An assembly/repair process takes place [as in Hobbes's ship of Theseus case]. Is the original ship identical with the ship with new parts, or with the newly assembled ship (or neither)?
> Data: Exactly one ship left port, and exactly two ships docked.
> Working assumption: There is no answer to the question whether the original ship is the ship with new parts, or whether the original ship is the newly assembled ship.

Denoting such-and-such property? Trenton Merricks has recently argued that this familiar ways of isolating vagueness in the semantics because of these properties.

Question (d) might seem completely baffling. After all, how can there *exist* objects such that it is indeterminate whether they *exist* - isn't that a little like there being a fish such that it is indeterminate whether it is a fish? Well, the language gets a little messy here. As Lycan puts it, we all agree that, in a sense, there are some things that don't exist. (We can even name some of them!) Whatever formal way we find to frame that question, we can use to frame (d). Perhaps (d) is baffling for another reason, that we cannot make sense of the idea of there being an object whose very existence is indeterminate. But notice that positive answers to either (b) or (c) pushes us very close to a positive answer to (b). Assume that there are two objects such that it is indeterminate whether they entirely overlap. Let *x* be the mereological difference of one from the other. Then it is indeterminate whether *x* exists. Assume that there are object-property pairs such that it is indeterminate whether the object satisfies the property. Then it is indeterminate whether the state of affairs (or fact, if you prefer) <<That property (That object)>> exists. Neither argument here is conclusive, since the first relies on a pretty liberal conception of mereological difference, and the second on a pretty liberal conception of the existence of states of affairs, but you can see how we might be pushed to vague existence.

1.5. The Problem of the Many

The p-cats that Parsons discusses raise particular problems, enough to justify naming a separate problem after them. As Parsons notes, it seems to be determinately true that there is exactly one cat on the mat, even when there are billions of p-cats there. This is a different puzzle to the puzzles about indeterminate identity, since it is possible (perhaps even probable) that we can solve it by getting clear on the semantics, not on the metaphysics. To set up the problem, I will somewhat carefully argue for the claim that there are billions of cats on the mat, so the possible solutions should all appear as rejections of the premises I am using. (I ignore, for now, the possibility that the argument here is invalid.) The lines with stars are assumptions, the other lines are inferences. This might be helpful for keeping track of where we can reject the argument.

*1. There are billions of objects, no two of which are identical, that are all p-cats on the mat.
*2. If there are billions of objects, no two of which are identical, that are all p-cats on the mat, then there are billions of p-cats on the mat.
3. There are billions of p-cats on the mat.
*4. There is nothing on the mat that is a cat except the p-cats.
*5. There is at least one cat on the mat.
6. At least one of the p-cats is a cat on the mat.
*7. If one of the p-cats is a cat on the mat, then all the p-cats are cats on the mat.
C. There are billions of cats on the mat .

Five assumptions, so five possible moves. Or six, if you think accepting the conclusion is a live option (and I do). We won't go into this in any detail until later chapters, but I'll say just a little about each option. Rejecting 1 means rejecting the idea that any arbitrary collection of material objects has a fusion. Rejecting 2 means rejecting the standard analysis of ⌜There are *n* *F*s⌝ as ⌜There are *n* objects, no two of which are identical, and each of these is an *F*⌝, perhaps in favour of an analysis of counting sentences in terms of non-overlap. Rejecting 4 means countenancing vague objects, like cats, not identical with any precise object. Rejecting 5 probably means being like Frege and rejecting all claims containing vague terms, like 'cat'. And rejecting 7 means insisting that the p-cats, despite being indistinguishable to the human eye, might be different enough that one of them is a cat and the other not. And accepting the conclusion means one has some explaining to do concerning our ordinary linguistic practices. I guess that the most popular option is to reject 7, and claim that predicates like 'cat' name what Ted Sider calls *maximal* properties. A property is maximal iff no large part of a thing that satisfies the property satisfies the property. This might be plausible for 'cat', but before we sign off on this solution to the problem, it is worth thinking about whether the same solution holds if we run the problem with other predicates in place of ‘cat’. For a quick list of interesting examples, consider ‘cube’, ‘mountain’, ‘pile of garbage’ and

1.6. Vagueness and other defects

It is worth distinguishing vagueness from two other properties of terms which behave a little like vagueness, but are at least plausibly distinct effects. Arguably, vagueness is not ambiguity, and it is not merely indeterminacy.

You all know what an ambiguous term is, I hope. It turns out to be a little tricky to precisely say what ambiguity comes to, and even harder to test for it. Well, actually it isn’t that hard to test for ambiguity, but philosophers have made a bit of a hash of it in recent times. Let’s take a clearly ambiguous word, ‘pike’. This seems to be ambiguous between two nouns, one of which denotes a class of fish, and the of which denotes a class of weapons. Well, I guess each of these nouns is vague as well, especially the second. Pikes generally had pretty sharp points, that’s how they could be weapons, but I don’t suppose that there’s a clear standard for just how sharp the point must be, or for that matter how long the staff must be, for something to be a pike. Why not say that pike is three-way ambiguous, not two-way ambiguous? On one reading it denotes the class of fish, on another a narrowly defined class of weapons, and on a third a more broadly defined class of weapons? I’ll leave it to you to decide why (or whether) we should regard the fact that ‘pike’ can be used sometimes to denote fishes and sometimes to denote weapon is sufficient for it to be ambiguous, but the fact that it can sometimes be used to denote a large class of weapons, and sometimes a smaller class, is not sufficient.

One difference in the way theorists regularly treat vagueness and ambiguity concerns their contributions to the truth conditions of sentences. Say I take a particular weapon, clearly a pike, and say, “This is a pike”. What I say is true, even though it wouldn’t be if you took ‘pike’ to have its other meaning. Now imagine that the object I hold up has only a dubious claim to being a pike. It is in the large class of weapons I mentioned above, but not in the smaller class, because it is not clearly a pike (though perhaps it is also not clearly not a pike). This strikes some people as sufficient reason to say that my utterance is not true, though perhaps it is also not false. This is not a universal reaction: many people do not think vagueness has this effect on sentences, and Kit Fine has argued that even ambiguous sentences are only true if they are true on all disambiguations, so my claim would have been false in the first instance.

I mostly bring this up to note that it does little to distinguish vagueness from indeterminacy of the kind we’ve grown used to from discussions in Quine, Field, and to a lesser extent, Kripke. I’ll

s example, since it seems the most plausible. Like most scientists, Newton didn’t exactly define many of the key terms that he uses. Rather, he lets his terms be defined by their role in his theory. But his theory isn’t entirely true, which causes some interpretative difficulties. We can’t say,

mass, Newton meant to denote the property that plays such-and-such a role in the world”, where the such-and-such is replaced by a description of the role mass plays in Newtonian mechanics. The problem with this analysis is that no property plays precisely the role that mass plays in Newtonian mechanics. As Field notes, it won’t do to say that therefore the word in Newton’s mouth is meaningless, or even denotationless. Many of the things Newton said using mass are clearly true (particularly empirical claims like *The mass of this object is roughly one pound*). And several others seem close to being true, even if they are false. Well, we all know the way out of this problem. We find the property that comes closest to playing such-and-such a role, and analyse Newton’s word *mass* as denoting that property. The problem with this move is that there is no *unique* property that comes closest to playing the role in question. In relativistic mechanics, there are two properties, called *rest mass* and *proper mass*, and both of them are pretty good approximations to Newton's concept of mass, and neither seems a significantly better interpretation than the others. Field concludes, not unreasonably, that Newton's term *mass* was indeterminate between these two meanings.

It's rather tempting to see the vagueness in regular terms, like *bald*, as being like this kind of indeterminacy. The picture is that *all* words get their meaning, in the last analysis, by their role in a theoretical system, in most cases the folk theory of the world. Folk theories tend to be (a) not exactly true (though close) and (b) not complete, so the theory does not determine a unique denotation for many words. Just like we can interpret *mass* in Newton's mouth in two ways without being uncharitable, we can interpret *bald* in each other's mouths in many ways without being uncharitable. Taking this kind of reasoning particularly seriously leads to the supervaluationist theory of vagueness. But even if one isn't a supervaluationist, this kind of picture linking vagueness with this Quinean indeterminacy seems attractive.

