UNIVERSITY OF CALIFORNIA

Los Angeles

The Nature and Logic of Vagueness

A dissertation submitted in partial satisfaction of the

requirements for the degree Doctor of Philosophy

in Philosophy

by

Paul Raymond Hovda

2001

The dissertation of Paul Raymond Hovda is approved.

Andrew Hsü

Edward Keenan

Donald A. Martin

Calvin Normore

Terence Parsons

David Kaplan, Chair

Universtity of California, Los Angeles

2001

Dedication

I dedicate my dissertation to my parents and step-parents, whose support has been unwavering and invaluable, and to three teachers who shaped me: George Bealer, Marian Keane, and Anne Moore.

Contents

Acknowledgments

For discussions of the subject and useful advice and criticism, I thank: Joseph Almog, Ben Caplan, Kit Fine, Andrew Hsü, Keith Kaiser, Youichi Matsusaka, Terry Parsons, Mark Rubin, Dominik Sklenar, especially Tony Martin, and most of all, David Kaplan.

VITA

April 28, 1970	Born, Frankfurt am Main, Germany
1992	B.A., Mathematics and Philosophy *summa cum laude*, Philosophy University of Colorado Boulder, Colorado
1993–1999	Teaching Assistant Department of Philosophy University of California, Los Angeles
1999	M.A., C. Phil., Philosophy University of California, Los Angeles
1999–2000	Instructor Department of Philosophy Occidental College Los Angeles, California
2000	Instructor Department of Philosophy University of California, Los Angeles

ABSTRACT OF THE DISSERTATION

The Nature and Logic of Vagueness

by

Paul Raymond Hovda

Doctor of Philosophy in Philosophy

University of California, Los Angeles, 2001

Professor David Kaplan, Chair

The dissertation considers both metaphysical and logical issues related to the vagueness of natural language. The principle metaphysical claim is that the vagueness of language is, at least in some cases, a direct result of indeterminacy in the subject matter of the language, rather than any sort of flaw of the language. A limited defense of this claim is given, as well as criticism of alternative views.

A number of logical issues are addressed. First, the relationship between the notion of determinacy and the idea of an unsharp line is considered, and it is suggested that the relationship is not as simple as it may at first seem, and that the idea of the unsharp line may be irreducible. Next, it is urged that there is a methodological fork in the road for the systematic treatment of vague language, including formal semantics. On one path, we accept certain intuitively puzzling propositions, exemplified by "This is red or it is not the case that this is red, though it is indeterminate which." On the other path, we reject classical principles of reasoning in our own reasoning both in and about vague language. Some limited arguments are given for taking the former path, and a formal system relevant to this path is motivated and discussed. Finally, both the metaphysical and logical work of the prior parts of the dissertation are brought to bear on the subject of the indeterminacy of identity.

Chapter 1

The Nature of Vagueness

This chapter considers the major philosophical perspectives on the nature of vagueness and gives arguments in favor of the view that vagueness in language sometimes arises from indeterminacy in the world.

1.1 Vagueness and its ubiquity

Vagueness manifests itself to us primarily in the feeling that arises when we consider questions about "where to draw the line" between things of a sort and things not of the sort. The famous *sorites* paradox and its most familiar variations reveal that we are liable to accept propositions which seem to entail a contradiction, but their deeper significance is that they indirectly raise the question “Where to draw the line?” and invite us to attempt to account for the peculiar intellectual feeling that there is no correct place to draw the line, that there is something wrong with drawing the line anywhere. A satisfying way of accounting for this feeling is a central part of a philosophical account of vagueness.

Though *sorites* paradoxes bring vagueness to our attention in an especially vivid way, the vagueness of a predicate can be seen without invoking any of the sort of general principles used in a typical *sorites* argument. We need only imagine a clear satisfier of the predicate, a clear non-satisfier, and imagine a series of things connecting the two, each member of which is ever so slightly different from the last. When we can imagine such a series, we will often have the sense that there is no place in the series to correctly draw the line between the satisfiers and the non-satisfiers, and thus we will see that the predicate is vague. So the existence of an easily stated principle like “For all x, if x is a heap of sand, then the object which results from removing one grain of sand from x is also a heap” is not essential to the vagueness of a predicate. What seems to be the heart of the matter is the *lack of sharp boundary*, and the resulting possibility of situations in which it is indeterminate whether the predicate is satisfied. This possibility of indeterminacy entails the possibility that there be cases of which it is problematic either to judge that the predicate is satisfied or to judge that it is not. The makings of this possibility exist wherever we can see an appropriate series of situations with extremely slight differences between any adjacent two, but with significant differences between the two at either end. The kinds of slight difference that are appropriate will vary from predicate to predicate, and for some cases there will be many ways to connect up a clear satisfier and a clear non-satisfier; there are many paths through the space of possibility, as it were, each step along which looks very relevantly similar to the last.[1] I will call such a series a "*sorites* series" below, but this should not be understood to connote an essential connection to a *sorites*-like argument.

Before discussing the various philosophical perspectives on vagueness, the ubiquity of the phenomenon of vagueness should be properly appreciated. Nearly all ordinary predicates are vague. With a little imagination, we can construct, for almost any one, a *sorites* series of objects or situations. So let us consider some examples. It is fairly commonplace that "bald", "red", and "short" are vague; I will discuss some others.

Consider "is alive". When we get down to milliseconds, it seems that there

[1]The path may proceed along some easily identifiable "dimension" of variation (as for a typical series for the predicate "is tall"), or it may not, as it might not for predicates like "clever" as applied to people or "built" as applied to buildings under construction. Vagueness is not merely the result of the existence of an easily identifiable *dimension* of fine-grained variation.

is vagueness about which second is that last or first second at which a typical organism is alive. Moreover there is probably vagueness about whether certain kinds of things, for example viruses, are ever alive.

Consider "is occurring". It is vague exactly when the party begins and ends. It is vague which day was the first day of the drought. And similar things may be truly said of just about any event or process. And this is independent of the fine-scale structure of time, for no matter how that goes, we can break time into small enough chunks to raise the relevant problems. An example due to Essenin-Volpin shows this: which of your heartbeats was the last heartbeat to occur when you were still a child? Childhood has a vague ending.

Consider "is a tiger". It cannot be true that every tiger has a tiger for a mother, for there have only been finitely many tigers, and there are no ancestral circles. But it is quite possible (though by no means certain) that even if we had been carefully observing the process of the evolution of that species, we would be unable to happily assert of any one animal that it was a tiger and its mother was not, and we would feel that this is not due to a lack of knowledge of the facts of tigers on our parts, but is rather a result of vagueness.

And here is an assortment of others: "speaks English", "is a sandwich", “loves Mary”, “lives in Los Angeles”, and “is a dangerous household chemical”.

I invite the reader to become convinced, if he or she is not already, that there is, or at least may well be, vagueness connected with most ordinary predicates. I do this not only to emphasize the urgency of the problems of vagueness, but also to remind us that vagueness is a vast, far-reaching phenomenon, and hence we should be careful to test our solutions on a variety of cases.

I will advance the doctrine that vagueness in language can spring from indeterminacy in the world. But I want to note here that I am prepared to find that in some cases, there is worldly indeterminacy lying behind the vagueness of some language, while in other cases not. For example, while the vagueness of “many”, or even “is tall”, may seem to have a flavor of non-worldliness or mere conventionality to it, the vagueness of “is a tiger”, on my model, will be the intrinsic vagueness of a natural kind. And in sections 1.4 and 1.5, I will consider two very different sorts of worldly causes of vagueness in language: smooth gradients and genuine indeterminacy.

1.2 General perspectives on vagueness

There are three general perspectives on vagueness: the epistemic, the semantic, and the worldly conceptions of vagueness. They identify three putative sources of vagueness: deficiency in our knowledge, the character of language, or indeterminacy in the world, respectively. To illustrate, consider a borderline case for the application of a predicate. The epistemic view is that in a borderline case there is a fact which we do not know, and hence we cannot say whether the predicate is satisfied or not. On this view, the predicate is either satisfied in virtue of this unknown fact, or is not satisfied, in virtue of this unknown fact. The semantic view is that the ultimate explanation of the borderlinearity of a borderline case *essentially* has to do with the predicate involved; the borderlinearity is ultimately explained as the semantic character of the predicate. The worldly conception of a borderline case holds that the world is lacking a robust, determining fact, where one is wanted by the predicate, and hence the predicate is neither determined to be satisfied or determined not to be satisfied. The three views offer three explanations of the apparent fact that we cannot correctly, with good justification, say that the predicate is satisfied or that it is not.

1.2.1 The epistemic view

According to the epistemic view of the nature of vagueness, the vagueness of a predicate consists in our being ignorant of certain facts. For example, the vagueness of “is a red tile in the series” is supposed to consist in our not knowing, for certain objects, whether those objects do or do not satisfy “is a red tile in the series”. The ignorance is supposed to be of a special kind: it is not that we are ignorant because we cannot see the tiles well enough, or are blind, nor because we do not understand the physics of light well enough, or anything like that.

The central commitment of the epistemic view

It should be noted that on the worldly and semantical views of the nature of vagueness, the vagueness of “is a red tile in the series” entails that there are objects about which we do not know whether or not they satisfy “is a red tile in the series”, in the sense that we know neither that the object in the case satisfies “is a red tile in the series” nor that the object in the case satisfies “is not a red tile in the series”. (And thus, there is an object about which we do not know that it is a red tile in the series, nor that it is not a red tile in the series.) The distinctive mark of the epistemic view is that vagueness *consists in* this ignorance. If α is an object that we cannot know to be a red tile in the series, and cannot know to be not a red tile in the series, and β is an object which is perfectly obviously, paradigmatically, unexceptionably, a red tile in the series, then the epistemic view is committed to the following:

> There is a fact of the matter as to whether α is a red tile in the series in just as robust a sense as there is a fact of the matter that β is a red tile in the series. There is no way in which this fact about α is metaphysically inferior to the fact about β.

If the fact about α is instead supposed to be metaphysically second-rate in some important way, and unknowable as a result, then we do not have an epistemic view of the *nature* of vagueness. The view advocated in this dissertation has it that while either α is a red tile in the series or not, there is no metaphysically robust fact of the matter that α is a red tile in the series, and no robust fact that α is not a red tile in the series. There is an ultra-thin sense in which there is a fact of the matter: either it is true that α is a red tile in the series or it is true that α is not a red tile in the series. This view is not an epistemic conception of the nature of vagueness, because it holds that the *reason* the fact about α cannot be known is that there is no robust fact that determines that α is a red tile in the series, and no robust fact that determines that α is not a red tile in the series.

There is no good reason to believe the epistemic view

The most obvious problem with the epistemic view is that it seems implausible on its face. It is far-fetched that there is some unknown sharp boundary between those items which are red and those that are not red. Why does this suggestion seem so far-fetched? In part because familiar cases of ignorance seem so different from borderline cases. Vagueness would have to be a *sui generis* type of ignorance. And it would have to be such an odd genus of ignorance that one doubts that it is in fact a sort of ignorance.

Timothy Williamson attempts to give a satisfying account of vagueness as ignorance in his book *Vagueness* [31]. Roughly put, Williamson's central ideas are as follows.

First, Williamson gives logical/semantical arguments aimed to show that bivalence holds, so that every sentence is true or false. We will not discuss the details of these arguments here. I give an argument similar in spirit to Williamson's in 3.1. While I find the arguments very plausible, they are far from incontrovertible, and as I suggest in 3.1, one can reject them if one is willing to reject some intuitively correct steps of reasoning. One then accepts the puzzling proposition that some of these intuitively correct steps are incorrect; but one can thus avoid other puzzling propositions that cannot be avoided if bivalence is accepted. It is safe to say that Williamson's logical/semantical arguments are unlikely to convince those who believe in truth-value gaps.

Williamson takes it that if a vague language is bivalent, then its vagueness must be epistemic. He goes on to attempt to explain the cause of our ignorance of truth-values in borderline cases.

Williamson's explanation is roughly as follows: The extension of a vague predicate is fixed by linguistic practice, and this fixing is sensitive to extremely subtle facts about the practice. Since we do not know these subtle facts, we do not know which of various similar candidates for the extension of "is red", for example, is its actual extension. Moreover, had practice been slightly different, things would have seemed (to one of us) exactly the same, and yet the extension of the predicate would have been different. Williamson holds that this further fact shows that we cannot (without a radical change of cognitive power) know of any one set that it is the extension of "is red". He suggests that the situation here is similar to the situation with a more familiar sort of ignorance connected with "margin for error" principles. If one is in a stadium full of people, one cannot know, just by looking with the naked eye, that there are exactly 17,306 people in the stadium in part because had there been just one more person in the stadium, things would have seemed exactly as they do. We can only know that there are 17,306 people in the stadium if our means of knowing does not have a "margin for error". Similarly, since our knowing of some set S that it is the extension of "is red" is incompatible with practice's being slightly different than it in fact is, and yet things would have seemed precisely the same to us had practice been slightly different, we cannot know, even if we can truly believe, that S is the extension of "is red".

There are a number of problems with Williamson's account, but the one that I think is most important is that Williamson's idea, that linguistic practice in fact fixes an exact, sharp extension (in the classical sense) for the predicate "is red", is no more plausible an idea than the idea that there is a hidden sharp boundary between those things which are red and those which are not. Williamson's account does help to show how vagueness generally could be a form of ignorance, but only by assuming that the indeterminacy in the connection between linguistic practice and the details of the semantic features of a language is *itself* an epistemic indeterminacy. In effect, Williamson's account of vagueness, generally, as ignorance presupposes that vagueness, in the connection between predicate and extension, is ignorance; it therefore cannot help convince us that vagueness is a kind of ignorance. It looks as if the epistemic view of vagueness cannot overcome the charge of simple implausibility. Let me explain.