I don't want to draw any positive conclusions about whether vagueness and this kind of indeterminacy are the same phenomenon, though I will note one difference between the two. There are usually two substantial (and related) differences between the range of meaning options left open by vagueness and those left open by indeterminacy. When we have a vague term, there are normally many many ways to interpret the word consistent with charity. When we have an indeterminate term, like *mass*, there are generally a small number of possible interpretations. So that's the first difference, in how many possible interpretations there are for the word. The second difference is that in the case of vagueness, those possible interpretations tend to be, loosely speaking, convex. If I_1 and I_2 are possible interpretations of *bald*, then interpretations 'between' these two generally will be as well. Generally, with vague terms, our practices narrow the interpretations down to an area, and anything in that area counts as an acceptable interpretation. With indeterminate terms, we don't get that effect. Newton definitely did not mean (rest mass + proper mass)/2 by *mass*. Now I'm not sure just how big a deal these differences are, but they exist, and they are worth noting.

1.7. Vagueness is Pervasive

As Russell realised, virtually every word in English is vague. Perhaps the logical connectives are not vague, though if you believe both logical pluralism and an inferential role semantics for the connectives, then you will think they are at least indeterminate. (That might be a reason for not mixing these views, but I do not want to take sides in that debate right here.) Perhaps numbers, if they are not logical terms, are another class of words that are not vague. But beyond that vagueness seems to be everywhere, especially if we take Quinean (or Kripkean, or Fieldean) indeterminacy to be vagueness

For an illuminating example of this, consider the predicate 'has more children than the average'. Theresa Robertson uses this in a recent paper as an example of a predicate that is not vague. And to be fair it is, in a loose sense, less vague than the other terms she has been considering in the context, but it surely is vague. It inherits its vagueness from the vagueness of the predicate *child of*. It is not too hard to come up with borderline cases of this. For example, if *a* is *b*'s biological child, and *b* is *c*'s spouse, does that make *a* a child of *c*, even if *a* and *c* have no other relation? I think in most cases, the answer is *Yes*, though it does become much less clear if *a* does not live with *b* and *c*, and especially if *a* is an adult. If *a* is *b*'s adopted child, and *b* is *c*'s spouse, and *a* is a 19 year old college student, who lives with her other (adopted) parent over summer break, I very much doubt we would say *a* is a child of *c*. Though if the circumstances were the same, except that *a* is 16 years old, and living with *b* and *c*, we would say that *a* is a child of *c*. It isn't too hard to come up with borderline cases between these, where it just seems vague whether the relation holds. And we can easily vary the situation enough so that whether *a* is *c*'s child determines whether *c* has more children than average or not.

Not only is vagueness pervasive, Sorites paradoxes are pervasive. Scott Soames notes that there is a difference in kind between the vagueness in a term like *tall* or *red*, where there is an obvious Sorites paradox, and the vagueness in *vehicle*. Soames thinks that although it is vague whether skateboards are vehicles, this does not imply that we can generate a Sorites paradox, because we cannot produce the small steps between skateboards and paradigm vehicles needed to generate a Sorites. Whether or not we can produce a chain of small steps that turns a skateboard into a vehicle, we certainly can produce a Sorites paradox for vehicle, as follows.

A car that is not presently functioning, perhaps because the battery has been disconnected, is certainly still a vehicle. So if you start taking pieces off a car, you don't cause the vehicle to go out of existence the first time you do something that makes it dysfunctional. But if you take away enough pieces you will have caused the vehicle to cease existing. A CD changer, for example, is not a vehicle. But when do you destroy the vehicle? If you took away small enough pieces at a time (maybe an atom at each step!) then it seems very implausible at each step that you would have destroyed the vehicle just then. That is, it seems very plausible that if there is a vehicle there after *n* atoms have been removed, then there is a vehicle there after *n*+1 atoms have been removed. The Sorites argument that we can generate should now be obvious.

Soames does seem to be on to an interesting, and possibly important, distinction here, but it is not the distinction between vague terms that support Sorites paradoxes, and those that do not. As far as I can tell, all vague terms support Sorites paradoxes. Stating what Soames's distinction amounts to requires a little bit of terminology from later chapters, so this may sound a little mysterious to some. What Soames has in mind is the distinction between vague terms whose precisifications are linearly ordered by strength, and those that are not. Or, equivalently, the distinction between terms whose precisifications vary in just one dimension, from those whose precisifications vary in many dimensions. Even in the latter case, there will still be Sorites paradoxes around, essentially because we can always focus on just one dimension to construct the paradox.

Vagueness even seems to infect theorising about vagueness. Grice's definition of vagueness made important use of borderline cases. A term is vague if there are cases to which it is not determinate that it applies. Russell called these penumbral cases, and this terminology has stuck. So a natural picture is that a vague term like *red* divides objects into those that clearly satisfy it, those that do not clearly satisfy it, and the penumbral cases. But this will not do. Just as is there is no sharp line between the red and not red lights, there is no sharp line between those lights that are clearly red, and those that are penumbral cases. So it seems there should be cases that are not clearly cases of being clearly red, and that are not clearly penumbral cases. These are cases of what is known as higher-order vagueness. As we shall see, accounting for higher-order vagueness is a serious stumbling block for several theories of vagueness.

Having said all that about how pervasive vagueness is, I'll conclude this section by arguing against one recent claim to discover even more vagueness. John Broome argues that apparent incommensurability is really vagueness. For some comparatives '*F*er than', there are cases where *a* is not clearly *F*er than *b*, and *b* is not clearly *F*er than *a*. To reuse an example we gave above, arguably it is not clear that Churchill was a better leader than Roosevelt, or that Roosevelt was a better leader than Churchill. Now this does not immediately imply vagueness. It certainly *seems* consistent to hold that the logic of 'is a better leader than' is the same as the logic of 'is an ancestor of'. Just as Churchill is not an ancestor of Roosevelt, nor Roosevelt an ancestor of Churchill, perhaps Churchill is just not a better leader than Roosevelt, nor Roosevelt better than Churchill. In some formal settings, we get cases much like this. For example, we normally say that L_1 is a stronger logic that L_2 iff everything you can prove in L_2

you can prove in L_1, but not *vice versa*. This leaves open the possibility that L_1 will not be stronger than L_2, nor L_2 stronger than L_1. (For example, the modal logics KB and S4 are such that neither is stronger than the other.) If we think that *x* is a better leader than *y* if *x* could do everything *y* could do (relative to political leadership) but not *vice versa*, then there will often be pairs where neither is a better leader than the other. Alternatively, we might insist that one of these two is better than the other, but that it is vague which of them is the better. Broome argues that certain properties of the better-than comparatives force us to adopt the second interpretation. The main argument Broome adopts for this is that the alternative view, on which neither is a better leader than the other, removes too much vagueness.

Broome argues that if there are pairs *x* and *y* such that neither is *F*er than the other, then there can be no borderline cases of being *F*er than *x*. In this case, *y* is not a borderline case, it is *not F*er than *x*. We can set out Broome's argument by developing our case of better leader than. Broome assumes, reasonably, that even if there are incomparable pairs like this, we can still create a Sorites series of (possible) leaders, each a little better than their predecessor, around any given candidate. So we can imagine Roosevelt´, who is as good as Roosevelt in all respects bar one, in which he is a little better, and Roosevelt´´ who is as good as Roosevelt´ in all respects bar one, in which he is a little better, and so on. Intuitively, eventually we should get to a person in this series for whom it is genuinely indeterminate whether he is a better leader than Churchill. Rather than use all the primes, I will call this person Lincoln, though if you disagree with my historical judgements here, feel free to substitute your own candidates.

Remember that it is clear, given the way the example, that Churchill is not a better leader than Lincoln, because Lincoln is better than Roosevelt, and Churchill is not better than Roosevelt. It is also meant to be indeterminate whether Lincoln is better than Churchill. Broome thinks this is impossible, because he holds the following principle.