Consider the connection between the linguistic practice concerning the predicate "is red" and the issues, for each tile in a series of 10,000 tiles very gradually ranging from red to green, whether that tile satisfies "is red". I believe the most intuitive thing to say is that the facts about practice determine that the first tile satisfies the predicate, and they determine that the last tile does not, but that there are very likely tiles for which practice does not determine whether the tile satisfies the predicate. Williamson asserts that practice *does* determine, for each tile, whether it satisfies the predicate; practice determines a *sharp line* between those things that satisfy and those things that do not satisfy the predicate.

When one considers for a moment what actual linguistic practice is like, Williamson's proposal seems hard to believe and unmotivated. It is almost as if one adopted the following bald epistemicist picture: the predicate "is red" expresses a property with sharp boundaries, and we simply do not know what those sharp boundaries are. To the extent that one is convinced by the arguments for bivalence, one might be tempted to say "There *must* be such a property!"; Williamson seems to have looked at these arguments and said "Practice must fix a sharp boundary." The suggestion that the determination of the sharp boundary comes from linguistic practice is more *testable* than the suggestion that the sharp boundary comes from a property that the predicate expresses, but when one considers the facts, they seem to tell against

Williamson. One can imagine a linguistic practice much different from the actual one, in which some mechanism comes pretty close to determining such a sharp boundary. If everyone followed the practice of not allowing a predicate to enter his personal vocabulary until it's truth conditions had been set out with the most extreme precision possible by a well funded linguistic legislature, and if everyone always used language only according to the linguistic law set out by the legislature, and if the legislature did a very good job, then perhaps practice would fix a precise boundary for the predicate.

But our actual practice is not currently so orderly. When one considers for a moment how one might go about trying to discern the location of the sharp boundary Williamson supposes is drawn by the practice governing "is red", by polls, questionnaires, oral examinations, brain scans, or what have you, one realizes that it is extremely implausible that the facts of practice determine a sharp boundary. Consider a reddish/brownish borderline case. A big *simplification* of the apparently relevant facts about practice might look like this: 80% of people will call a tile of that color "red" in some trials, 80% will call it "not red" in some trials, 40% will call it "red" more than 75% of the time, 38% will call it "not red" more than 80% of the time, 10% will call it "red" 95% of the time, 3% will call it "not red" 99% of the time, etc. It seems quite unlikely that facts like these determine a sharp boundary for "is red". The bald epistemic view, that the predicate gets its semantic power by expressing a property which establishes sharp boundaries for it, is probably more plausible.

1.2.2 The worldly and semantic views

To get a better feel for these two sorts of accounts of vagueness, let us fill them in a bit by considering how they might account for the vagueness of a vague predicate. For illustration, we will make our example artificially simple. We will suppose that for our predicate—"is alive"—there are, determinately, only three sorts of things: things which determinately satisfy it, things which determinately dissatisfy it, and things of which it is indeterminate. There is no further or higher indeterminacy about this predicate.[2] Let us suppose that the things of which it is indeterminate are the viruses.

An obvious way the vagueness of "is alive" might be thought to have a worldly nature would be as follows: the vagueness of the predicate is supposed to be explained by a more basic and irreducible indeterminacy in the property of being alive. The predicate expresses that property, and gets its semantic power through this attachment to the property. Thus, where the property is

[2] In section 2.1.2 and throughout chapter 2.1 we will discuss the fact that our sense of the lack of a sharp boundary of a vague predicate involves what has been called "higher-order" vagueness.

indeterminate—where the property is not determined to be had nor not to be had—the predicate is vague.

There are two importantly different models of the semantic conception of vagueness: reductive and non-reductive. A reductive idea, appropriate to our simplified example, is arrived at by revising the simple model on which a predicate gets its representational power by being attached to a property. The reasoning at work is something like this: there is no indeterminacy about which things have or lack a property, and hence a vague predicate cannot fit the simple model, but must have a different representational connection to the world. One obvious picture of these connections that might account for the phenomena has it that the vague predicate is tied to *two* properties instead of one. The properties themselves are perfectly precise, and there is no indeterminacy in the connection between predicate in world; it is just that the semantic character of the predicate is not as simple as the simple model would have it.

In the example, the predicate is supposed to be, as it were, *hitched to* two properties instead of to a single property. The one property is determinately had by viruses and determinately living things, and determinately lacked by everything else; the other property is determinately had by determinately living things and is determinately lacked by the viruses and everything else.

To embellish the hitching metaphor, imagine a carriage hitched to two horses instead of one. When the horses go in the same direction, the carriage moves as if it were hitched to just one horse. But when the horses disagree, and go in opposite directions, the carriage does not move. Similarly, on our model, where the properties agree—where an object either has both or lacks both—the predicate is "moved" either to truth or to falsity; but where the properties disagree—where an object has one but not the other—the predicate does not determinately move to truth or falsity. So the fact that it is vague whether viruses are alive is explained by the non-vague, semantic connection of the predicate "is alive" to non-vague properties. The deviations in the behavior of the carriage from that expected of a one-horse carriage is not explained by an unexpected kind of horse, but by a different mechanism of connection between carriage and horses.

I call this model "reductive" because the vagueness of the predicate is supposed to be the result of something which does not involve any vagueness or indeterminacy: just the unexceptionably precise and determinate, if complicated, workings of semantic machinery. The model is in keeping with the view that the world itself is neither "vague" nor characterized by any sorts of indeterminacy that might be called "vagueness". On such a view it might even be a category error to call the world "precise" or "determinate", and these adjectives would be properly applicable only to language or other sorts of representation. Yet language is part of the world, and so the vagueness of language must ultimately consist in some phenomena that are in principle precisely describable, and which are not themselves in any way indeterminate.

There is another sort of model in which vagueness is primarily linguistic or representational rather than worldly, and yet which does not *reduce* vagueness to a special sort of non-vague, non-indeterminate mechanism. On this model, the vagueness of language is explained by the indeterminacy of the connection between language and the (rest of the) world. To illustrate: In our simplified example, "is alive" might be supposed to have its extension determined by the language users' dispositions regarding the word, in such a way that the predicate's being satisfied by a thing amounts to its users' being disposed to assert the predicate of the thing, in certain conditions. Now if it is intrinsically "vague" or indeterminate just what the dispositions of the language users in this connection are, then the predicate will be vague. The vagueness of the predicate arises from indeterminacy in the linguistic dispositions of language users, *not* from the indeterminacy of some entirely language-independent property that might be called "the property of being alive". Yet the vagueness of the language is not explained by something which does not involve any indeterminacy, since, in our case, it is the intrinsic indeterminacy of the dispositions of language users that explains the vagueness of the predicate. Thus the account is non-reductive.[3]

The non-reductive linguistic account and the worldly account share a notion of irreducible worldly indeterminacy: its being indeterminate whether so and so, is, at least in some cases, a brute fact, and not a fact about the representation of so and so. On the linguistic account, this indeterminacy is found only in the connections among language, language users, and the world. It might be called "worldly vagueness", since it is indeterminacy which is not itself explained as the workings of a mechanism for representing the subject matter in which it is found. The subject matter is the character of the linguistic dispositions or conventions of language users, but the vagueness in our description of *it* is not explained by the semantic workings of the (meta-) language which describes this subject matter; rather it is taken as an inherent feature of the subject matter. Supposing that "is red" is a predicate the account of which essentially involves some reference to us and our reactions to things, we can look at the non-reductive linguistic account of the vagueness of "is red" as follows: it turns out that the property of being red is not an *independent* worldly property, but rather is or depends upon the property of

[3]Delia Graff's suggestion that vagueness in language is (partially) a result of vagueness in our interests and intentions may be an example of a such a conception of vagueness. See [9].

being suitably related to us; thus the vagueness of being red is or depends upon the vagueness of being suitably related to us. So for my purposes, the non-reductive semantic account can be considered a sub-species of the worldly account, distinguished by its selection of a special sort of vague property as the semantic correlate of a predicate: a mind- or language-dependent property.

The real divide between semantic and worldly views of vagueness is between the reductive semantic account and the other two. Thus in what follows, when I speak of the "semantic account", I will have the reductive account in mind. The non-reductive semantic conception is of independent interest, but, as I suggest in section 1.3, there seems to be as much *prima facie* evidence that there is irreducible vagueness in the rest of the world as that there is irreducible vagueness in the relation between language and the world, or in the intentions and interests that mediate this relation.

1.2.3 Vagueness as context-dependence

Another sort of view of the nature of vagueness on which vagueness has a linguistic or semantic nature is the view that vagueness is a sort of context-dependence. There are various ways in which the view can be worked out. See for example [4], [9], [12], and [1]. I will not examine any of these views in detail, for it seems to me that vagueness remains after the context is as fixed as possible, and hence vagueness is not context-dependence.[4] It is well known that many adjectives are context-dependent. Whether something counts as "small", for example, is context-dependent: the same object is both small (for a dog) and not small (for a Chihuahua).

Different contexts might yield different standards for what should count as red, for example, but fix a typical context and there is still no sharp boundary. And it is typical that vagueness remains after the context is fixed. Moreover, there are predicates which do not seem context-dependent but which are probably vague in the relevant sense of lacking sharp boundaries: "member of the species *homo sapiens*" for example.

This is not to dismiss all of the work of those authors who have pursued conceptions of vagueness as context-dependence; indeed much of it may tell us some interesting things about the resolution of truth conditions by context. And we need not deny that the lack of contextual resolution accounts for a sort of linguistic vagueness. It is just that there are other sources of vagueness in language, including, as I will argue, indeterminacy in the non-linguistic world

[4]It might be objected that this does not show that vagueness is not context-dependence, for it might be asserted that it is not possible (for us) to fix the context well enough to get rid of vagueness. That is, it might be thought that our resources for fixing or determining the context are not great enough to fix or *determine* (completely) the context. Because it is a lack of determinacy that is the explanatory anchor, this looks like another species of non-reductivist account. But the required notion of "context" needs to be made out.

itself.

1.3 *Prima facie* evidence for the worldly view

In this section I will give some relatively simple arguments for the view that vagueness is as much a worldly phenomenon as a linguistic one.

1.3.1 A quick look at the opposition

In *The Plurality of Worlds* David Lewis writes

> The only intelligible account of vagueness locates it in our thought and language. The reason it's vague where the outback begins is not that there's this thing, the outback, with imprecise borders; rather there are many things, with different borders, and nobody has been fool enough to try to enforce a choice of one of them as the official referent of the word 'outback'. Vagueness is semantic indecision.[5]

This seems to give a pretty clear picture of the view I oppose. But its clarity is not perfect. Let us consider it carefully. The first sentence does little more than boldly assert a thesis.

The first clause of the second sentence rejects just the sort of view I would like to defend: There really is this thing, the outback, and it's a real fact about it that it has imprecise borders. Lewis says that the existence of such a thing

[5] [17] p. 212.

is not the reason why it is vague where the outback begins. Let me point out now that this thought of his is difficult to make out; upon reflection, the thought seems dangerously close to being of the form: P does not explain P.

"It's vague where the outback begins", I say, and Lewis would agree. But then he would add that this is not explained by there being something, the outback, with imprecise borders. But I would ask what difference there is supposed to be between its being vague where the outback begins and there being a thing, the outback, with imprecise borders. Technical maneuvers involving such things as the *de re* / *de dicto* distinction come to mind, but before a philosophical discussion of those matters gets started, I believe it should be noted that on the face of it

> It is vague where the outback begins.

seems to be very close to, if not just another way of putting

> There is this thing, the outback, with imprecise borders.

The second clause of the second sentence gives Lewis' preferred explanation of the fact that it is vague where the outback begins: there are many (almost coincident) things and nobody has tried to enforce a choice of one of them as the official referent of "outback". The third sentence gives a general statement of the view, and suggests that what's true in the case of the outback is true in other cases of vagueness.[6]

Let us think carefully about Lewis' suggestion that vagueness in language is simply a result of "semantic indecision". Reasonable as it may sound at first, I believe that reflection will show it to be a rather strange view.

1.3.2 The direction of explanation

Do changes in our use of language change the world? Of course they do, for language is part of the world. But they do not typically change the part of the world that the language is *about*. And typically the way we use language does not determine the way the world is, except insofar as the language itself is part of the world. I think it far from obvious that vagueness is exceptional in this regard; rather, it looks not to be.

The ripe fruit of an orange tree is (normally) orange. That it is orange has nothing at all to do with how we happened to use the word "orange". Had someone made the official meaning of "orange" be the same as the current meaning of "red", this wouldn't have had any effect at all on the fact that those fruits are orange. So much is as elementary as the old saw: If we were

---

[6]It is worth noting that I could, if only for the sake of argument agree with Lewis that there are a bunch of precisely bounded things no one of which has been enforced to be the official referent of "outback". Indeed, on my view too, none of *those* has been chosen as the referent of "outback": "Outback" refers to the outback, and the outback does not have a precise boundary.

to call a horse's tail a "leg" a horse would still have four legs. Now suppose someone *were* foolish enough to make "outback" refer to one of Lewis' many precisely bounded objects. Why should that have any effect on whether it's vague where the outback begins? The semantics of "outback" would have changed of course, just as they would have had we decided to make "outback" refer to the number three; but on the face of it, such changes shouldn't affect the outback itself.

Suppose I have a typical ripe orange, and I ask for an account of the fact that it is orange. It is not correct to say that it is orange because we decided to call it "orange". Certainly that doesn't initially seem to have much to do with why the orange is orange; the correct account turns out to have to do with the light-reflectance properties of its surface, and maybe some other things. Now suppose I had a peach whose flesh was a sort of orangeish-yellow color: "not definitely orange, but not definitely not orange", we might say. Now suppose that I ask for an account of the flesh's being neither definitely orange nor definitely not orange. Again, how we decided to use some word is not at issue; what is at issue is the peach flesh and its color. On the face of it, an answer that appeals to language does not start in an appropriate place.

Similarly, were we to go to an area that used to be quite definitely part of the outback, but now borders a city and has already been physically marked by developers who will soon begin to erect condominiums, the fact I report when I say that it is vague whether this area is still part of the outback is not about our language.

Further, note that as the world changes, what sentences are vague changes. Twenty years ago this area was definitely part of the outback; now *it* has changed, and the language has not—at least not in any obviously relevant way. Thus the facts about the border of the outback have changed, and not because some facts about how we use the word "outback" have changed.