The collapsing principle: For any *x* and *y*, if it is more true that *x* is *F*er than *y* than that *y* is *F*er than *x*, then *x* is *F*er than *y*. (Broome 1997: 77)

To see the problem, set *x* to be Lincoln, and *y* to be Churchill. By hypothesis, it is entirely false that Churchill is better than Lincoln, but in a sense indeterminate whether Lincoln is better than Churchill. Hence, in a loose sense, it is truer that Lincoln is better than Churchill than that Churchill is better than Lincoln. Hence, by the collapsing principle, Lincoln is better than Churchill. This shows that there can be no members of the Sorites chain around Roosevelt such that it is indeterminate whether they are better leaders than Churchill. But of course there can be such cases, so the assumption that neither Churchill or Roosevelt is better than the other must be mistaken.

The argument here relies crucially on the collapsing principle, and this is where it goes wrong. Broome has a little argument for the collapsing principle, but I think we can see how this argument fails. To set up the argument, return to the hypothesis that neither Churchill nor Roosevelt is better than the other, and Lincoln is related to the two of them in the ways set out above. Imagine you have to decide which political leader will be inducted next into the political leaders Hall of Fame. Churchill, Lincoln and Roosevelt are clearly better leaders than any other eligible candidates, and your instruction is to find the best leader to be the next inductee. Since Lincoln is better than Roosevelt, we can cut the field down to two: Churchill and Lincoln. Broome thinks that since it is clearly false that Churchill is better than Lincoln, but it is indeterminate whether Lincoln is better than Churchill, then (a) you should choose Lincoln, and (b) this shows that Lincoln really is better than Churchill after all, since your instruction was to choose the best leader available. Generalising from this case in the obvious way, we get support for the collapsing principle.

I doubt that (a) really is obviously true, but even if it is there are several reasons to think that (b) does not follow from it. As Broome notes, if it really were the case that neither Churchill nor Lincoln is better than the other, presumably the right thing for the chooser to do would be to toss a coin. So I don't quite see why, since it is indeterminate whether it is the case that neither Churchill nor Lincoln is better than the other, it isn't indeterminate whether this is the right thing to do. Perhaps this is because there is an implicit instruction in these cases that the chooser should not rely on chance mechanisms to make the choice. But this would undercut the claim that (b) follows from (a): if there are more constraints on the chooser than just *choose the best leader*, then the fact that they should choose Lincoln does not show that Lincoln is better. In fact, I doubt that *choose the best leader* is even one of the instructions, let alone the only instruction, even if it was the explicit statement of the instructions. Any kind of ranking procedure like this adopts the pretence that it is possible to strictly (or at least completely) order the candidates. And there is a clear sense in which (F) is true in this case:

(F) According to the pretence that the candidates can be strictly ordered, Lincoln is better than Churchill.

Given the asymmetry between the two, that Churchill is clearly not better than Lincoln, but it is not clear whether Lincoln is better than Churchill, it is clear that if you had to make a linear ordering of the leaders, Lincoln should go above Churchill. And I think it is fairly clear, if often implicit, in these cases that the chooser should make her choices according to what the rankings would be according to just this ordering, which we have to pretend exists if the concept of a political Hall of Fame is to be coherent. So if Lincoln should be inducted, this is because (F) is true, and the fact that Lincoln should be inducted shows no more than that (F) is true, but (F) is obviously compatible with it being indeterminate whether Lincoln is better than Churchill. So there can be indeterminate cases of 'better than Churchill', even if neither Churchill nor Roosevelt is better than the other. So the only *argument* that these cases of apparent incommensurability are cases of vagueness fails. So vagueness is everywhere, but still it isn't quite as pervasive as some have argued.

1.8. Intuitionist Logic

I said above that I would take our default logic for reasoning about logical systems to be intuitionist logic. Some readers may be unfamiliar with this logic, so I finish this chapter with a quick guide to it. Those of you who already are familiar with intuitionist logic, or who are prepared to take everything I say using it on trust, can jump ahead to chapter 2.

Intuitionist logic (sometimes called constructive logic) is a lot like classical logic, but with one interesting difference. In it, one can never prove a disjunction without being able to prove one or other disjuncts, and more generally, one can never prove an existentially quantified formula without being able to prove an instance of it. This makes for a sharp divide with classical logic, where (15) and (16) are both theorems.

(15) $A \vee \neg A$

(16) $\exists x\, (Fx \rightarrow \forall y\, (Fy))$

We'll focus on the propositional version of intuitionist logic here, but it is worth keeping (16) in mind when we're discussing the logic further. It can often seem that intuitionists just differ from classicists in how they treat negation, perhaps leading to the suspicion that the two sides just mean something different by negation. But as (16) shows, the two sides differ over sentences that do not contain any negations, so the debate is a little deeper than that.

There's a pretty straightforward argument that (15) is a logical truth, so intuitionists have to find a place to deny it. The argument only uses fou r logical principles, set out here:

($\vee I_R$) From A infer $A \vee B$
($\vee I_L$) From A infer $B \vee A$
(DNE) From $\neg\neg A$ infer A
(RAA) For any set of sentences Γ, if Γ and A imply B and $\neg B$, then Γ implies $\neg A$

Here is the argument, again with assumptions starred
*1. $\neg(A \vee \neg A)$
*2. A
3. $A \vee \neg A$ From 2 by ($\vee I_R$)
So from premises 1,2 we have $A \vee \neg A$ and $\neg(A \vee \neg A)$. Hence by (RAA) $\neg A$ follows from 1 alone. We continue with that, noting that we are no longer relying on premise 2.
4. $\neg A$
5. $A \vee \neg A$ From 4 by ($\vee I_L$)
So from premises 1 alone (or plus the null set if you prefer) we have $A \vee \neg A$ and $\neg(A \vee \neg A)$. Hence by (RAA) $\neg\neg(A \vee \neg A)$ follows from the null set alone. We continue with that, noting that we are no longer relying on premise 1.
6. $\neg\neg(A \vee \neg A)$
7. $A \vee \neg A$ From 6 by (DNE)

Since we don't like the conclusion, we had better object to one of the steps in the proof. And the step intuitionists reject is the last one, the elimination of the $\neg\neg$ from the front of the formula. This might seem like an odd place to get off the boat, but it makes for a rather interesting logic. And it is justifiable if we understand content in the way intuitionists typically do. Intuitionist logic is normally linked with one or other kind of anti-realism such as verificationism in metaphysics or constructivism in philosophy of mathematics. We'll run with the verificationist version for now.

The idea is that the content of a sentence is just its verification conditions. How does this carry over to the meanings of the connectives? Well, the way to verify a conjunction is to verify one or other conjuncts. The way to verify a conditional $A \rightarrow B$ is to have a method for converting any verification of A into a verification of B. The way to verify a disjunction $A \vee B$ is to verify A or to verify B. This is sometimes weaken ed so that one can verify $A \vee B$ by developing an procedure that one can verify will, when run, yield in a finite amount of time a verification of A or a verification of B. And the way to verify a negation $\neg A$ is to verify that no verification of A will be forthcoming. So the way to verify $\neg\neg A$ is to verify that no verification that there is no verification of A will be forthcoming. (You might want to read that a couple of times to make sure you've got the scopes right!) And we can do that by, for example, showing that one could never rule out A, even if we haven't actually ruled A in. That's exactly what happens with instances of the law of excluded middle. We can never rule out the possibility that someone, someday, will either verify A or verify $\neg A$. (The universe is going to be around for a while, and people are smart, so you know, we should err on the safe side.) In fact, as the above argument shows, it is inconsistent to hold that both A and $\neg A$ have been ruled out, just because what it is for $\neg A$ to be verified is for A to be ruled out. So we have ruled out $\neg(A \vee \neg A)$, i.e. we have $\neg\neg(A \vee \neg A)$. But we don't have an effective procedure for working out which, so we shouldn't claim to have verified $A \vee \neg A$. Hence double negation elimination, DNE from now on, is dubious.

The same kind of point arises in the quantified version of intuitionist logic. Classically, we can infer from not everything is F to something is not F. This won't work in intuitionist logic. To verify that something is F, you must produce the F, or at least have an effective procedure for producing it. Now one can have a way of ruling out the possibility that everything is F, but without having any idea how to produce the thing that is not F. Let's take a concrete instance of this. Consider pairs of adjacent light beams in our original example, and say such a pair is congruent if they are either both red or both not red, and incongruent otherwise. Given that the first beam is red, and the last is not red, it cannot be the case that all pairs are congruent. (Remember that we have RAA and modus ponens as rules of inference, so we can easily derive a contradiction from the claim that all pairs are congruent, so that hypothesis is false.) The intuitionist *denies* that we can infer from this that some pair is incongruent, at least in the absence of an effective procedure for determining the pair. And I don't think such a procedure is forthcoming any time soon.

At last, I can imagine you thinking, a connection with vagueness. It turns out intuitionist logic has a few surprising connections with vagueness. Hilary Putnam argued that this was a solution to the Sorites a while ago, though the idea hasn't done very well in the interim. It's not immediately clear whether intuitionist solutions to the other puzzles mentioned fare just as badly. If you think the right answer to the issues about the anomalies is that instances of the law of excluded middle sometimes fail to be true, and instances of the law of non-contradiction are always true, then you might well think intuitionist logic is the logic of natural language. Even if you don't think intuitionist logic provides much help, you might think that a small modification to its proof theory or semantics will provide the logic that solves all the problems. Since a few of people have thought this, we'll look at the proof theory and semantics for intuitionist logic.

But first to come back to the morass about how to have debates about logic. As noted, there seems to be something of a consensus that when reasoning *about* logic, one is entitled to use a logic at least as strong as intuitionist logic. This is a kind of disjunctive consensus: most everyone thinks classical logic is appropriate for the meta-language, except the intuitionists, who think intuitionist logic is appropriate. Since classical logic is stronger than intuitionist logic, my claim about the consensus is true. (There is something noble, I think, about those brave heretics in the intuitionist camp who insist on the logical standards they hold to be true, even in debates within philosophical theory. I'd be embarrassed if I was a devotee of some deviant logic or other who routinely slipped into classical reasoning as soon as I entered the philosophy classroom.) So that's the standard I'll use if we can generate problematic conclusions using intuitionistically acceptable rules of proof from theoretical claims made by a particular theory, that will count as a substantial cost of the theory.

What are those rules of proof? Well, provided you've been taught classical logic the right way, i.e. without too many redundant rules, they are just the rules of proof you learned for classical logic, with two exceptions. The first exception is that DNE is not an acceptable rule. The second exception, in fact an exception to the exception, is that the rule *ex falso quodlibet*, or EFQ, is acceptable. This rule says that if $\Gamma \vdash A \wedge \neg A$, then $\Gamma \vdash B$. The proviso above is that these are the only exceptions to a fairly stripped down version of classical logic. Anyone who's learned logic via a system in which various things called De Morgan laws are accepted as rules of proof might have to learn a few more exceptions. Rather than go into all the possible levels of detail, I'll just state for the record a proof system for classical logic, and then state how to modify it to produce classical logic. We will use a Gentzen-style natural deduction system, where we infer the validity of certain sequents from the validity of other sequents. We'll take sequents to be ordered pairs of sets of formulae and formulae, and write $\Gamma \vdash A$ to mean that the sequent $<\Gamma, A>$ is valid. (For some purposes, it is better to take the first member of the pair to be a sequence of formulae, not a set of formulae, and then have rules to confirm that the order of the sequence, and how many times the formulae appear, are irrelevant. We ignore these concerns here, because they do not seem relevant to vagueness.) Here are the rules, with some suggestive names. Note that any biconditional iff always takes wide scope in the following statements.

A. If $A \in \Gamma$, then $\Gamma \vdash A$
W. If $\Gamma \subseteq \Delta$, and $\Gamma \vdash A$, then $\Delta \vdash A$
$\wedge$. $\Gamma \vdash A$ and $\Gamma \vdash B$ iff $\Gamma \vdash A \wedge B$
$\rightarrow$. $\Gamma \cup \{A\} \vdash B$ iff $\Gamma \vdash A \rightarrow B$
$\vee$I. If $\Gamma \vdash A$, then $\Gamma \vdash A \vee B$, and $\Gamma \vdash B \vee A$
$\vee$E. If $\Gamma \vdash A \vee B$, and $\Gamma \cup \{A\} \vdash C$, and $\Gamma \cup \{B\} \vdash C$, then $\Gamma \vdash C$
RAA. If $\Gamma \cup \{A\} \vdash B \wedge \neg B$, then $\Gamma \vdash \neg A$
DNE. If $\Gamma \vdash \neg\neg A$, then $\Gamma \vdash A$
$\forall$E. If $\Gamma \vdash \forall x(\phi x)$ then $\Gamma \vdash \phi t$
$\forall$I. If $\Gamma \vdash \phi t$, and t does not occur in ϕ or in any member of Γ, then $\Gamma \vdash \forall x(\phi x)$
$\exists$E. If $\Gamma \cup \{\phi t\} \vdash C$, and t does not occur in ϕ or C or any member of Γ, then $\Gamma \cup \{\exists x(\phi x)\} \vdash C$
$\exists$I. If $\Gamma \vdash \phi t$ then $\Gamma \vdash \exists x(\phi x)$

To get intuitionist logic, we drop DNE, and add the following rule, EFQ.

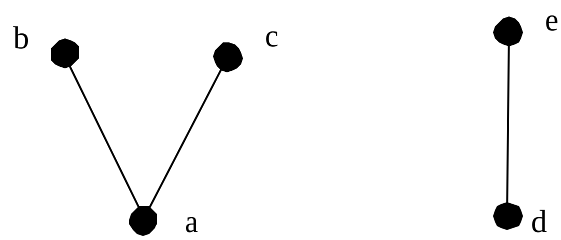
EFQ. If $\Gamma \vdash A \wedge \neg A$ then $\Gamma \vdash B$

Because intuitionists do not accept excluded middle, they cannot say that an argument is valid whenever there is no line on the classical truth table where the premises are all true and the conclusion false. So it might be wondered just what semantic model they could provide for their logic. At one stage it was thought that some many-valued truth tables like those we'll consider next week could do the trick. But Kurt Gödel showed that this was impossible: intuitionist logic is not equivalent to *any* n-valued logic, for any value of n, and for any truth tables you can construct. So the semantic models for intuitionist logic will be a little trickier than those for classical logic. The easiest models to understand were developed by Saul Kripke in the 1960's. We'll just look at the semantics for the propositional part of intuitionist logic, though if anyone likes we can also talk about the semantics for predicate logic.

The models consist of sets of points (call the set S) that are ordered by an reflexive, anti-symmetric, transitive relation (call the relation R), and a valuation function on those points. A relation is reflexive iff for all points a, a bears the relation to itself. A relation is anti-symmetric iff there are no distinct objects a and b such a bears the relation to b and b bears the relation to a. A relation is transitive iff for any three objects a, b and c such that a bears the relation to b and b bears the relation to c, then a bears the relation to c. If a bears the relation to b, we will write aRb for short. One can think of the points informally as states of knowledge, and aRb holds iff b is a state of knowledge one could get to from a. The valuation function says which propositional constants (p, q, r, etc.) are true at which points. Strictly, the valuation function just says whether a propositional constant is *true* at a point a; whether it is false at a is left undetermined by the valuation function for a. So the function associates a set of propositional constants with each point, the set of constants that are *true* at that point. It is a restriction on models that if p is true at a, and aRb, then p is true at b. The idea is that if we know p at a, and we can get from *a* to *b* by learning more things, then we still know *p* at *b*. The real work is done by the definitions of truth of compound propositions in these various states of knowledge. We will write $x \vDash A$ for *A* is true in state of knowledge *x*.

$x \vDash A \wedge B$ iff $x \vDash A$ and $x \vDash B$
$x \vDash A \vee B$ iff $x \vDash A$ or $x \vDash B$
$x \vDash \neg A$ iff for all *y* such that *xRy*, $y \nvDash A$
$x \vDash A \rightarrow B$ iff for all *y* such that *xRy*, if $y \vDash A$ then $y \vDash B$.

Let's see a few of quick examples of this to see these semantics provide the distinctively intuitionist results. Consider the following two models.