These points about the direction of explanation are not meant to be decisive, only to re-orient one's intuitions and show that the view of the opposition is far from obvious, and some distance from the natural. I turn now to some other examples that I believe intuitively suggest that there are worldly sources of linguistic vagueness.

1.3.3 Nature's joints and certain gradual changes

Consider the following model of one of the tasks of science: expressed metaphorically, our job, as students of nature, is to cut reality at its joints. The metaphor is familiar and suggestive. On the face of it, there is a natural geographical "joint" between the Gobi desert and the rest of Asia, as there is not between the perfectly rectangular state of Colorado and its surroundings.

This is a gross example, but it is easy to see that we could mistake a relatively insignificant boundary for a significant one, and it is even easier to see that we could fail to notice some very significant ones. But when we look for nature's joints, we may find that they are not perfectly sharp.

When we look at the details, we find that the joint between the Gobi and its surroundings is not perfectly sharp: there simply are no facts that determine *exactly* where the Gobi begins—not at the level of millimeters. Such a finding regarding any geographical entity is quite possible, and, it should be noted, contingent if actual. And so it goes for the objects of study in geology, biology, astronomy, and other sciences.[7]

So nature's joints may not be sharp, it seems. And if so, this might explain why certain parts of our language are vague. (We will consider this thought a little more in section 1.5.)

Another potential source of linguistic vagueness is gradual change.

Many changes take time, and many are gradual. There are two importantly different sorts of time-consuming change, however, one which seems to require indeterminacy and one which does not. The kind that does not is exemplified

---

[7]Some philosophers dispute the question whether there is some "objective" notion of natural (as opposed to unnatural) boundaries. I do use the notion of *naturalness*, however, as well as a related notion of *arbitrariness*. I cannot, in a few pages, convince doubters; I hope they play along as far as they can.

by the process of becoming sixty years old. There is a sense in which growing to be sixty years old is a very slow, gradual process, for it takes sixty years for something to undergo it. Yet the change from being not yet sixty to sixty *itself* takes very little or no time, and so there may be no indeterminacy involved in such a change. The kind of time-consuming or gradual change which involves indeterminacy analogous to the indeterminacy about the spatial boundaries of objects is gradual or time-consuming change from the obtaining of a state of affairs to the failing to obtain of that state. This kind is exemplified in some changes from being alive to not being alive.

A lioness chases down an antelope. During the chase the antelope is clearly alive, though it suffers a progression of ever-more serious wounds. Eventually, the lioness manages to bring the antelope down, and though still alive, it has clearly begun to die. Its blood pressure begins to drop rapidly, its breathing becomes labored and ineffective, its heart beats erratically and its body has been severely and irreparably damaged. At this point it is not so clear that the antelope is alive; but it is not clear that it is not alive either. Finally the heart beats for a last time, and then brain activity slowly comes to an end. After a short time, it is clear that the antelope is no longer alive.

There is a period during which the antelope was dying but not yet dead. Such a period does not yet give us an example of worldly indeterminacy. But intuitively, the passage from being alive to not being alive may itself take time, and if so, it is indeterminate just when the antelope passes from being alive to not being alive. And this explains why the predicate "is alive" has a certain vagueness. It is because of this that it neither straightforwardly applies nor straightforwardly fails to apply at all times during this change.[8]

The death of the antelope is a dramatic example, not only because of its inherent drama, but because of the logical drama it exemplifies: it is not simply the losing or gaining of a property: it is the going out of existence of a thing. Less logically dramatic changes may be gradual or time-consuming changes from the having of a property to the failing to have it, or vice-versa. For example, shortly after its birth, the antelope was not a biological adult. Yet a few years later it is a biological adult, but the change is gradual, and so there does not seem to be a sharp line between the times at which it is not a biological adult and the times at which it is.

---

[8]I am influenced here by Warren Quinn's "Abortion: identity and loss" [24], which includes a nice discussion and defense of the idea that the coming into existence of a typical human being is gradual, in the relevant sense. Quinn is careful to note that this idea is not (merely) the idea that the sortal "human being" is vague, but that it is rather a supposition about language-independent facts in the world. He does not explicitly suggest, as I do, that these facts in the world *explain* (at least in part) the vagueness of "human being" or "is alive".

1.3.4 Micro and macro

A few words should be said about what might be called the "ontological level" of the objects and phenomena at issue here. They are not atomic, in the sense of having no parts. They are composed of parts, and many of their most important properties are determined by the facts about their parts. Those philosophers who feel that there are not *really* any objects above and beyond the simplest things that there are (sub-atomic particles?) will think that many of my examples concern things which don't really exist or exist only in our linguistic and conceptual conventions.[9] In a sense, my examples cannot convince them that vagueness can have a *worldly* source, because they are convinced that the sorts of facts I point to—facts about such macro-level objects as the outback—are facts about things which themselves exist only in some attenuated sense ("by convention" perhaps).[10] I will not address reductionism or idealism of that depth, and I will assume that macro-level objects *exist* in a robust sense, even if facts about them are in some important ways determined by lower-level objects and facts. I believe that once it is

---

[9]Democritus seems to have clearly stated the view, and there have been and are many philosophers who, for various sorts of reason, agree.

[10]There is no reason in principle why there could not be vagueness at the level of atoms, sub-atomic particles, or even at the level of the maximally fundamental type of matter, if there is such a thing.

granted that there are tables and chairs in addition to sub-atomic particles, it will emerge that some or perhaps all of these macro-level objects have vague boundaries in space and time.

I turn now from objects and their boundaries to another sort of example: vocabulary for covering and dividing a range of similar properties.

1.4 The smooth gradient

I claim that there are contingent, empirical facts about the world that explain some of the vagueness of natural language. The general shape of one sort of worldly condition that leads to vague language emerges when we consider the sorts of difficulties we might run into were we to design a language meant for ordinary human use. I will call that condition the "smooth gradient".

1.4.1 A simple example: color vocabulary

Imagine that we have been given the task of designing a language for a soon-to-exist race of human beings who will inhabit a world somewhat similar to earth. (Don't try to make the scenario overly realistic: in particular, do not worry too much about how the people are supposed to *initially* learn the language; maybe we learn it ourselves first and teach it to their initial generation, and then leave the scene.) We work on the team whose job is to design their ordinary color vocabulary; our specialty is predicates.

Simplicity and taxonomical completeness

The vocabulary we give them is to be a simple, everyday human language, and so must not be overly complex. In particular, we are allowed at most twelve simple predicates. Yet the vocabulary is to provide, among other things, a nicely structured taxonomy of the colors of the world; at least one subset of our predicates should provide a non-trivial exhaustive taxonomy. That is, the members of this subset must be mutually exclusive, but must together exhaust the colors, so that everything is in the extension of one of them, and nothing is in more than one.

Fregean goals

Suppose our lead language designer is a disciple of Gottlob Frege. We lesser engineers are told that we must strive to make our language precise. This means at least that, for each predicate and each object, the object must either satisfy or fail to satisfy the predicate, *and* there is to be no indeterminacy about which is the case.

A cooperative world

The facts about color in the world of the prospective language users may or may not help us. If color and light are very simple our job may be very easy. Suppose for example that, in their world, there is only white light. There is nothing transparent except for perfectly transparent air. Suppose that every last object in their world is of one of two shades: one that we would call a clear case of "black" and one that we would call a clear case of "white". There are no colors besides these, and there are no shades of gray. There are no prisms, or anything else that could produce a color other than one of the two shades. For this world, our job is trivial: we need only two color words, and it is obvious what their extensions will be. (Perhaps we should throw in a third word, covering everything that is not transparent.)

A less cooperative world

Now consider a world just like the black-and-white world except that there are also some shades of gray, all more or less equally common among the salient objects of their world. Suppose there are 101 of them, more or less evenly distributed between pure white and pure black. Let us enumerate them, with number one as the blackest and number 101 as the whitest. How shall our color predicates divide up the shades? We need to provide a non-trivial taxonomy, and hence we will need a set of no more than twelve simple predicates that are exhaustive and mutually exclusive.

Now, if we want to have two color words in our taxonomy, it will be unclear what their extensions should be. Surely the most natural thing to do would be to put the darker shades under one predicate and the lighter shades under the other. So we might have a predicate corresponding to "dark" that would cover at least shades one through fifty inclusive, and a predicate corresponding to "light" that would cover at least shades fifty-two through 101 inclusive. But there are 101 colors to deal with, so there will be one extension larger than the other. Where should we put shade number fifty-one? There does not seem to be any better reason to place it with the darker shades than with the lighter shades. Perhaps we should try to avoid this arbitrariness by creating a three-word taxonomy instead.

But the shades do not divide evenly into three groups, nor into any number of groups less than 101, except of course for the trivial all-inclusive "division".

We must provide a taxonomy. But we do not see any reasonable hope for thinking that there is some one possible taxonomy that is better—more rational, more natural, more useful—than any of the other possibilities. We will be forced to choose arbitrarily.

1.4.2 The extent of arbitrariness in design

In 1.4.1 we saw that an empirical fact—the smooth gradient of a large number of shades of color—can affect whether there is a best solution to a given language-design problem. Upon reflection it becomes clear that our world is probably more like the "uncooperative world" with the 101 shades of gray than it is like the "cooperative world". After all, the colors in the world with the 101 shades of gray are much more orderly than those of the world we live in. If we were designing a language for ourselves to speak, then if we wanted a relatively compact, humanly usable language, and if we wanted our predicates to be bivalent, we would be forced to make some arbitrary choices.

Could we avoid arbitrariness in design if we gave up other goals? Well, we *might* be able to escape arbitrariness if we allowed ourselves a less human language. For example if we were designing the predicates for the world with the 101 shades of gray, we might have a predicate for every single shade. And we might includes predicates for each set of adjacent shades in the series, or even for each subset of the shades. And it is true that, for certain *human* purposes, it would be desirable or even necessary to have a predicate for each shade. (If there are a great many shades, it will be impractical to have so many grammatically simple predicates. We would use instead a complex system for referring to shades.) Nevertheless, it is very desirable to have an all-purpose, general, "ordinary" vocabulary for colors as well.

Could we avoid arbitrariness in the design of everyday language if we gave up the desire to make a Fregean language? This is a difficult question to consider in the abstract. I believe that the answer depends upon the details of the subject matter. But in the world of the 101 shades, it is difficult to see how any language could avoid arbitrariness.

It is fairly clear that the most obvious alternatives to the Fregean possible languages do not help with the issue of arbitrariness. Consider for example a trivalent language, in which sentences can have one of three truth-values: True, Middle, and False. Never mind how the logical connectives work in this language; arbitrariness will arise in the construction of a color language in the uncooperative world. Though it is not clear how to think of the goal of providing a taxonomy when we use a trivalent language, it is clear that some such goal is reasonable. (And if we drop the goal, then there are even fewer rational constraints and thus there is even more arbitrariness.) No matter which way we divide up the 101 shades into groups of three or groups of groups of three, we face the prospect of making an arbitrary choice, just as we did with (groups of) groups of two.

Things are somewhat less clear if we shift to a language with many truth values. This is partially because it is hard to understand just what it means for there to be many truth values. The obscurity is partially due to obscurity about what a "truth value" really is, but let us set that aside.[11] It is also unclear what the requirement that there be a taxonomy of colors would amount to.

Will any of the possible ways of designing the language stand out as a single most reasonable or most natural possibility? Though it is hard to be sure until we carefully consider specific proposals, I think the prospects are not very good.

Suppose our truth values are the real numbers from zero to one inclusive. Here is a fairly natural "extension" for a color predicate for the world with the 101 shades of gray: the darkest shade (number 1) satisfies it to degree 1, the lightest shade (number 101) satisfies it to degree zero, and in general shade number n satisfies it to degree (101-n)/100. But it would seem that there should also be a predicate which is straightforwardly true (i.e. true to degree 1) of some of the shades of gray between the darkest and the lightest, and the problem of giving a "cut-off" for shades that satisfy the predicate to degree 1 will now look just like the problems we faced when we were considering

[11] I think I understand well enough the idea of a sentence's or a claim's *being true*, but I do not really know what *The True* might be.

bivalent predicates. And since there are so many possible extensions when there are so many truth values, arbitrariness seems inevitable.[12]

Perhaps a vague language is just like a precise language, except that it has more truth values or a different logic, or both. Perhaps not. Either way, there may well be many possible vague languages that are very similar. And if so, then where the world presents the sort of smooth gradient considered above, there may well be parallel issues of arbitrariness for vague and precise languages. In section 1.5 I will suggest that there are sources of linguistic vagueness that are not like the smooth gradient, and that for these cases there is an important difference between precise languages and vague languages.

1.4.3 The connection to vagueness in natural language

Perhaps there is no best possible semantic structure for a given part of human language: no matter how the logic works, no matter whether there are many truth values or just two, a language designer aiming to create a good human

[12]Another interesting system of truth values to consider has the truth values being the shades themselves. There are a few possible extensions that seem very natural: one on which a predicate is satisfied by a shade exactly to the "extent" of that very shade; one on which a predicate is satisfied by any shade exactly to the extent of the darkest shade; and so on. Such a system might get us some way toward avoiding arbitrariness, but of course it is hard to know what it amounts to for a sentence to be "true" in such a system. Moreover, it is difficult to see how the "truth values" of our fragment of the language will combine with "truth values" of other parts of the language. The idea may have some potential, but it looks very hard to develop.

language would have to choose arbitrarily from among alternatives when assigning "extensions" (and meanings) to a certain part of the vocabulary. Let us suppose it is indeed so. Does this fact have anything to do with our *actual* language?

I believe that it does: the vagueness of our language is partially *explained* by the facts about the world that make it true that arbitrary choices would have to be made by anyone constructing a language for us. Consider again the worlds for which we were designing color vocabulary, now focusing on how a natural human language might evolve in these worlds. It is very plausible that in the two-color world, the color-vocabulary will come to have the shape that we would have given it.