In the left hand model, the relation *R* is such that *aRa*, *bRb*, *cRc*. *aRb* and *aRc*. The valuation function says that *p* is true at *b*, and says nothing else. A few things follow from this. $p \vee q$ is true at *b*, since *p* is true at *b*. $\neg p$ is true at *c*, since no point accessible from *c* has *p* true at it. But $\neg p$ is true at *a*, since *aRb* and *b* is true at *a*. Nor is *p* true at *a*, since the valuation function did not assign *p* to *a*. Hence $p \vee \neg p$ is not true at *a*. So there are points in models at which $p \vee \neg p$ is not true.

In the right hand model, the relation *R* is such that *dRd*, *eRe* and *dRe*. The valuation function says that *p* is true at *e*, and says nothing else. Now at *d*, we do not have $\neg p$ true, since there is a point accessible from *d* at which *p* is true. Indeed, there is no point accessible from *d* at which $\neg p$ is true. Hence $\neg\neg p$ is true at *d*. But *p* is not true at *d*. So $\neg\neg p \rightarrow p$ is not true at *d*. So there are points in models at which $\neg\neg p \rightarrow p$ is not true.

As you may have guessed by now, a formula is a theorem of intuitionist logic iff it is true at all points in all models. A sequent is valid iff there is no point in any model at which all the premises are true and the conclusion is false. So the right-hand model also shows that $\neg\neg p \vdash p$ is not valid in intuitionist logic. The proof theory described above is sound and complete with respect to these models.

2. Many Valued Logics

I suppose few of us learned classical logic by *first* learning the rules of proof listed above. Rather, we first learned the truth tables, or something much like them (like semantic tableaux), and only later bothered about messy things like plausible sounding proof theories. This is certainly how *I* teach introductory logic courses, and I get the impression it's a pretty common practice. And it is easier to get oneself worried about the applicability of the truth tables to a vague language than to get worried about any of the rules of inference mentioned above. Using the truth tables requires taking for granted that every sentence in the language is either true or false, an assumption that vagueness seems to challenge. Although we can *prove* that all instances of the law of excluded middle are true using the above rules, we do not *assume* this, so there is nowhere as obvious to challenge the proof theory as there is to challenge the semantics.

As the name suggests, many-valued logics deviate from classical logic by altering the semantics. They keep the idea of truth tables, the idea that the truth value of compound sentences is a function of the truth values of their constituents. (Quantified versions also keep the idea that the truth value of a quantified sentence supervenes on the truth values of all of its instances, but we will focus here on propositional versions.) They differ from classical logic in two respects. First, they hold that there are more than two truth values that sentences can take. Secondly, they disagree with classical logic about which sentences take maximal truth values. There are a few different ways this disagreement can be turned into a positive theory. I'll first sketch one of the simpler versions, and then generate a taxonomy of possible variations.

2.1. The Kleene Logic

The following logic is based on truth tables developed in the 1950s by S. C. Kleene, a mathematician. Kleene did not intend these to be the foundations of a logic – he just used them as a technical convenience to state some results concerning combinations of functions. But whatever their origins, they make for a pretty enough theory. In the theory, we have three values that sentences can take, which for now we will denote as 1, ½ and 0. If you like you can think of 1 as truth, 0 as falsity, and ½ as some intermediate state, though just how the values should be interpreted will be a matter of some debate in what follows. If sentences can take more than two truth values, then the normal Tarskian definition of truth of compound sentences is incomplete. The Tarskian definition, in effect the traditional truth tables, tells us what the truth value of a compound sentence is when all of its parts take traditional truth values. But it falls silent on the question of the truth value of sentences whose parts take non-traditional truth values. So we need new tables. The following are from Kleene, I assume that their interpretation is fairly obvious.

A	$\neg A$
1	0
½	½
0	1

$A \wedge B$	1	½	0
1	1	½	0
½	½	½	0
0	0	0	0

$A \vee B$	1	½	0
1	1	1	1
½	1	½	½
0	1	½	0

$A \rightarrow B$	1	½	0
1	1	½	0
½	1	½	½
0	1	1	1

Note a few things about these tables. First, they are tables! The truth value of a conjunction, disjunction, negation or conditional is determined entirely by the truth values of the individual parts. This is a non-trivial fact about these logics, one that not all the theories we shall look at below will share. Secondly, there is a nice algorithm underlying each of the tables. The value of $\neg A$ is 1 minus the value of A; the value of *A & B* is the lower of the values of *A* and *B*; and the value of $A \vee B$ is the greater of the values of *A* and *B*. There is a wide consensus that if we are to have a three valued logic based on tables, then these should be the table for $\neg$, & and $\vee$. There is less consensus about what should be done with $\rightarrow$. In the Kleene table the identification of $A \rightarrow B$ with $\neg(A \wedge \neg B)$ is preserved. One might wonder why, if one was keeping *any* feature of the classical truth tables, why one would choose to keep just *this* morsel, but that is what is done.

Thirdly, for *any* sentence of any complexity, if all the atomic sentences within it get value ½, then the whole sentence gets value ½. So if *Louis is bald* gets this new truth value, then so do (1), (2) and (3). This might strike some of you as, well, odd.

(1) Louis is bald or Louis is not bald.
(2) It is not the case that Louis is bald and Louis is not bald.
(3) If Louis is bald then Louis is bald.

It is no surprise that (1) gets this value, indeed this result seems to be one of the major attractions of the theory to its proponents, but that (2) and (3) fail to be perfectly true, well that's a little strange.

If we think that these truth tables, rather than the classical truth tables, are appropriate for English, then we have something to say about the Sorites arguments. We can have it the case that light of wavelength 700nm is red, and light of wavelength 400nm is not, while there is no value of *n* such that it is true that light of wavelength *n* nm is red, and light of wavelength *n* - 0.1 nm is not. Of course, for many values of *n* this conjunction will not be false, but there is no value of *n* for which it is *true*. That at least looks like progress. And it is a benefit shared by all of these many-valued theories. The other costs and benefits turn a little more on the details of the theories in question, so we will stop to survey the options available to the would-be many-valued theorist.

2.2. A Little Taxonomy

To a first approximation, we can classify many-valued theories on the basis of their answers to the following five questions. The questions are not *entirely* independent, some combinations of answers will sound particularly incoherent, so don't think that with five questions we have thirty-two live options to consider!

(a) What value does ½ $\rightarrow$ ½ get?

In Kleene's tables, if *A* and *B* get the value ½, then so does $A \rightarrow B$. This might seem like a flaw in the system, since it means that if *A* gets this new value, then so does $A \rightarrow A$. But surely that's a logical truth, and surely its status as such is not threatened by *vagueness*. If these kinds of concerns are compelling, and personally I think they should be, then one might prefer the logic developed by the Polish logician Jan ukasiewicz. ukasiewicz was interested in three valued logics not because of concerns about vagueness, but because he wanted to develop Aristotle's idea that future contingent sentences whose truth value depends on free actions by humans do not have a (traditional) truth value. His logic was just like Kleene's except for the answer it gives to the above question. So we have the following tables:

A	$\neg A$
1	0

$A \wedge B$	1	½	0
1	1	½	0

$A \vee B$	1	½	0
1	1	1	1

$A \rightarrow B$	1	½	0
1	1	½	0

½	½	½	½	½	0	½	1	½	½	½	1	1	½
0	1	0	0	0	0	0	1	½	0	0	1	1	1

For three connectives, the same values are applied, so the same algorithms still apply. The value of a conjunction is still the minimum value of the two conjuncts, the value of a disjunction is still the maximum value of the two disjuncts, and the value of a negation is still one minus the value of the negated sentence. But the conditional now no longer fits into these categories. Indeed, it is trivial to prove that $A \rightarrow B$ is not equivalent to any sentence expressible using $\wedge$, $\vee$ and $\neg$. (Exercise: Prove it.) Roughly, the value of $A \rightarrow B$ is a measure of how far away from the truth you go by moving from A to B. If you move no further away (or even closer), the conditional gets the value 1, otherwise its value is 1 minus the difference between the value of A and the value of B.

(b) Are there three values or more?