But in the world with the 101 shades of gray, it is less clear how the everyday language will evolve. (Of course eventually there will come to be a technical vocabulary for designating each of the shades, but it is unlikely that that vocabulary will *supplant* the smaller, more primitive ordinary color language.) Just as we would have to make a number of arbitrary design choices to settle the exact extensions of the color vocabulary, so too will language makers inhabiting such a world. And since the development of natural language proceeds, at least initially, in anything but the orderly way our design committee would proceed, it seems implausible that there will come to be a well-established convention that will give the everyday color vocabulary precise boundaries. Thus it seems very likely that it naturally will be a vocabulary for which precise boundaries are not in fact drawn, and hence a vague vocabulary.

1.4.4 The character of the smooth gradient

I think it is fairly clear that the world of the 101 shades of grey has a character that makes for arbitrariness in language design. But it is not clear *exactly* what it is about this world that makes that so. While there is not room fully to pursue the matter, we may say that it seems to be something like this: there are 101 items that are in some sense equivalent, and that come in a natural linear ordering. One can now see that the project of giving a small taxonomy of a group of things like this will face arbitrariness unless there are some other relevant facts that would make some taxonomy stand out.

I chose the number 101 for dramatic effect. Since it is prime, there is no way non-trivially to partition the group *evenly*. But there being 100 shades instead would not have changed the essential point, I believe. For it is hard to see how to choose between a taxonomy of five even groups and one of ten even groups. But what if the numbers "work out"? There are numbers that are divisible by only one number between two and thirty. Maybe there is, as it happens, one non-arbitrarily best solution to our language design problem if there are exactly 81 shades of grey. But it is not perfectly clear that having precisely equal groups in the taxonomy is of high enough value to rule out some non-even taxonomies as equally good solutions to our problem.

In reality, there are smooth gradients with many more equivalent elements. Perhaps in come cases, due to a lucky numerical fact, there happens to be a unique way to split the elements evenly into twelve or fewer groups, but I think the reader will grant that the possibility of such cases does not much affect my main point. There may also be gradients with infinitely many elements, and for these similar problems will arise. This is true even if there are natural notions of "less" and "more" and "near" and "between" for them. For how many elements should there be in the taxonomy? Unless there are further relevant facts about the gradient or things related to it, it is hard to see any more reason to break the elements into four rather than five groups, for example.

In this section I have argued that there are empirical facts about the world that can make vague language natural, almost inevitable. In the absence of a serious effort, on the part of language designers, to precisify the language, the lack of non-arbitrary breaking points in the world seems likely to cause a certain amount of vagueness in the language. But this does not reveal anything particularly *good* about vague language. In the next section I will supplement the arguments of section 1.3 with further arguments to the effect that there are empirical facts about the world that make vague language *better* than precise language.

1.5 Vague language can be better than precise language

Vague language is clearly superior to precise language if there are important subject matters that any precise language *mis*-represents, and that, for at least this reason, are more accurately described using vague language. Arguments over whether there are objects with this character threaten to be nothing more than brute metaphysical head-butting; I say *there are* such things, my opponent says *there aren't*, and purported arguments on either side look like "question begging". I will nonetheless attempt to defend my side, and I will try to present evidence for the superiority of vague language that does not simply say "there are vague objects and so there should be vague language." Among other things, toward the end of the section I will address a possible opponent who is willing to entertain the existence of vague objects, but sees no reason to think that vague language is any better than precise language.

I focus on an example not discussed in section 1.3: biological species. The examples of section 1.3 could be given somewhat similar discussion, but species are among the more important and more vividly recognizable cases of things the description of which calls for vague language, and it is useful to give them a separate, more detailed treatment.

1.5.1 Mother and daughter species

Suppose that for a couple hundred thousand years a species inhabits an ecologically stable geographical region; a certain coastal area, say. A series of ancestrally connected creatures live in this area and, with a few exceptions, are pretty much genetically and morphologically the same. They exhibit pretty much the same behavior, relate to the other creatures in the area and the larger environment in pretty much the same way. They are reproductively isolated: no other creatures are capable of interbreeding with any of them, yet they can interbreed.

Now suppose that after this period of stability, some relatively catastrophic environmental change occurs—not in a single day, but over the course of about a thousand years. After this, the species, we would say, struggles to survive, and most of the creatures die off. But some happen to inherit unusual sets of genes that had little or no relevance when present in their ancestors, but, as it happens, now give them abilities better to exploit the new environment and (just barely) survive and reproduce. The population is quite unstable and becomes much less uniform than before; there are many variations on the genetic themes, and many combinations of them, most of which do not "work" at all. Some of the descendents of the original species cannot even mate with other equally direct descendents. Eventually, after a few thousand generations, stability is regained. There is much less genetic variation to be found among the reproductively isolated creatures than in the transitional period; a certain set of new genes has "won out" so to speak. The new creatures would not be able to reproduce with the members of the mother species, and their behavior and morphology are substantially different. A single daughter species has been born, and, let us suppose, it remains stable for a couple hundred thousand years.

Whether evolution ever *actually* happens this way is not relevant for our purposes; only that it could. And I will take it for granted that it could. I will argue that the best description of a world in which it does occur this way will involve two names for two species, and that *since* it is vague exactly what the "boundaries" of these species are, there will be a certain amount of vagueness in certain possible sentences involving these names.

1.5.2 There might not have been non-arbitrary "species" boundaries

In a world in which the story of 1.5.1 is true, there are species whose members all share a rough physiological structure. But, if we consider the logical space of all possible physiological structures in the abstract, no single non-trivial taxonomy will stand out. The situation is in this respect like that in the world with the 101 shades of gray; but here there are very many dimensions along which smooth gradients extend. Consider the physiological structure of a typical garden snake. We can imagine a series of pairwise very similar structures leading from the structure of the four-foot garden snake to the structure of a big anaconda. And we can imagine a series of pairwise very similar structures from that of the garden snake to that of a small sea snake. And also from that of the garden snake to that of a small lizard. There are very many dimensions in the logical space of possible physiological structures; each structure is "near" very many similar structures, and a little reflection will reveal that if we simply consider possible structures *as such*, in the abstract, so to speak, there is no single most natural taxonomy except for the trivial ones that lump all structures together or which fragment the space maximally, into its individuals.

It is with reference to the actual course of history that some possible structures deserve to be separated out as those belonging together as the typical structures of creatures *of a species*. If there were a world in which there were somehow no inheritance, in which creatures randomly popped into existence with random structures, there would be no reason to lump structures together as typical structures of members of a species. There would be, it seems to me, no species at all.[13]

But in a world like the one of section 1.5.1, there is not simply a smooth gradient; there is a relatively small but continuous transition between two well-defined and much larger "regions".

1.5.3 There might have been sharp boundaries for species

We can easily imagine worlds in which species have sharp boundaries. Suppose that what it is for a living thing to come into existence is for one of a finite number of distinct *forms* to become present in some matter. The living thing is a sort of composite of the form and the matter, and the presence of the form

---

[13] A similar example is used for related purposes by Michael Ayers in Vol.2, Chapter 7 of his *Locke* [3]. Locke was interested in borderline cases of species-membership, and wanted to use such examples to argue that the *mind*, rather than the natures of the things classified, fixes the boundaries drawn by classifications. Leibniz replied that borderline cases show no such thing. Timothy Williamson reports briefly on the issue between them in [31], pp. 34 & 35.

causes the composite to tend to assume a certain shape and behave in certain characteristic ways: the form guides the matter toward a certain perfection. But because of inevitable imperfections in the matter, the composite is unlikely to achieve full perfection, and will accidentally have lots of features not dictated by the form, some of which may be contrary to the image the form pushes the matter toward. The composite—the living thing—dies when the accidental stresses on the matter cause the form to lose its grip on the matter, so that the matter is no longer enformed by that form. To be of a species is to have the form that defines or is that species.

In such a world, there will be no indeterminacy about what species a given living thing is of. The question is simply resolved by which form is the form of that living thing, and the forms are all quite determinately distinct from one another.

But our world is not like that, and what makes a living thing be in a given species in our world seems to involve the possibility, and probably the actuality, of indeterminacy.

1.5.4 Arbitrariness and vague language in our example

Consider the project of describing the biology of the world of 1.5.1. The mother species slowly goes extinct, and at some point the daughter species is flourishing. But because of the relatively slow and gradual nature of the speciation process, for each species, there is no natural sharp boundary to be drawn between the creatures that are in it and those that are not. Thus if we were to use only precise language in attempting to describe this world, we would have to choose arbitrarily some extension for "is a member of species S" (where S is meant to name either of our species). Is vague language any better off?

I suggest that it is. One would be missing something about this world if one did not notice these two species in it. There really are, to put it picturesquely, two calm pools of creatures historically separated by a stream of gradual change.[14] The first "pool" forms a unity, and so does the second, but there simply is no line that deserves to be called the boundary of one of these "pools". Instead of a smooth gradient of very many items on a par, we have a small, gradual transition between exactly two items on a par: the two adjacent pools connected by a short, wide, murky stream, not a long, straight,

[14]This image was suggested to me in conversation by Daniel Dennett. Dennett's work includes some ideas sympathetic with mine: see [5], the first half of which covers some themes related to those in this essay, and also [6].

I disagree with Dennett's construal of a contradiction between the intrinsic vagueness or indeterminacy of the boundaries of species and the traditional view called *essentialism*. For I think many—though certainly not all—of the ideas traditionally associated with the term "essentialism" are correct, and in fact are enhanced by the recognition that natural kinds, real properties, and other essentialist paraphernalia essentially involve a certain amount of indeterminacy. But the issue may be largely terminological.

distinctly flowing river.

Choosing any one possible precise language to describe this world requires choosing arbitrarily. But there is a single choice for the semantics of a language aimed at describing the situation, a choice that naturally stands out, with no arbitrariness, as the most accurate, best choice. The critical semantic facts about a non-arbitrary language are that it has common nouns or predicates naming the two species and that it has referring expressions referring to, and quantifiers ranging over, the various animals. These facts guarantee that it is vague, and that it has the best resources for describing (part of) the situation. Thus if we choose a vague language to describe this world, we make a better, non-arbitrary decision.

1.5.5 Only vague language lets us express the important facts

A straightforward way of putting my thesis about the example is this: it's a fact that there are these two species, and that they do not have sharp boundaries, and thus if you have a name for the species in your language, and a name for each creature, and a word for the membership relation, you will have a vague language: for *if a language can, for each animal, express the propo-*

sition that that animal is a member of the first species, then it can express a proposition that is neither determinately true nor determinately false. If a sentence expresses a proposition that is either determinately true nor determinately false, then the sentence is neither determinately true nor determinately false. Thus any language all of whose sentences determinately have one or another truth value, and hence any precise language, must therefore be failing to allow us to straightforwardly and accurately represent the creatures, the species and the facts about them.

I thus agree with Gareth Evans' description of what it amounts to to suppose that the vagueness of language is a direct result of a sort of vagueness in the world.

> It is sometimes said that the world might itself *be* vague. Rather than vagueness being a deficiency in our mode of describing the world, it would then be a necessary feature of any true description of it.[7]

Understood properly, this is just about right. It is not quite literally true that "any true description of the world" will be vague, if a "description" can be as short as a single sentence. For there are true sentences that are perfectly determinately true, and could not have been otherwise, which describe the world. (The sentences of mathematics, for example.) But understand it to mean that any language which can express even all of the salient facts about the world will itself necessarily be vague, and it is a good expression of the idea I am defending.

1.5.6 The challenge of the many

There is an important line of objection to my claim of the non-arbitrariness about a vague language. My claim is that (relative to the project of describing the relevant part of a world like the one of 1.5.1) vague language differs from precise language in this: there is no single best, most natural possible precise language, yet there is a single best, most natural vague language. The challenge is this: just as there are very many, very similar, equally good, possible precise semantic profiles for a language that differ only slightly on where they "draw the line" (between, for example, members and non-members of a species), there are very many very similar possible *vague* semantic profiles that differ only slightly on where they "draw the line". Whatever my favored language's drawing a "vague line" turns out to amount to, the objection runs, there could be another vague language just like it but differing just slightly in its semantic profile, and thus differing just slightly with regard to the character of the "vague line". For example, perhaps in the first language the sentence that approximates "Charlie is a member of S" is determinately neither determinately true nor determinately not true, but in the second, the corresponding sentence is not so.

My response is that we really have no good reason to believe that there are so many very similar, almost identical, possible vague semantic profiles, of which there are many equally good candidates for being the best or most relevant, since there are not many very similar, almost identical, items, any of which are equally good candidates to play the role of the species.[15] I give a little more discussion of the problem of the many in 5.5.

[15] Or the role of Charlie, or the role of the membership relation.

Chapter 2

Indeterminacy and boundaries

In this chapter we examine the notion of the lack of sharp boundary which seems characteristic of vagueness, and its connection with determinacy. It is argued that the lack of sharp boundary is a deep phenomenon that cannot be easily reduced to any easily describable lack of determinacy.

2.1 Unsharp boundaries

One the main things that makes a *sorites* series of objects puzzling and interesting is the intuition that there is an *unsharp boundary*, for example, between the red and the non-red in a line-up of color tiles that ranges from a paradigm case of red to a paradigm case of green.

Let us focus on the following example: We have attached to the back of a truck a machine that sprays a one foot wide fan of paint onto the ground immediately behind the truck. The paint it sprays is mixed from two large vats of paint in the truck, and the mixture is precisely controlled by a computer connected to the odometer of the truck. We have a straight ten mile stretch of white concrete road to paint on, and we have programmed the computer to continuously vary the mixture of paint from the two vats (A and B) so that: at the beginning of our strip, the mixture is 100% from A and 0% from B, after five miles, the mixture is 50/50, and at the end of the ten mile long strip, the mixture is 0% from vat A and 100% from vat B. Vat A is filled with paradigmatically red paint, and vat B is filled with paradigmatically green paint.