Two truth values wasn't enough to capture all of the structure of vague language: we needed a third value for the intermediate cases. But perhaps there is even more structure that we should consider. The following example, due to Roy Sorenson, brings this out. Sorensen discusses a case of two sisters, call them Amy and Betty. They are identical twins, but Amy is born a little before Betty. They each grow up to be basketballers, but Betty is always a little shorter than Amy. (To be precise, at any time she's just the height that Amy was an hour ago, but Amy is growing.) Imagine a time when Amy and Betty are somewhere around the penumbra, that vague area between the tall and the not-tall. We do want to say that if Betty is tall then so is Amy, but it is not that clear that if Amy is tall, then so is Betty. On the Kleene tables, both these conditionals get a value ½, on the ukasiewicz tables they both get a value 1. Basically, it seems we should be able to distinguish between the two conditionals, and the many-valued approach cannot.

There is an easy way to fix this. Rather than just recognising one new truth value, we can claim there are an infinity of truth values between 0 and 1, one for each real in [0, 1]. As Amy gets taller, the truth value of *Amy is tall* gets higher and higher, reaching 1 when she definitely becomes tall. With all these truth values we have to provide a new series of accounts of the truth values of compound sentences in terms of the truth values of their parts. But this is easy given the algorithms we've established. Letting $V(A)$ be short for the value of A, we have the following rules for the Kleene and ukasiewicz logics. You can confirm these are just formal statements of the principles given above.

Both: $V(\neg A) = 1 - V(A)$
$V(A \vee B) = \max(V(A), V(B))$
$V(A \wedge B) = \min(V(A), V(B))$
Kleene: $V(A \rightarrow B) = \max(1 - V(A), V(B))$
ukasiewicz $V(A \rightarrow B) = \min(1, 1 - V(A) + V(B))$

As you might have guessed, $\max(x, y)$ is the larger value of x and y, and $\min(x, y)$ is the smaller.

The ukasiewicz theory now makes the conditional *If Betty is tall then Amy is tall* perfectly true at every point in time, while its converse *If Amy is tall, then Betty is tall* is not perfectly true. On the Kleene theory, the first conditional always gets a higher truth value than the second, but it's value can be rather low at times. This seems bad, just like it seemed bad that $A \rightarrow A$ wasn't always perfectly true.

(c) What do valid inferences preserve?

Fixing a semantics for the truth values of compound sentences tells us which sentences get which truth values in which circumstances, but doesn't yet tell us all that we need to know about the *logic*. Since logic is the study of inferences, we haven't yet been told a crucial detail of the logic until we are told what counts as a valid inference. There are a few options here, but two are particularly salient.

First, we might say that an argument is valid iff whenever the premises all receive truth value 1, then so does the conclusion. This is the idea that validity is the preservation of truth. Secondly, we might say that an argument is valid iff there is no valuation on which the value of the conclusion is lower than the value of any of the premises. This is the idea that validity is the preservation of truth value: a valid argument can never take you further from the truth than you started. In two-valued systems, these two positions are equivalent, but they are not equivalent here. For example, in either the three-valued or the real-valued ukasiewicz systems, $A, A \to B \vdash B$ is valid on the first way of understanding validity, but not on the second. Now this might at first look like an argument for the first concept of validity until we remember that this very inference seemed to get us until trouble in the Sorites. I don't want to take sides here, just note that there is a reason to think of validity as preservation of truth-value, not merely as preservation of truth.

There is a third option which strikes me as a little ad hoc, but might seem plausible to you. This is only an option if we have decided to go with infinitely many truth values, rather than just three. We could say that there is a special value, let's say 0.9 for the sake of argument, above which sentences are alidity is the preservation of 'almost truth'. Again, $A, A \to B \vdash B$ will not be valid in ukasiewicz logics on this concept of validity. The truth-preservation concept of validity says that an argument is valid iff it preserves 'almost truth' on every way of understanding almost truth.

(d) Are these new values? Part I - Are they values at all?

The above three questions all concern the technical issues in setting up the new logic. The following two questions concern the philosophical interpretation of the formalism. Return for a minute to the three-valued theories with which I opened. There is a division between proponents of these theories as to whether sentences that take the 'value' ½ have a new truth value, or simply fail to take a truth value.

I don't have a great deal to say on this subject, as I'm not sure it's particularly exciting. I'll just note that the idea that ½ is not a truth value might be attractive if one is a little squeamish about departing too far from classical logic. On this picture, you can still say that classical logic is appropriate when all the sentences under consideration take truth values. The only mistake classicists make is to think that sentences infected by vagueness are necessarily truth valued. If this helps you feel better about being a heretic I suppose it can't hurt too much.

(e) Are these new values? Part II - Are they different *values*?

Michael Dummett noted that the adoption of an apparently three-valued logic does not in fact commit you to denying *bivalence*, the claim that there are only two truth values. In particular, just providing the tables with 1, ½ and 0 (or even more values, but we'll stop with three for now) does not show that bivalence fails, unless it is the case that 1 is truth and 0 is falsity. And, Dummett suggests, nothing in the 'three-valued' approach provides support for that last claim. Rather, he argues, within the framework these theorists establish, the most plausible theory holds that sentences with the value 1 are true, and sentences with *any other value* are false. On this picture, all the values other than 1 do not represent separate truth values, but rather represent *ways* of being false. The reason we have the truth-tables is that the theory is no longer truth-functional. The truth value of $\neg A$, or of $A \to B$, does not just depend on which truth-values (i.e. true or false) that A and B take, but on the way they take those values. For instance, if $V(A) = 0$, then A is false and $\neg A$ is true. But if $V(A) = ½$, then A is false but $\neg A$ is also false.

Dummett was writing specifically about the three-valued theory, and in particular about Strawson's use of a three-valued theory in his account of definite descriptions, but the idea carries over to the infinite valued case as well. Adopting one of the eight logics described above still makes for substantial formal differences with classical theory whether we describe the non-1 values as being somehow intermediate between truth and falsity, or as being ways of being false. But how well the theory does at solving problems associated with vagueness *is* crucially affected by this decision on how to describe these values. We now turn to seeing just how well these theories do on that point.

2.3. Problem Solving

So how do these various theories do on the puzzles outlined above. They do fairly well on the Sorites. Different versions of the theory will say different things about what goes wrong in the argument. But they can all say that the reason we think it is a sound argument is that for every step there is some plausible theory, a theory very close to the truth in fact, on which that step is correct. Let's illustrate this on the infinite-valued version of the ukasiewicz logic.

It should be easy to see that on this logic, every premise in the Sorites argument has a truth value of very nearly one. If we take validity to be preservation of the value 1, then the argument is valid, but it is not sound, because the premises are not (quite!) valued 1. If we take validity to be preservation of truth-value, then the argument is invalid, because modus ponens only preserves perfect truth, not truth-value. So either way, we can explain why the argument *feels* valid, and it *feels* like it has true premises, even though one of these feelings in mistaken. In my book, that's a pretty good account.

On the logical anomalies, these accounts get some of the way towards accounting for the logical anomalies. The 'law' of excluded middle is no longer a law, so we are not forced to say that (1) is true. On the other hand, and we will say more about this below, they also deny that other apparent logical truths are truths. The Kleene tables even deny that instances of $A \rightarrow A$ are logical truths, which seems like a good enough reason to reject them in my book. So the verdict here is mixed.

These theories have a little to say on the puzzles concerning vague objects, but it is getting a bit ahead of ourselves to say just what it is. Very briefly, and we will be coming back to this more later in the book, the idea that there are objects that are indeterminately identical seems to be inconsistent with the principle that for any x, y, and any property P, if x is P and y is not, then x is not y. And that principle is (classically!) a consequence of Leibniz's Law, that if x is y and x is P, then y is P. However, this last inference does not work in many-valued logics. More generally, the law of contraposition, that $A \rightarrow B$ entails $\neg B \rightarrow \neg A$ fails in the Kleene logic, but even in the ukasiewicz logic the relevant inference, from $(A \wedge B) \rightarrow C$ to $(\neg C \wedge B) \rightarrow \neg A$ fails. (For a counterinstance, let $V(A) = 0.8$, and $V(B) = V(A) = 0.5$.) So we can accept indeterminate identity and Leibniz's Law if we accept these logics. Some people regard this as a fairly substantial benefit.