If nothing goes wrong with our equipment, we will produce about as good an example as one could hope for of an unsharp spatial boundary. For we will have a colored stripe of paint ten miles long, and certainly some stretch of the stripe will be red. But the spatial boundary of the red stretch of the stripe is unsharp.

We could also use our machine to produce a series of ten thousand uniformly colored tiles, any adjacent two of which are indistinguishable by the human eye, but which range from red to green. These too can be used to produce an example of an unsharp boundary, but the notion here seems to be not purely spatial. There is no sharp boundary for red tiles, and this holds no matter how the tiles are arranged in space. Scatter them to the four corners of the earth and one does not change the "boundary" in the relevant sense.[1]

Such examples provide standard introductions to the sort of vagueness philosophers are interested in. Everyone wants to say of these, or at least of some such examples, that they give us situations in which there is an unsharp boundary or a lack of a sharp line. In our example, there is an unsharp boundary for the red tiles. One of the main philosophical questions about vagueness is then: What does the lack of sharpness in a boundary amount to?

2.1.1 Reduction of the unsharp boundary

The notion that concerns us in the study of vagueness is not the notion of a boundary *per se*, but rather the notion of a *lack of sharpness* in a boundary, whether spatial or logical. It is natural to seek a way to reduce or capture talk of the lack of sharpness in a boundary in other vocabulary: for example, we might try to find a way to render "The red tiles have an unsharp boundary" in some other terms.

---

[1]Though there may be also the unsharp spatial boundary of the mereological sum of the red tiles, or some such scattered object.

What are the first ideas one might consider about explaining the lack of sharp boundary by some other facts? Let us focus on the example of the 10,000 colored tiles.

A non-starter

One suggestion about the lack of sharpness in the boundary of the red tiles is that it amounts to there being things "in between" the red and the non-red. Thus there would be no pair of adjacent tiles such that the one is red and the other not red, and so, in a sense, no sharp boundary between the red and the non-red.

The problem with this suggestion is that it is not clearly intelligible. For suppose that tile number 5000 is one of the tiles between the red tiles and the non-red tiles. Then it looks as though tile number 5000 is not red, since if it were, it would not be between the red and the non-red, but would be among the red. But then it is not red, and hence not *between* the red and the non-red, but rather, *among* the non-red.

Now one can make sense of the suggestion with the notion of "choice" negation, so that

Tile number 5000 is neither red nor not red.

ends up saying that tile number 5000 neither satisfies "is red" nor "is not red" (where the "not" expresses the supposed "choice" negation, while the negative sense of the "neither... nor" is that of the supposed "exclusion" negation).[2] I take up this way of understanding the suggestion, though not in the vocabulary of "choice" (as opposed to "exclusion") negation, just below, in the subsection after next. To the extent that I understand the distinction between the two kinds of negation, understand "exclusion" negation to be a meta-linguistic notion, and hence I treat it as just the notion of some sentence's being not true or some thing's failing to satisfy some predicate, and I express it this way.

There's at least one that's neither determinately on one side nor determinately on the other

We can improve the above suggestion: we can interpret it as meaning, in our example, that there are things in between the tiles that are *determinately* red and the tiles that are *determinately* not red. If tile number 5000 is one of these other tile, then it is not determinate that tile number 5000 is red, and it is not determinate that tile number 5000 is not red.

In general then, the unsharpness of the boundary of the F's would consist

[2]Practitioners of three-valued logic commonly make the distinction between "choice" and "exclusion" negation, which have the following characteristics: the "choice" negation of a sentence is true just in case the sentence is false, neither true nor false when the sentence is neither true nor false, and false just in case the sentence is true; the "exclusion" negation of a sentence is true just in case the sentence is not true, and is false otherwise.

in there being things that are neither determinately F nor determinately not F.

This suggestion is more intelligible than the first one, for one can make *some* sense of the suggestion that a thing x be neither determinately F nor determinately non-F. This does not seem like a self-contradiction, the way the assertion that something is *neither F nor not-F* does. But it is puzzling nonetheless, for if x is such a thing, then it seems that x could not be F; if it were, then it would be determinately F, wouldn't it? But this consideration strongly suggests that x is not F. The advocate of the suggestion must address this seemingly plausible consideration.

There's at least one that's neither truly on one side nor truly on the other

An apparently similar suggestion is that there being an unsharp boundary for the red things, for example, is for there are things that are neither truly "red" nor truly "not red". This suggestion goes naturally with the logical treatments of vagueness with truth-value gaps or alternative truth-values.

This suggestion is just like the most recent one, except that it replaces "determinately" by "truly". It thus faces a similar task: to explain why the inference from "x is not truly 'red' " to "x is 'not red' " fails. Again, it would seem that if x is not truly "red", then x cannot be red, and so must be not red.

These two suggestions are formally similar, but I think there is good reason to think that the latter is not as plausible. For on the face of it, when we say that there is an unsharp boundary between the red tiles and the rest of the tiles, we are not saying something about a piece of language. What we say would be true even if the English language had never existed, it seems, and hence even if there had not even been the predicate "is red". Moreover, what we say would seem to be reportable in languages other than English, at least in principle. That would not be the case were we really saying something about the predicate "is red".

The standard view: reduction to determinacy or to truth

The suggestion that the lack of sharp boundary for the red tiles somehow boils down to facts expressible with the notion of determinacy or with the notion of truth, as something like the existence of things which are neither determinately (truly) red nor determinately (truly) not red can be said with some justice to be the "dominant view" in the current philosophical literature about vagueness. Some such view can be found explicitly or implicitly in the work of Gareth Evans, Kit Fine, David Lewis, Vann McGee, Terence Parsons and Peter Woodruff, and Michael Tye, among others.[3]

A few authors have expressed views which suggest that this is not the case; most famously Timothy Williamson, who suggests instead that the unsharpness reduces to facts about *unknowability* rather than *indeterminacy* or lack of truth value.[4] But on Williamson's view there is something *formally* identical with determinacy, playing an identical role: knowability. R. M. Sainsbury has registered some doubt of the idea in [26], as has Kit Fine;[5] neither of these authors are motivated by epistemicism.

2.1.2 Why these suggestions do not tell the whole story

> Someone might seek to obtain precision in the use of words by saying that no word is to be applied in the penumbra, but unfortunately the penumbra itself is not accurately definable, and all the vaguenesses which apply to the primary use of words apply also when we try to fix a limit to their indubitable applicability.
>
> Bertrand Russell [25]

Our suggestion, that there is an unsharp boundary for the red things just in case there are some things that are neither determinately F nor determinately not F, seems to capture some part of what we think when we think that the

[3]See [7], [8], [14], [20], [23], [27] and other works by the same authors.

[4]In [31] and elsewhere.

[5]In conversation, March 2000.

red things have an unsharp boundary. But it does not seem to tell the *whole* story in a typical case of vagueness. To the extent that we can make sense of it, it does not seem to be sufficient to fully express the felt lack of *any* sharp boundary between the red and the non-red.

The suggestion is compatible with there being a sharp boundary for the things that are determinately red and for things that are determinately not red. But if that's the way it is, then the lack of sharp boundary for redness is remarkably shallow. All we would have to do is choose to mean "determinately red" by "red", and there would no longer be an unsharp boundary for our word "red". This would be a very slight change of usage. (Slight, since everything we would have before called "red" we would still call "red", and everything we would have before called "not red" we would still call "not red". But a change nonetheless, for we would now say that *there is* a sharp line between the "red" and the "non-red".)

If it were really *that* easy to give "red" a sharp boundary, then one wonders why we wouldn't have done so long ago, or why the word does not naturally "gravitate" toward having such a meaning. But the fact is that it is not so easy to give "red" a sharp boundary, and the unsharpness of the boundary for redness seems to run much deeper than any lack of sharpness which is so near a presence of sharpness. Our intuitive sense of the unsharpness of the boundary between the red tiles and the rest of the tiles is not captured by a picture in which there is a completely sharp partition of the things into three classes: the determinately red, the neither determinately red nor determinately not red, and the determinately not red.

Instead, it seems that there is just as much of an unsharpness of boundary for the determinately red things as for the red things. To the extent that we can make sense of the existence of things which are neither determinately red nor determinately not red, we ought to be able to make sense of the existence of things which are neither determinately determinately red nor determinately not determinately red. Moreover, the unsharpness of the boundary between the determinately red and the rest of the things seems to be *part of* or to *flow from* our original sense of a lack of a sharp boundary between the red things and the rest of the things.

2.1.3 Immediate and grand unsharpness

Our suggestion was meant to capture or explain what it is for there to be an unsharp boundary. It may be defended against the immediately foregoing charges of inadequacy by appealing to a distinction between two ideas of unsharpness: immediate and grand. The suggestion plausibly captures the first, but not the second.

The intuitions which seemed to tell against our suggestions are seen as intuitions that there are many *immediately* unsharp boundaries. The somewhat vague intuition that there are *no* relevant sharp boundaries between the red and the non-red we can call the intuition that there is a *grandly* unsharp boundary. Thus our sense that there is an unsharp boundary between the determinately red and the not determinately red would be seen as a sense that there is a *second* immediate unsharp boundary, additional to the immediate unsharp boundary between the red and the not red. Our sense that there are *no* sharp boundaries between the red and the non-red would be seen as the sense that for some broad range of boundaries, all are immediately unsharp.

The grandly unsharp boundary seems to require an immediately unsharp boundary for the red (and for the non-red), an immediately unsharp boundary for the determinately red and for the determinately not red, an immediately unsharp boundary for the determinately determinately red and for the determinately determinately not red, and so on. What exactly the "and so on" covers is not immediately clear, but this formulation is a step towards a formulation of an important intuitive notion of a grandly unsharp boundary. We will now consider in greater detail how the formulation should go.

2.2 Trying to capture the grand lack of sharp boundary

Let us suppose that our intuition that there is an unsharp boundary between the red and the non-red does indeed split into two: the intuition that there is an immediately unsharp boundary between the red and the non-red—an idea that can be directly captured by the notion of indeterminacy—and a less well defined intuition that there are also many other immediately unsharp boundaries, for such things as the determinately red, the determinately determinately red, and so on. Let us consider how how this second intuition might be made more definite.

Many relevant issues will emerge when we consider the following

Question Given no other assumptions but that there is a grandly unsharp boundary for a predicate F, how many objects must there be?

As we will see, the question is not easy to answer.

We will consider the problem for a modal first-order formal language which contains one symbolic predicate letter F, formula operators $\sim$, $\mathcal{D}$, and $\&$, the quantifier $\exists$, variables, name letters, and . The operator $\mathcal{D}$ is meant to represent "it is determinately the case that".

There does not seem to be an interesting correlate of our question concerning how many objects there *can* be given that there is a vague predicate.

There is no reason to think there is an upper limit to that. But could there be a vague predicate and yet no things, or some very small number of things?

2.2.1 Minimal assumptions about the logic of determinacy

Background Assumptions

Classical Logic For now, we will allow ourselves to use classical logic in our discussion. That is not at all to say that we assume that classical logic is valid within the language we are discussing; only that we will use it in *our* language.

Notions of model, satisfaction, and truth We assume that there is some notion of a *model*, and associated notions of truth and satisfaction, which may or may not be bivalent. Thus we will speak of a sentence's being true at a model, an open sentence's being satisfied by an object in the domain of the model (relative to a variable assignment), and so forth.

These Background Assumptions will make our discussion much easier than it would be without them. It should be kept in the back of the mind however, that ultimately, the assumption that Classical Logic is valid in *our* language may not be justifiable, as we will see in section 3.2.

Next we consider a set of basic and presumably uncontroversial assumptions about the determinacy operator and the logical expressions. I do not assume bivalence, nor lack of bivalence, nor classical logic, nor the failure of classical logic—within the object language.[6] That is, it is not assumed that a model makes a given sentence either true or false, nor that a vague language must have sentences which fail to be either true or false, and so on. (The cost of this neutrality is a little complexity.) I will write "$\mathfrak{M} \models \phi$" for "model $\mathfrak{M}$ makes true (is a model of, models) sentence ϕ". I will write "$\phi \models \psi$" for "every model of ϕ is a model of ψ". Next, on its own separate page, is a schematic presentation of the Minimal Assumptions. ϕ and ψ are arbitrary formulas of the object language, and $\mathfrak{M}$ is an arbitrary model. (I have in mind here that if ϕ is an open sentence like Fx, then to say that ϕ is true at a model is to say that it is true relative to a given assignment at that model, and that, for example, $Fx \models \sim\sim Fx$ means that for all models and assignmnents, if the assignment makes Fx true, it makes $\sim\sim Fx$ true.)

[6]In this connection, those attracted to three-valued logic will want to take the negation ($\sim$) in the object language to be "choice" negation, so that the Minimal Assumptions do not entail bivalence.

$$\mathcal{I}\phi \quad =_{def} \quad ({\sim}\mathcal{D}\phi \,\&\, {\sim}\mathcal{D}{\sim}\phi)$$

$$\mathcal{A}\phi \quad =_{def} \quad {\sim}\mathcal{D}{\sim}\phi$$

D1 $\mathcal{D}\phi \models \phi$

N1 $\mathfrak{M} \models \phi \Rightarrow \mathfrak{M} \not\models {\sim}\phi$

DNE $s{\sim}{\sim}t \models st$ where s is any (possibly empty) string consisting of nothing but $\mathcal{D}$'s and $\sim$'s and t is any sentence.

DNI $st \models s{\sim}{\sim}t$ where s and t are as in DNE.

C1 $\mathfrak{M} \models (\phi \,\&\, \psi) \Leftrightarrow (\mathfrak{M} \models \phi$ and $\mathfrak{M} \models \psi)$

E1 For any variable v, $\exists v\phi(v)$ is true iff there is an assignment to v of an object that satisfies $\phi(v)$

I1 $\mathcal{I}\phi \models \mathcal{I}{\sim}\phi$ (Follows from **C1** and **DNI**.)

I2 $\mathcal{I}{\sim}\phi \models \mathcal{I}\phi$ (Follows from **C1** and **DNE**.)