We will now look at three substantial problems for this theory. First, we'll be a little more critical of its treatment of the logical anomalies. Secondly, we'll say a little about how conditionals are handled in these logics. Finally, and most importantly, we'll look at the problems this theory has with false precision, and some of the not-so-successful attempts to deal with this problem.

2.4. Anomalies

These many-valued theories tend to say pretty implausible things about contradictions. The problem is that if *A* takes an intermediate truth value, then so will $\neg A$, and the only way a conjunction can be perfectly false is if one or other conjunct is perfectly false. So in these cases $A \wedge \neg A$ will not be perfectly false. So not all contradictions are perfectly false, which is absurd. (Even the dialethicists think all contradictions are false, though they think some of them are true as well.) The point doesn't just arise with direct contradictions like this, but also with inconsistent predicates. This is the point Williamson is making in the following passage.

> The sentences *He is awake* and *He is asleep* are vague. According to the degree theorist, as the former falls in degree of truth, the latter rises. At some point they have the same degree of truth, an intermediate one … the conjunction *He is awake and he is asleep* also has that intermediate degree of truth. But how can that be? Waking and sleep by definition exclude each other. *He is awake and he is asleep* has no chance at all of being true…Since the conjunction in question is clearly incorrect it should not have an intermediate degree of truth…How can an explicit contradiction be true to any degree other than 0? (1994: 136)

Note that this argument could be avoided if we adopted the Dummettian interpretation outlined under question (e) above. If all the 'values' other than 1 are just ways of being false, then there is no intuitive reason that a contradiction must be false in one of these ways rather than another. I assume here that we have no intuitions about ways of falsehood, and provided they do a decent amount off theoretical work, the theorist is entitled to introduce them via any kind of definition she likes. Still, the Dummett move only delays the problem. We now have that *He is awake and he is asleep* is always false, but we don't have that *It is not the case that he is awake and he is asleep* is always true when the negation takes widest scope. And intuitively, it is *very* plausible that this is always true.

We should be careful about how we state the intuition that is being appealed to here. It is *not* the case that our intuitive reaction to any contradiction is that it must be false, and that its negation must be true. The exceptions are instructive: the way they appear, and the ways they can be made to disappear suggest that, despite the existence of these cases, contradictions are always false and their negations are always true. The two most prominent exceptions concern idiomatic expressions, and compound contradictions.

Williamson (1994: 136) notes that we can use expressions like "He is and he isn't" to describe borderline cases. To get a feel for this, imagine the following conversation. "Is Louis bald?" "Well, … he ght push intuitions like this to get someone to feel *He is awake and he is asleep* is true, and hence its negation is not. As Williamson perceptively notes, this tendency can be overcome merely by using different names, or generally different referring devices, to pick out the subject in each conjunct. This is fairly good evidence that we are dealing with an idiomatic usage here. So whatever one thinks about (4), (5) should seem definitely odd.

(4) ?Louis is bald and he is not bald.
(5) ??Louis is bald, and the King of France is not, and Louis is the King of France.

To the extent that one can make sense of (5) at all, it is by assuming that *bald* picks out an intensional property, while the identity clause only implies an extensional equivalence between Louis and the king. This is almost certainly a false assumption, but it seems the most charitable assumption around if one is interpreting (5). One certainly does not hear utterances of (5) as in any way conveying that Louis is a borderline case of baldness, as one might hear the idiomatic, “He is and he isn’t.” And this is the evidence that “He is and he isn’t” is functioning as an idiom - if our commitment to it reflected a commitment to the truth of the proposition it seems to express, then we should also accept sentences that seem to express the same proposition. But the clear fact is that we do not accept these sentences, suggesting that phrases like, “He is and he isn’t” generally communicate something other than their apparent meaning.

There are another class of exceptions to the rule that we always judge contradictions to be false. In some cases of compound contradictions, vagueness can lead to us thinking that the sentence fails to be false. So (6) can sound not too bad if you are in a favourable mood.

(6) It is not the case that Louis is bald, but nor is it the case that Louis is not bald.

We won’t deal with this here, but there are some arguments that our judgements here are sensitive to pragmatic, rather than semantic, features of the sentence. We will come back to this in later parts of the discussion. In any case, even if (6) failed to be false, this is no help for the many-valued theorist, because their theory still looks to be in trouble because of its false predictions concerning simple contradictions.

Michael Tye attempted to defend the many-valued theorist from this objection. Tye’s theory is based on the 3-valued Kleene tables, so he has $A \rightarrow A$ not always being true, and $A \wedge \neg A$ not always being false. But, he notes, the first of these is never false, so it is a ‘quasi-tautology’. A sentence is a ‘quasi-tautology’ iff it is false on no line of the truth table. And $A \wedge \neg A$ is never false, so it is a ‘quasi-contradiction’, which by definition is a sentence that is never false. Tye suggests that our feelings that $A \rightarrow A$ is a logical truth and $A \wedge \neg A$ a logical falsehood can be explained by the fact that they are quasi-tautologies and quasi-contradictories, respectively. This *might* get the many-valued theorist out of this problem, though I have my doubts about the plausibility of the explanation. A bigger problem is that Tye’s theory predicts that we accept all instances of the law of excluded middle, since it is a quasi-tautology as well. If we do accept this, then we have little reason to adopt the Kleene tables. If we don’t, then the ‘quasi’ theory just changes the false prediction that Tye’s theory makes.

2.5. Conditionals

There hasn’t been as much attention paid to conditionals as to contradictions in the literature on these logics. Still, I think it is worth briefly noting the problems the theories have with conditionals. It seems to me as safe a point as any in logic that all instances of (7) are logical truths, and all instances of (8) are equivalent to the matching instance of (9).

(7) $A \rightarrow A$
(8) $(A \wedge B) \rightarrow C$
(9) $A \rightarrow (B \rightarrow C)$

We already noted that the Kleene tables do not validate (7), so they seem inappropriate already. They do better with (8) and (9) - these really are equivalent on the Kleene tables. The ukasiewicz tables provide the opposite answers. On that theory, (7) is a logical truth. And on those theories (8) entails (9), no matter which of the accounts of entailments noted above we adopt. But (9) does not entail (8), which seems rather surprising. For a counterexample, let $V(A) = 0.8$, $V(B) = 0.6$ and $V(C) = 0.4$. Then the truth value of (9) is 1, and the truth value of (8) is 0.8. I'll leave it to you to decide just how embarrassing this is.

2.6. False Precision

The point of adopting these intermediate truth values is so we can say that, in some good sense, there is no fact of the matter whether, say, Louis is bald. And whatever problems end up accruing as we develop the theory, this doesn't seem like an ignoble motivation. How strange then, to try to solve it by positing facts about the degree to which Louis is bald. There are several ways to state the problem that we are facing here, and I think they all come down to pretty much the same point. What the various versions of the problem all have in common is that the motivation we have for adopting the many-valued approach at the first order seems to provide a motivation for adopting at the second-level.

To make this concrete, reflect for a minute on why you don't find it plausible that there is some hidden, but sharp, boundary between, say, the places that satisfy the predicate 'around here' and those that do not. There are a few reasons you might think this. (If, indeed, you do!) You might think that if such a boundary existed, someone would know about it, but no one does. You might think that the conventions that establish linguistic meaning are not detailed enough to create such a boundary. You might have some combination of these reasons: that even if there are unknowable things (like the exact location and momentum of a particular electron), anything established by convention, like a meaning fact, is knowable. Now consider a new predicate 'definitely around here', that is satisfied by all and only objects *x* such that if *n* is a name for *x*, then ⌜*n* is around here⌝ is true to degree 1. Your reason for denying that there is some hidden, but sharp, boundary between the places that satisfy the predicate 'around here' and those that do not is also a reason for denying there is some hidden, but sharp, boundary between the places that satisfy the predicate 'definitely around here' and those that do not. But the many-valued theorist, as their theory now stands, is committed to the existence of this boundary. We will spell out this commitment in a couple of ways.

(a) Excluded Middle

Every sentence has a truth value. That truth value either is 1 or it is not 1. So for any *x*, either *x* is definitely around here, if the truth value of ⌜*n* is around here⌝ is 1, or *x* isn't definitely around here, if that truth value is not 1. But there is no more reason to believe in this version of the law of excluded middle than in the more traditional version that the many-valued theorist rejects.