A1 $\phi \models \mathcal{A}\phi$

A2 $\phi \models \psi \Rightarrow \mathcal{A}\phi \models \mathcal{A}\psi$

M1 $\mathcal{D}(\phi \,\&\, \psi) \models (\mathcal{D}\phi \,\&\, \mathcal{D}\psi)$

M2 $(\mathcal{D}\phi \,\&\, \mathcal{D}\psi) \models \mathcal{D}(\phi \,\&\, \psi)$

M3 $\phi \models \psi \Rightarrow \mathcal{D}\phi \models \mathcal{D}\psi$

M4 $\mathcal{A}\mathcal{D}\phi \models \mathcal{A}\phi$ (Follows from **D1** and **A2**.)

The symbols $\mathcal{I}$ and $\mathcal{A}$ are not taken as primitive. Their "definitions" can be taken as mere abbreviations in our language. $\mathcal{I}$ represents indeterminacy, and $\mathcal{A}$ represents admissibility. Formally, admissibility and determinacy are duals just like possibility and necessity in "modal logic", and indeterminacy is formally related to them as contingency to possibility and necessity.

Assumptions M3 and A2 deserve special mention. They are perhaps the most controversial of all of the Minimal Assumptions. M3 says, in effect, that if ψ is a logical consequence of ϕ, then $\mathcal{D}\psi$ is a logical consequence of $\mathcal{D}\phi$. A2 says something similar about admissibility.

DNE and DNI are together quite powerful. They are generalizations of the most familiar rules governing double negation. They would be disputed by those who attempt to give an intuitionistic logic for the determinacy operator, but they are otherwise uncontroversial, and are used later in this chapter.

2.2.2 Immediate and grand lacks of sharp boundaries

Above we considered what is probably the simplest conception of the unsharp boundary between F's and non-F's: as the existence of a thing which is indeterminate with respect to being F. Thus:

Immediately unsharp boundary for F, Minimal conception There is something such that it is indeterminate whether that thing is F. So a model represents F's having an immediately unsharp boundary if it is a model of $\exists x \mathcal{I} F x$.

But we also considered a less well defined notion of a "lack of any sharp boundaries" which requires an immediately unsharp boundary between F and non-F, an immediately unsharp boundary between things that are determinately F and things which are not, and so on. It is not obvious how the "and so on" is to be filled in. Let us consider filling it in this way:

Grand lack, conception one There is a grand lack of sharp boundary between F's and non-F's just in case all of the following are true:

There is an immediately unsharp boundary for the F's and for the non-F's.

There is an immediately unsharp boundary for the determinately F's and for the determinately non-F's.

There is an immediately unsharp boundary for the determinately determinately F's and for the determinately determinate non-F's.

and so on...

Note that the grand lack is here conceived in terms of the immediately unsharp boundary and the notion of determinacy. Hence a different conception of the immediately unsharp boundary would yield a different specification of the grand lack, through the above conception.

An immediately unsharp boundary requires only one thing

Having fixed our Minimal Assumptions about the semantics of our language, let us consider what must be, given only that there is an immediately unsharp boundary for the predicate F. We have said that for that to be is for there to be at least one thing a such that a satisfies $\mathcal{I}Fx$ (given our assumptions, this means just that $\exists xFx$ is true). Thus it looks as though there might be a vague predicate with an immediately unsharp boundary and just one thing.

A grand lack with just one thing

Given the Minimal conception of the immediately unsharp boundary and our first conception of the grand lack of sharp boundary, we get that a model represents a grand lack of sharp boundary for F just in case it makes all of the following true:

$$\exists x\mathcal{I}Fx \,\&\, \exists x\mathcal{I}{\sim}Fx$$
$$\exists x\mathcal{I}\mathcal{D}Fx \,\&\, \exists x\mathcal{I}\mathcal{D}{\sim}Fx$$
$$\exists x\mathcal{I}\mathcal{D}\mathcal{D}Fx \,\&\, \exists x\mathcal{I}\mathcal{D}\mathcal{D}{\sim}Fx$$
$$\vdots$$

Given the current conception of the grand lack of sharp boundary for F, how many objects must there be for there to be a grand lack of sharp boundary between F's and non-F's? The answer is "one". For a single object can witness all of the infinitely many required sentences. Considering the first, it is easy to see that our Minimal Assumptions allow that a single object could satisfy both conjuncts; indeed I1 and I2 require it. Now consider whether that object could satisfy the first conjunct of the second sentence. Suppose the object were named b. Then we would need to have

$$\mathcal{I}Fb \,\&\, \mathcal{I}\mathcal{D}Fb$$

that is,

$$\sim\mathcal{D}Fb \,\&\, \sim\mathcal{D}\sim Fb \,\&\, \sim\mathcal{D}\mathcal{D}Fb \,\&\, \sim\mathcal{D}\sim\mathcal{D}Fb$$

be true. The simultaneous truth of both the first and fourth conjuncts is the only odd thing here, for it requires that something both be the case and be not determinately the case. But our minimal assumptions do not rule that out. Thus we can have one object witness both the first conjuncts of the first two sentences required for the grand lack of sharp boundary. Similar comments hold for satisfying the second conjuncts of the two sentences. As for an object's witnessing both conjuncts of the second sentence—$\mathcal{I}\mathcal{D}F$ and $\mathcal{I}\sim\mathcal{D}F$—there is no conflict with the minimal assumptions. And since the two conjuncts of the first sentence are equivalent, no further considerations are needed to show that there could be an object that simultaneously witnesses both the first two sentences of the series.

Similar considerations will reveal that it is possible for a single object to witness all of the infinite series of sentences. But the considerations are rather complicated, and a better way to see the fact is by building a model in a system that conforms to our Minimal Assumptions about the semantics. That can be done, but the apparatus for doing so will not be presented until 4. The Boundary Semantics (section 4) conforms to the Minimal Assumptions, and it easy to construct a model that uses one object to witness all of the sentences in the series.[7]

2.2.3 The Principle of Instantiated Tripartition

It might be argued that for there to be an unsharp boundary for a vague predicate, there must be something that satisfies it, something that dissatisfies it, and also something indeterminate with respect to it. So consider this schematic principle:

---

[7]The construction of the model is given here: Let n be a classical model with one object b in its universe and that assigns $\{b\}$ to F. Let m be just like n except that it assigns $\emptyset$ to F. Let the Boundary Model b be as follows: $b_1 = \{n, m\}$ and $b_2 = \{\{n\}, \{m\}, \{n, m\}\}$; in general, $b_{i+1} = \mathcal{P}(b_i) - \emptyset$ (for $i > 0$). It does not matter which of n and m we choose for b_0. It is fairly easy to see that b is a model for every sentence in the series.

Principle of Instantiated Tripartition If there is an immediately unsharp boundary between the ϕ's and the non-ϕ's, then there are three *distinct* things x, y, z with the following character:

$$\begin{cases} x \text{ is determinately } \phi; \\ y \text{ is neither determinately } \phi \text{ nor determinately not } \phi; \\ z \text{ is determinately not } \phi. \end{cases}$$

We may picture the typical situation thus:

$$\longleftarrow \mathcal{D}\phi \longrightarrow\longleftarrow \mathcal{I}\phi \longrightarrow\longleftarrow \mathcal{D} \sim\phi \longrightarrow$$

where it is understood that there is at least one thing in each of the three partitions (determinately ϕ, indeterminate with respect to ϕ, and determinately not ϕ). Our picture suggests what seems to be the case, that with respect to being ϕ, the things form a space on which there is some notion of being more or less ϕ, so that the left-to-right direction of the diagram would move us from more ϕ to less ϕ.

On this Principle, there must be at least three objects in order for there to be an immediate lack of sharp boundary between the F's and the non-F's. But there need be no more.

What happens when we turn to the grand lack of sharp boundary, conceived as the existence of infinitely many immediately unsharp boundaries, as in 2.2.2? Consider an immediately unsharp boundary between the determinately red and the not determinately red. For this, by our Principle, there must be three things: one determinately determinately red, one indeterminate with respect to determinate redness, and a third determinately not determinately red. Similarly if we think that there is no sharp line between the determinately not red and the things which are otherwise. And so if we think that each of the two boundaries introduced when we say that there is no sharp line between the red and the not red is itself unsharp, then it looks as though there will be at least five things. Intuitively, the resulting five-way partition would look like this:

$$\leftarrow \mathcal{DD}F \rightarrow \leftarrow \mathcal{ID}F \rightarrow \leftarrow \mathcal{DI}F \rightarrow \leftarrow \mathcal{ID} \sim F \rightarrow \leftarrow \mathcal{DD} \sim F \rightarrow$$

(where "F" represents "is red".) But our picture assumes some things; for example, we have assumed that it is not the same thing that is indeterminate with respect to being determinately red and that is indeterminate with respect to being determinately not red. (Further, it suggests that the five partitions fall into a linear ordering with respect to being more or less ϕ. For example, something that is determinately determinately ϕ is more ϕ than something that is indeterminate with respect to being determinately ϕ.)

Yet a grand lack of sharp boundaries does not entail that there be more than three objects when the immediately unsharp boundaries are governed by the Principle of Instantiated Tripartition. Let me illustrate. Suppose that objects a, b, and c witness the instance of the Principle for the immediately unsharp boundary of the predicate "red". Now consider the x, y, and z that witness the instance of the Principle for the predicate "determinately red". There is no obvious reason why it cannot be that x, y, and z are a, b, and c. For it seems coherent on the face of it that a be both determinately red and determinately determinately red, that b be both indeterminate with respect to being red and indeterminate with respect to being determinately red, and that c be both determinately not red and determinately determinately not red. The only odd one of these three propositions is the one that concerns b. Here the situation is just like that discussed above, regarding the very most basic conception of a lack of sharp boundary. Given the Minimal Assumptions, we can have a picture like this:

a	b	c
F	?	$\sim F$
$\mathcal{D}F$	$\mathcal{I}F \,\&\, \mathcal{I} \sim F$	$\mathcal{D} \sim F$
$\mathcal{DD}F$	$\mathcal{ID}F \,\&\, \mathcal{ID} \sim F$	$\mathcal{DD} \sim F$
$\mathcal{DDD}F$	$\mathcal{IDD}F \,\&\, \mathcal{IDD} \sim F$	$\mathcal{DDD} \sim F$
$\vdots$	$\vdots$	$\vdots$

The highly indeterminate b witnesses all of the lacks of determinacy, while the highly determinate a and c witness all of the determinate matters.[8]

[8]A boundary model for the envisaged situation is easy to construct: Let b be the boundary model used in section 2.2.2 to show that the simplest conception of the

2.2.4 A controversial assumption

It is interesting to consider what further assumptions would rule out the situation just pictured. The obvious assumption that would do this is:

$$\phi \models \mathcal{D}\phi$$

Were this correct, then if b satisfies $\sim\mathcal{D}F$, then b satisfies $\mathcal{D}\sim\mathcal{D}F$, and hence cannot satisfy $\mathcal{I}\mathcal{D}F \quad (= \ \sim\mathcal{D}\mathcal{D}F \,\&\, \sim\mathcal{D}\sim\mathcal{D}F)$. And it will be clear after the next subsection that this assumption entails that there will have to be infinitely many things, given a grand lack of sharp boundaries—whether or not we are also given the Principle of Instantiated Tripartition.

But this assumption is somewhat controversial, for if it is true, then if a model makes true $\mathcal{I}\phi$, then the model makes true neither ϕ nor $\sim\phi$. Thus given that something is indeterminate, a certain form of bivalence would fail, and there are respectable approaches to the logic of vagueness on which bivalence holds. There is another principle that is similar but not controversial in the same way, and that is the Principle of Paradigms.

---

grand lack of sharp boundaries does not require that there be more than one object. Now modify it as follows: let n be the classical model with the three objects a, b, c in its universe that assigns $\{a, b\}$ to F and let m be the classical model that is just like n except that it assigns $\{a\}$ to F. Let the structure of the boundary model be as before.

2.2.5 The Principle of Paradigms

This principle is much stronger than the Principle of Instantiated Tripartition, but is weaker than the controversial assumption just discussed.

The Principle of Paradigms If there is an object that satisfies ϕ, then there is an object that satisfies $\mathcal{D}\phi$.

If we combine this principle with the most straightforward conception of the immediately unsharp boundary, then if there is an immediately unsharp boundary between the F's and the non-F's, we get, at a minimum, a picture like this:

a	b	c
F	?	$\sim F$
$\mathcal{D}F$	$\mathcal{I}F$	$\mathcal{D} \sim F$
$\mathcal{D}\mathcal{D}F$	$\mathcal{D}\mathcal{I}F$	$\mathcal{D}\mathcal{D} \sim F$
$\mathcal{D}\mathcal{D}\mathcal{D}F$	$\mathcal{D}\mathcal{D}\mathcal{I}F$	$\mathcal{D}\mathcal{D}\mathcal{D} \sim F$
$\vdots$	$\vdots$	$\vdots$

We still do not need more than three objects.

But if we add that there is a *grand* lack of sharp boundary between F's and non-F's, we must have more objects. For example, there must be something in each of the five parts of the partition

$$\leftarrow \mathcal{D}\mathcal{D}F \rightarrow\leftarrow \mathcal{I}\mathcal{D}F \rightarrow\leftarrow \mathcal{D}\mathcal{I}F \rightarrow\leftarrow \mathcal{I}\mathcal{D} \sim F \rightarrow\leftarrow \mathcal{D}\mathcal{D} \sim F \rightarrow$$

Note that anything in the middle part cannot be in either of the parts adjacent to it; for example, if $\mathcal{IDF}b$ then $\sim\mathcal{D}\sim\mathcal{DF}b$, but if $\mathcal{DIF}b$ then $\mathcal{D}\sim\mathcal{DF}b$ (assuming what seems apt for inclusion among the minimal assumptions, that $\mathcal{D}$ "distributes over" &). (Oddly, it turns out that what we have assumed does not require that the things in the second and fourth parts be distinct.)

Since the grand lack of sharp boundary requires also that there be no sharp boundary between the $\mathcal{DDF}$ and the rest, there must be something that satisfies $\mathcal{IDDF}$, and hence, by the Principle of Paradigms, something that satisfies $\mathcal{DIDDF}$. It turns out that this thing cannot also satisfy any of $\mathcal{DDDF}$, $\mathcal{DDIF}$, $\mathcal{DDI}\sim F$, or $\mathcal{DDD}\sim F$.

That there must be infinitely many things (given the Principle of Paradigms, the Minimal Assumptions, and a grand lack of sharp boundaries for F) can be proved as follows. First we will need an auxilliary fact that we will use again later in our discussion.

Fact 1. Given the Minimal Assumptions, no object satisfies more than one of the following open formulas.

$$\begin{array}{ll} \phi_0 & \mathcal{DI}Fx \\ \phi_1 & \mathcal{DID}Fx \\ \phi_2 & \mathcal{DIDD}Fx \\ \vdots & \vdots \\ \phi_n & \mathcal{DID}^nFx \\ \vdots & \vdots \end{array}$$

Proof. First observe that for $m < n$, if an object satisfies $\mathcal{AD}^n\phi$, then it satisfies $\mathcal{AD}^m\phi$; this is evident from Minimal Assumption M4. Now suppose that object a satisfies ϕ_n and object b satisfies ϕ_m, with $m < n$. We show that $a \neq b$. Let ψ be $\mathcal{D}^mFx$. First, we have

$$\begin{array}{lr} \quad a \text{ satisfies } \phi_n & \\ \Rightarrow a \text{ satisfies } \mathcal{DID}^nFx & \\ \Rightarrow a \text{ satisfies } \mathcal{ID}^nFx & \text{by D1} \\ \Rightarrow a \text{ satisfies } \mathcal{ID}^{n-m}\psi & \text{by def. of } \psi \\ \Rightarrow a \text{ satisfies } {\sim}\mathcal{D}{\sim}\mathcal{D}^{n-m}\psi & \text{by def. of } \mathcal{I}\text{, M1, and C1.} \\ \Rightarrow a \text{ satisfies } \mathcal{AD}^{n-m}\psi & \text{by def. of } \mathcal{A} \\ \Rightarrow a \text{ satisfies } \mathcal{AD}\psi & \text{by the initial observation} \end{array}$$

Next, we have

$$
\begin{aligned}
& b \text{ satisfies } \phi_m & \\
\Rightarrow\ & b \text{ satisfies } \mathcal{DID}^m Fx & \\
\Rightarrow\ & b \text{ satisfies } \mathcal{DI}\psi & \text{by def. of } \psi \\
\Rightarrow\ & b \text{ satisfies } \mathcal{D}{\sim}\mathcal{D}\psi & \text{by def. of } \mathcal{I} \text{ and M2} \\
\Rightarrow\ & b \text{ satisfies } {\sim}{\sim}\mathcal{D}{\sim}\mathcal{D}\psi & \text{by DNI} \\
\Rightarrow\ & b \text{ satisfies } {\sim}\mathcal{AD}\psi & \text{by def. of } \mathcal{A}
\end{aligned}
$$

Thus a satisfies $\mathcal{AD}\psi$ and b satisfies ${\sim}\mathcal{AD}\psi$. Putting together E1, C1, and Con2, we see that since, if a and b were identical then something of the form $\phi \,\&\, {\sim}\phi$ would be true, $a \neq b$. Since n and m ($< n$) were arbitrary, we have that no object satisfies more than one of the ϕ_n. □

Combining the Principle of Paradigms with the current conception of the grand lack of sharp boundaries, each of the open formulas must be satisfied by some object. Since they are all satisfied, there must be infinitely many objects.

2.2.6 A second conception of the grand lack of sharp boundary

Let us consider now a slightly different way of casting the felt lack of any sharp boundary for a predicate F. We rendered the immediately unsharp boundary for the F's minimally as the existence of something indeterminate with respect to being F. This brings to mind three new mutually exclusive predicates or "classifications", as I will call them, beyond being F. They are: being determinately F, being determinately not-F and being indeterminate with respect to being F. Thus there is pattern of three-from-one, which (using symbols now) takes us from the one classification F to the three classifications $\mathcal{D}F$, $\mathcal{D}{\sim}F$, and $\mathcal{I}F$. We will now consider a conception of the grand lack of sharp boundary on which this pattern iterates indefinitely; there are immediately unsharp boundaries for F, the first three classifications that grow from F, the nine classifications that grow from those three, and so on.

Grand lack, conception two There is a grand lack of sharp boundary between F's and non-F's just in case there is an immediately unsharp boundary for all of the following:

The F's.

The determinately F's, the things which are indeterminate with respect to F, and the determinately non-F's.

The determinately determinately F's, the things indeterminate with respect to being determinately F, and the determinately not determinately F's; the things determinately indeterminate with respect to being F, the things indeterminate with respect to being indeterminate with respect to being F, and the things determinately not indeterminate with respect to being an F; the determinately determinately non-F's, the things indeterminate with respect to being determinately non-F, and the determinately not determinately non-F.

and so on...

Let us suppose that immediately unsharp boundaries are governed by the

Principle of Instantiated Tripartition (2.2.3). Then, given a grand lack of sharp boundary for F on the second conception, and treating the language as we have been above, we can use Fact 1 to show that there must be infinitely many things, for all of the open sentences $\mathcal{DID}^n Fx$ must be satisfied.

2.2.7 How one is apt to picture things

When we imagine a series of things ranging steadily and slowly from red to not red, we are apt to imagine the unsharpness of the boundary of the red by rejecting a picture like this:

$$\longleftarrow\!\!\!-\!\!\!-\!\!\!-\; 1 \;-\!\!\!-\!\!\!-\!\!\!\longrightarrow\!\longleftarrow\!\!\!-\!\!\!-\!\!\!-\; 2 \;-\!\!\!-\!\!\!-\!\!\!\longrightarrow$$

in favor of a picture like this:

$$\longleftarrow\; 1 \;\longrightarrow\!\longleftarrow\; 2 \;\longrightarrow\!\longleftarrow\; 3 \;\longrightarrow$$

That is, we represent the unsharpness of the boundary of the red things by replacing a picture of a simple division of things into two (red and not-red) in favor of a three-way partition. We then have the idea that the two boundaries introduced in the three-way partition are also unsharp, and so should also be pictured as having an in-between region of indeterminacy. And we think that this sort of unsharpness may hold of the resulting boundaries, and so on. Throughout we imagine things laid out in order, and have the idea that the ordering has some significance like this: the things to the left are more red than the things to the right.

Thus we might wonder whether we can capture these quasi-geometrical ideas—the various partitions—with the "classifications" formed with the determinacy operator, negation, and so on (like "indeterminate with respect to redness"). If this is possible, it must be that many of classifications are related in a way that corresponds to the "is more red than" relation among things. And it looks as though many are so ordered; for example, it seems that something that is indeterminate with respect to being red must be at least as red as something that is determinately not red, and at most as red as something that is determinately red.

But it is not the case that any two classifications are related in this way. Consider something that is indeterminate with respect to being indeterminate with respect to being red—i.e., something that satisfies $\mathcal{I}\mathcal{I}Fx$. Intuitively, such a thing could be either more or less red than something that satisfies $\mathcal{D}\mathcal{I}Fx$. In effect, it could be on "either side" of a thing that is determinately indeterminate with respect to being red.

Nevertheless there would seem to be some interesting patterns among the classifications. Symbolizing "x is at least as red as y" as $x \geq_F y$, the following are plausible:

$$\mathcal{D}^n Fx \,\&\, \mathcal{D}^m Fy \quad \models \quad x \geq_F y \qquad (n \geq m)$$

$$\mathcal{D}^n{\sim}Fx \,\&\, \mathcal{D}^m{\sim}Fy \quad \models \quad y \geq_F x \qquad (n \geq m)$$

and also

$$\mathcal{D}Fx \,\&\, \mathcal{I}Fy \quad \models \quad x \geq_F y$$

$$\mathcal{I}Fx \,\&\, \mathcal{D}{\sim}Fy \quad \models \quad x \geq_F y$$

Now let us see whether we can express with "classifications" the quasi-geometrical intuitions about the unsharp boundaries. We revise a picture like this:

$$\longleftarrow\!\!\!-\!\!-\!\!- \text{ red } -\!\!-\!\!-\!\!\!\longrightarrow\!\longleftarrow\!\!\!-\!\!-\!\!- \text{ not red } -\!\!-\!\!-\!\!\!\longrightarrow$$

to a tripartition:

$$\longleftarrow\!\!- \; 1 \; -\!\!\longrightarrow\!\longleftarrow\!\!- \; 2 \; -\!\!\longrightarrow\!\longleftarrow\!\!- \; 3 \; -\!\!\longrightarrow$$

The three classifications $\mathcal{D}F$, $\mathcal{I}F$, and $\mathcal{D}{\sim}F$ are natural renderings of the three regions, and they do seem to come naturally ordered by the "is at least as red as" relation. But we then regard each of the two boundaries in that picture as unsharp in the same way that the boundary pictured by the first picture, which suggests a picture like this:

$$\longleftarrow 1 \longrightarrow\!\longleftarrow 2 \longrightarrow\!\longleftarrow 3 \longrightarrow\!\longleftarrow 4 \longrightarrow\!\longleftarrow 5 \longrightarrow$$

The next step would be to regard each of the four boundaries in this last picture as unsharp, producing a picture with nine regions. But even at this stage it is not clear what classifications go with each of the five regions.

Suppose we focus on the transformation from

$$\longleftarrow\!\!\relbar\!\!\relbar\; F \;\relbar\!\!\relbar\!\!\longrightarrow\!\longleftarrow\!\!\relbar\!\!\relbar\!\!\relbar$$

to

$$\longleftarrow\; \mathcal{D}F \;\longrightarrow\!\longleftarrow\; \mathcal{I}F \;\longrightarrow\!\longleftarrow\; \mathcal{D}{\sim}F \;\longrightarrow$$

and then look at the left-most of the three classifications in the three-way division, then classifications 1 through 3 of the five-way division would seem to be $\mathcal{D}\mathcal{D}F$, $\mathcal{I}\mathcal{D}F$, and $\mathcal{D}{\sim}\mathcal{D}F$. These three do naturally fall under, in the desired way, the "is at least as red as" relation.

But there are problems. If we look at the second of the three classifications in the three-way partition, and focus on its unsharp left-side boundary, we would want to transform

$$\relbar\!\!\relbar\!\!\relbar\!\!\longrightarrow\!\longleftarrow\!\!\relbar\!\!\relbar\!\!\relbar\,\mathcal{I}F\,\relbar\!\!\relbar\!\!\relbar\!\!\longrightarrow$$

to

$$\longleftarrow\; \mathcal{D}{\sim}\mathcal{I}F \;\longrightarrow\!\longleftarrow\; \mathcal{I}\mathcal{I}F \;\longrightarrow\!\longleftarrow\; \mathcal{D}\mathcal{I}F \;\longrightarrow$$

But if 1 through 3 of the five-way partition are $\mathcal{D}\mathcal{D}F$, $\mathcal{I}\mathcal{D}F$, and $\mathcal{D}{\sim}\mathcal{D}F$, then where will the above three classifications go? And notice also that these three classifications are not naturally ordered by the "is at least as red as" relation.

Thus the quasi-geometrical idea that the left one of the two boundaries pictured in this picture

$$\longleftarrow\; 1 \;\longrightarrow\!\longleftarrow\; 2 \;\longrightarrow\!\longleftarrow\; 3 \;\longrightarrow$$

is unsharp seems to be ambiguous: there is the unsharpness of the right boundary of class 1, and the unsharpness of the left boundary of class 2, and each seems to produce its own classifications.

Let us take a different tack. What we need is a way of thinking of the transformation from a (unsharp) boundary

$$\xrightarrow{\hspace{4cm}}\xleftarrow{\hspace{4cm}}$$

to the resulting structure

$$\xrightarrow{\hspace{2cm}}\xleftarrow{\hspace{2cm}}\;\xrightarrow{\hspace{2cm}}\xleftarrow{\hspace{2cm}}$$

in terms of classifications, a way that gives the same result no matter which side we "approach" from, that intuitively corresponds with the notion of the unsharpness of the boundary, and that yields classifications which preserve the "ordering" implicit in the pictures.

2.2.8 A third conception of the grand lack of sharp boundaries

Here is one way to try to do it. We will think of the transformation as taking us from

$$\longleftarrow\!\!\!-\!\!-\; \phi \;-\!\!-\!\!\!\longrightarrow\longleftarrow\!\!\!-\!\!-\; \psi \;-\!\!-\!\!\!\longrightarrow$$

to

$$\longleftarrow \mathcal{D}\phi \longrightarrow\longleftarrow \mathcal{A}\phi \,\&\, \mathcal{A}\psi \longrightarrow\longleftarrow \mathcal{D}\psi \longrightarrow$$

The outermost of the resulting classifications arise from their parents in the original picture in the same way, and so things work out nicely so that we can simultaneously transform all of the boundaries of a picture. For example, if we transform the two boundaries of

$$\longleftarrow 1 \longrightarrow\longleftarrow 2 \longrightarrow\longleftarrow 3 \longrightarrow$$

the rightmost result of transforming the 1/2 boundary will be the same as the leftmost result of transforming the 2/3 boundary. Thus the notion of simultaneously transforming all of the boundaries will make sense.

And it looks plausible that our transformation preserves order. For example, if ϕ and ψ are ordered by the "is at least as red as" relation (i.e., if it is gauranteed that every ϕ is at least as red as every ψ) then

$$\longleftarrow \mathcal{D}\phi \longrightarrow\longleftarrow \mathcal{A}\phi \,\&\, \mathcal{A}\psi \longrightarrow\longleftarrow \mathcal{D}\psi \longrightarrow$$

should represent three classifications that are similarly ordered. Further, the transformation seems to answer to our intuitive idea that the ϕ/ψ boundary is unsharp.