(b) Sorites

Start with an ordinary sorites argument for 'definitely around here'. For example, any point that's within efinitely around here. If *x* is definitely around here, and *y* is within, say 1 inch of *x*, then *y* is definitely around here. But that implies that Tianamen Square is definitely around here, which is absurd. So one of the sorites premises here cannot be correct. It is easily provable that when we are working in a language where all of the atomic sentences take traditional truth values, these new logics all endorse classical reasoning. But the only predicates we've used are 'within 1 inch of' and 'definitely around here', and any atomic sentence formed using those predicates takes a traditional truth value. So if one of the sorites premises is incorrect, that implies (because it classically implies) that there are two points 1 inch apart such that one is definitely around here and the other is not.

As the old saying goes, things aren't as bad as they seem for the many-valued theorist. They're much much worse. There isn't just a hidden but sharp border between the points *x* such that ⌜*n* is around here⌝ is true to degree 1 and those where it isn't. (As above, we use *n* as a name for the object *x*.) For any *r* in [0, 1] there is a sharp but hidden boundary between those points *x* such that ⌜*n* is around here⌝ is true to degree *r* and those where it isn't. The problem we started with was that applying classical logic to the sorites cases forced us to admit the existence of a hidden but sharp boundary, and we didn't want to do that. So we gave up classical logic, with all of its virtues, and in exchange we got rid of one sharp boundary and replaced it with continuum many. I'm not seeing how this is a good trade. To be fair, only the theorist who accepts infinitely many truth values has bought continuum many sharp boundaries. The three-valued theorist has only bought two: loosely speaking, one between the true and the indeterminate, and another between the indeterminate and the false.

For another perspective on the problem, consider again what the many-valued theorist says about a particular penumbral case, say of a light that is somewhere between being red and being orange. The intuition that the many-valued theorist is trying to capture is that it is indeterminate whether the light ray is red. Now one way to capture this might by introducing some indeterminacy in whether the sentence *That is red* is true. But we shouldn't try and model this kind of indeterminacy by introducing a new kind of determinate fact, such as *It is true to degree 0.314 that that is red*. If there is no fact of the matter about whether the thing is red or not, there is no fact of the matter about the degree to which it is true that the thing is red.

The only way out of this is to retreat from the numbers. One option is to say that the numbers are really ordinal values not cardinal values. The idea is that there isn't any fact about what the truth value of *That is red* is, just some comparative facts about how its truth value compares to other truth values. Formally, we might introduce a new connective $_T$ as primitive, so p $_T$ q means, roughly, p is true to a higher degree than q. But this is only rough, because really the comparative notion is all we have, and the degrees are only heuristic guides to the behaviour of the connective.

Within any of the frameworks considered above, this idea is mistaken twice over. First, if we assume what Keefe calls a connectedness principle for this connective, then we are still committed to ridiculously many sharp but hidden boundaries. Secondly, if we want to keep anything like the above logics, there will only be one cardinal scale compatible with any complete ordinal scale, so we will not really have avoided the precise degrees.

Note that on all of the theories above, for any two sentences p and q, we have p $_T$ q or q $_T$ p, and, sort of as a corollary, we have p $_T$ q or $\neg(p$ $_T$ $q)$. (The meta-language is classical, after all.) Let q be any subject-predicate sentence where the subject is a penumbral case of a satisfier of the predicate. And, for convenience, assume the predicate is not a colour predicate. Now for any sentence of the form *Light rays of wavelength x nm are red*, this sentence will be either truer than q, or exactly as true as q, or less true than q. And it either will be, or will not be, truer than q. Say that a number x is F_q iff ⌜Light rays of wavelength x nm are red⌝ is truer than q. The reasoning considered just above seems to show that there is a sharp boundary between those numbers that are F_q and those that are not, *for any q at all*. So now there are only as many sharp but hidden boundaries as there are propositions. This project of removing the sharp but hidden boundaries is going smashingly well.

The retreat from cardinal scales (where we assign numbers to things) to ordinal scales (where we just rank things) only buys us something if there are many cardinal scales that can be associated with the ordinal scale. In perhaps the best known use of ordinal scales, for measuring preferences, this is true. A consistent preference ordering only determines an assignment of utilities to goods up to an affine transformation. That is, if the function f from goods to utilities is such that the function assigns a higher number to x than y iff x is preferred to y, then the function g, where $g(x) = ax + b$ (for $a > 0$ and b any real), also assigns a higher number to x than y iff x is preferred to y. So if we just start with the preference ordering, there is no reason to think f is *the* function from numbers to utilities than to think *g* is the function. Which is to say, there is no reason to think there are such things as numerical utilities, because a preference ordering that does not determine an assignment of numerical utilities does all the explanatory work that we need. Sad to say, the same happy story does not go through on the many-valued accounts in question. Specifically, there cannot be *distinct* functions *f* and *g* with the following properties (the proof is in Keefe 2000: 133):

(a1) $f(A) \quad f(B)$ iff $A \quad _T B$
(a2) $g(A) \quad g(B)$ iff $A \quad _T B$
(b) $f(A \rightarrow A) = g(A \rightarrow A) = 1$
(c) $f(\neg(A \rightarrow A)) = g(\neg(A \rightarrow A)) = 0$
(d1) $f(\neg A) = 1 - f(A)$
(d2) $g(\neg A) = 1 - g(A)$
(e1) $f(A \rightarrow B) = \min(1, 1 - f(A) + f(B))$
(e2) $g(A \rightarrow B) = \min(1, 1 - g(A) + g(B))$

That is, if *f* and *g* both respect the ukasiewicz theory of the truth values of compound sentences, and both assign 1 to logical truths and 0 to logical falsehoods, they must be identical. The last qualification seems fairly weak here - if the only way we can get away from degrees of truth is by denying that degrees fall in [0, 1], then it seems we haven't really gained very much. Keefe notes that on the Kleene definition of the conditional, there can be distinct *f* and *g* that satisfy the relevant constraints. But even there, there must be some constraints. In particular, we still have $f(A) > 0.5$ iff $g(A) > 0.5$. This is because $f(A) > 0.5$ iff $A >_T \neg A$ iff $g(A) > 0.5$. So there is still a sharp line between those sentences that are true to degree greater than 0.5 and those that are not.

You might think, after all this, that the problem was that we kept classical logic in the metalanguage. It was because we did this that we had sorites paradoxes around concepts like *Determinately around here* without a many-valued logic to fall back upon. If only the many-valued logic applied in the meta-language, maybe all these problems would go away. This position has been advanced, and even defended, by Michael Tye.

Unfortunately, as Williamson and Keefe have pointed out, the position is either seriously incomplete or completely incoherent. To see why this is so, we just need to reflect on how we normally read truth tables. Imagine I had given you a truth table for the four connectives that looked like this:

A	$\neg A$
T	F

$A \wedge B$	T
T	T

$A \vee B$	T
T	T

$A \rightarrow B$	T
T	T

This wouldn't tell you a whole lot about the logic that was to be developed. And it wouldn't do so because I haven't covered all the cases. For truth tables to generate a unique logic, they must be complete. So for a three-valued truth table (of the kind Tye uses to generate his preferred logic of vagueness) to generate a unique logic, it must be the case that every sentence takes one of these three truth values. And, at the very least, that means that for any sentence, the disjunction: This sentence takes value 1, or value ½, or value 0, must be true. But, if it can be indefinite which of these three values some particular sentence takes, as seems to be the suggestion, and if a disjunction is only true if one of the disjuncts is true, as all the many-valued theories assert, then this particular disjunction is false. So Tye's suggestion is dramatically incomplete.

Actually, things are worse than that. It is not altogether clear just how Tye's suggestion could be completed. Any truth table based account will require the truth of some disjunction of the form: all sentences take this truth value, or this one or… And any account that looks like the accounts above will say that no disjunction can be true unless one or other disjunct is true. So accounts based on such truth-tables *must* accept that for any sentence, something of the form 'The truth value of this sentence is thus-and-so' will be true. This seems to doom any possibility of accounting for higher-order vagueness within these logics. And that in turn seems to doom these accounts.