This solution to the problem of casting the quasi-geometrical intuition of the boundary in terms of the determinacy operator and other logical operators can be used to give us a third conception of the grand lack of sharp boundaries for a predicate F. But first let us be precise about the solution. We may describe it as follows. First, recursively define a notion of "classification based on a predicate F" ("classification", for short) as follows: Fx itself is a classification; if ϕx is a classification, then so is the result of prefixing to ϕx any of the operators $\mathcal{D}$, $\mathcal{A}$, and $\sim$; and if ϕx and ψx are two classifications, then so is the string $(\phi x \,\&\, \psi x)$. Thus classifications are basically open formulas with no quantifiers, and only the variable x.

We now define recursively a notion of an "approximation of a grandly unsharp boundary" ("approximation" for short). Thus: the ordered 2-tuple $(Fx, {\sim}Fx)$ is an approximation; and if an ordered n-tuple $(c_1, c_2, \ldots, c_n)$ is an approximation, then so is the ordered $(2n-1)$-tuple $(d_1, d_2, \ldots, d_{2n-1})$ that arises from it as follows:

$$d_n = \begin{cases} \mathcal{D}c_{\frac{i+1}{2}} & \text{for odd } i \\ (\mathcal{A}c_{\frac{i}{2}} \,\&\, \mathcal{A}c_{\frac{i+2}{2}}) & \text{for even } i \end{cases}$$

Let us call $(Fx, {\sim}Fx)$ the *zeroth* approximation, $(\mathcal{D}Fx, \mathcal{A}Fx \,\&\, \mathcal{A}{\sim}Fx, \mathcal{D}{\sim}Fx)$, the *first* approximation, and so forth. We may now state the

Grand lack, conception 3 For every $n \geq 0$, for every classification c in the n^{th} approximation, $\exists xc$ is true.

Fact 2. On the third conception of the grand lack of a sharp boundary, a grand lack entails that there are infinitely many things.

Proof. Once again, we use Fact 1. We just need to show that all of the open sentences $\mathcal{DID}^n Fx$ are satisfied. Let ϕ_n denote the second member of the n^{th} approximation. Our first step is to show that

$$\text{For } n \geq 0,\ \ \phi_{n+1} \models \mathcal{ID}^n Fx$$

(Here, the "$\models$" relation relates a pair of open sentences if on every model and every assignment, if the first sentence is true, so is the second, and $\mathcal{D}^0 Fx = Fx$.)

We show this by induction on n.

Base case: $n = 0$

$$\begin{array}{rll} \phi_1 = & \mathcal{A}Fx \,\&\, \mathcal{A}{\sim}Fx & \\ \models & {\sim}\mathcal{D}{\sim}Fx \,\&\, {\sim}\mathcal{D}{\sim}{\sim}Fx & \text{unabbreviating } \mathcal{A} \\ \models & {\sim}\mathcal{D}{\sim}Fx \,\&\, {\sim}\mathcal{D}Fx & \text{by DNE} \\ \models & \mathcal{I}Fx & \text{def. of } \mathcal{I} \end{array}$$

Induction step: We assume for n and show for $n + 1$. The induction hypothesis is that $\phi_{n+1} \models \mathcal{I}\mathcal{D}^n Fx$.

$$
\begin{aligned}
\phi_{n+2} = \quad & \mathcal{A}\mathcal{D}^{n+1}Fx \,\&\, \mathcal{A}\phi_{n+1} && \text{construction of approximations} \\
\models \quad & \mathcal{A}\mathcal{D}^{n+1}Fx \,\&\, \mathcal{A}\mathcal{I}\mathcal{D}^n Fx && \text{by induction hyp., C1, and A2} \\
\models \quad & \mathcal{A}\mathcal{D}^{n+1}Fx \,\&\, \mathcal{A}{\sim}\mathcal{D}\mathcal{D}^n Fx && \text{by def. of } \mathcal{I} \text{ and C1 and A2} \\
\models \quad & {\sim}\mathcal{D}{\sim}\mathcal{D}^{n+1}Fx \,\&\, {\sim}\mathcal{D}{\sim}{\sim}\mathcal{D}\mathcal{D}^n Fx && \text{unabbreviating } \mathcal{A} \\
\models \quad & {\sim}\mathcal{D}{\sim}\mathcal{D}^{n+1}Fx \,\&\, {\sim}\mathcal{D}\mathcal{D}\mathcal{D}^n Fx && \text{by DNE and A2} \\
\models \quad & \mathcal{I}\mathcal{D}^{n+1}Fx && \text{def. of } \mathcal{I}
\end{aligned}
$$

Thus $\phi_{n+1} \models \mathcal{I}\mathcal{D}^n Fx$. Now observe that the third element of the $(n+1)^{\text{th}}$ approximation is $\mathcal{D}\phi_n$. Let ψ_n be the third member of the $(n+1)^{\text{th}}$ approximation. Now by M3 and the above fact, $\psi_n = \mathcal{D}\phi_n \models \mathcal{D}\mathcal{I}\mathcal{D}^n Fx$. Since on the third conception of the grand lack of sharp boundary every member of every approximation is satisfied, all of the sentences of Fact 1 are satisfied. □

Thus on the third conception of the grand lack, the grand lack entails infinitely many things.

2.2.9 Summary of results about the "How many?" question

We have seen that with the Background Assumptions and the Minimal Assumptions in place, we get the following results:

- Given the Minimal account of the immediately unsharp boundary, the existence of an immediately unsharp boundary for the F's requires only that one thing exist.

- Given the first conception of the grand lack of sharp boundary, and nothing more, there need be only one thing.

- Given the Principle of Instantiated Tripartition and a grand lack on the first conception of the grand lack of sharp boundary, there need be only one thing.

- Given the Principle of Paradigms and a first conception grand lack, there must be infinitely many things.

- Given a grand lack of sharp boundary on the second conception, plus the Principle of Instantiated Tripartition, there must be infinitely many things.

- Given a grand lack on the third conception, there must be infinitely many things.

2.2.10 Remarks about the "How many?" question

Our discussion so far has shown that when we render the notion of the unsharp boundary with the notion of determinacy, natural assumptions lead to either:

the somewhat unintuitive idea that there could be a grandly unsharp boundary for a predicate F and yet there be only one thing; or the seemingly problematic idea that if there is a grandly unsharp boundary for a predicate F, there must be infinitely many things.

The idea of the first disjunct is unintuitive because the familiar sorts of situation that make us want to say that there is an unsharp boundary for a predicate are ones that involve a great many things. We imagine thousands of colored tiles, a long series of men, each with one less hair than the last, and the like. It is surprising that all of the many "layers" of indeterminacy that the grand lack of sharp boundary seems to require should be present in a single thing.

The idea of the second disjunct is problematic, for if grand lacks of sharp boundary require the existence of infinitely many things, then a world that contains finitely many colored tiles is not a world in which there is a grandly unsharp boundary for the red things. And since it is quite conceivable that the real world in fact contains only finitely many relevant objects, it looks as though the assertion that there are grandly unsharp boundaries is much more contentious than it would have seemed. Thus if we are to take seriously the idea that there be grand lacks of sharp boundary, then we should be prepared to reject the assumptions which lead to the grand lack's entailing the existence of infinitely many objects.

Should the existence claims be read possibilistically?

This raises the question whether we should read all of the propositions under discussion as being about possible objects (so to speak) rather than actual ones. For example, the grand lack of sharp boundary for "red" would entail only that it be possible, for each of the infinitely many classifications of 2.2.8, that there be something in that classification. Thus, conceptions of the grand lack that entail that there be infinitely many things would entail only that there be infinitely many existential possibilities, and hence not that there *actually* be infinitely many things.

While this reading takes the bite out of the entailment of infinity, it does not seem to me to be a satisfactory reading of what we have in mind when we think that there is a (grand) lack of sharp boundary for a predicate like "is a red tile in the series". When we look at the actual series of 10,000 tiles, we think that there is a (grand) lack of sharp boundary for the red tiles of the series, and that this has nothing to do with what (other) things could exist, but only with what actually exists. It is worth noting that the possibilistic reading might be appropriate for the thought that there is no sharp boundary for *the color red* or perhaps for the predicate "is red", considered in the right way. We do have an intuition that "is red" would still be *vague* even if all red objects went out of existence. But it seems to me that our sense of the lack of sharp boundary bears directly on the *actual* red things, even if there turn out to be finitely many of them.

2.2.11 An argument against certain gappy logics for vagueness

The facts about the "How many?" question can be used to make an argument against certain approaches to the logic of vague language: those on which

$$\phi \models \mathcal{D}\phi$$

holds. For if that holds, then given a grand lack of sharp boundary (with any conception of the grand lack, and with the minimal conception of the immediately unsharp boundary) there must be infinitely many things. To see this, consider again Fact 1). On any conception of the grand lack, given the current assumption, all of the sentences in the crucial series would have to be true, and so the proof would go through as before.

Thus on any gappy approach that goes for the above assumption, a grand lack of sharp boundary would require infinitely many things. But this is an intolerable requirement, so those approaches are intolerable.

The current observation has an interesting corollary. Since

$$\phi \models \mathcal{D}\phi$$

has intolerable consequences, we will want to assert that it is not the case. But if it is not the case, then there must be a model $\mathfrak{M}$ such that $\mathfrak{M} \models \phi$ but $\mathfrak{M} \not\models \mathcal{D}\phi$. This model represents a situation in which ϕ is true and yet $\mathcal{D}\phi$ fails to be true. That there could be such a situation is puzzling. (Such puzzling situations will be addressed further in section 3.1.)

2.2.12 Morals of this section

We set out to try to capture our sense that the lack of sharp boundary for the F's is not exhausted merely by the existence of things that are neither determinately F nor determinately not F. We saw that some of the most intuitive ways of making out our sense that there is *no* relevant sharp boundary for the F's lead to there being infinitely many things. This result seems unacceptable, or counter-intuitive at the very least. There are not infinitely many tiles in the example of the 10,000 tiles, but we want to say that there is no sharp boundary for the red tiles. What should we make of this?

There are a number of options. Two that are worth considering are as follows. First, the formulations of what it amounts to for there to be no sharp boundary for the F's that entail infinitely many things might be wrong as formulations of that intuitive idea. Despite the way we are apt to picture things, it may be that the third conception of the grand lack is simply too strong, and the fact that it entails infinitely many things shows that to be so.

Second, we might reconsider the idea that the notion of no sharp boundary for the F's can be captured properly with the determinacy operator and the standard logical apparatus. It may be that despite the fact that a determinacy operator is (at present) more formally tractable than an irreducible notion of a lack of sharp boundary, the notion of the lack of sharp boundary is really irreducible, and cannot be fully captured with such an operator. It is not clear where this path leads, and it will not be further explored in this dissertation.[9]

In the rest of this dissertation, we will focus primarily on the determinacy operator instead of the notion of the lack of sharp boundary, chiefly because of its formal tractability and its prominence in the literature. And it should be noted that even if the second, more radical, option is on the right track, lacks of determinacy are still likely to be central to, if not at the very center, of vagueness. For if every tile is either determinately red or determinately not red, what sense could there be to there being no sharp boundary for the red

[9]In recent conversation, Kit Fine expressed something like this view, and suggested that it might lie close to the truth. As far as I know, he has not yet worked out where the path leads.

things?

Chapter 3

Indeterminacy and logic

This chapter addresses some general issues regarding logic, semantics, and vague language. The indeterminacy that is characteristic of vague language, it is argued, is a deep phenomenon that inevitably leads to puzzlement. In sections 3.2 and 3.3, it is suggested that we must make a choice between two different sorts of puzzling propositions: either classical reasoning is not valid, or there must be true instances of a certain puzzling scheme. The second of these two options is embraced by an approach to the logic of vagueness I call the "Pure Qualification Approach". This and other approaches are discussed in section 3.4 and arguments in favor of it are given in 3.5.

3.1 The puzzling schemes and inferences

The puzzling schemes and inferences are closely related to the Tarski (T) biconditionals. Consider the four inference patterns

$$
\begin{array}{l}
(1)\ \phi \therefore \mathcal{T}\phi \\
(2)\ {\sim}\phi \therefore \mathcal{F}\phi \\
(3)\ {\sim}\mathcal{T}\phi \therefore {\sim}\phi \\
(4)\ \mathcal{F}\phi \therefore {\sim}\phi
\end{array}
$$

where $\mathcal{T}$ represents "it is true that" and $\mathcal{F}$ represents "it is false that".

Inference patterns (1) and (2) seem incontrovertible. Pattern (3) initially seems plausible, and can be thought of as a contrapositive form of (1). Yet if we allow all (1)–(3), we run into serious problems in understanding how it could be, for any proposition ϕ, that it is neither true nor false that ϕ. For we have

$$
\begin{array}{rcll}
{\sim}\mathcal{T}\phi \,\&\, {\sim}\mathcal{F}\phi & \Rightarrow & {\sim}\mathcal{T}\phi & \\
& \Rightarrow & {\sim}\phi & \text{by (3)} \\
& \Rightarrow & \mathcal{F}\phi & \text{by (2)}
\end{array}
$$

Thus it seems that

$$
{\sim}\mathcal{T}\phi \,\&\, {\sim}\mathcal{F}\phi \quad \Rightarrow \quad \mathcal{F}\phi \,\&\, {\sim}\mathcal{F}\phi
$$

which strongly suggests that it cannot be the case that

$$\sim\mathcal{T}\phi \,\&\, \sim\mathcal{F}\phi$$

The Gappy approaches to the logic of vagueness reject inference (3), though they accept inference (1). The idea that the latter inference is valid while the former is not is rather puzzling, since the former is just a contrapositive form of the latter, and this sort of contraposition, intuitively, seems valid.

But if we do accept all three patterns of inference, it looks as though we will be unable to hold that when a proposition is indeterminate, it is neither true nor false. This suggests (but does it entail?)[1] that we will be forced to concede that indeterminate propositions are either true or false. But this raises a new puzzle. For if we have

$$\mathcal{I}\phi \,\&\, \mathcal{I}\sim\phi$$

and we also have

$$\mathcal{T}\phi \vee \mathcal{F}\phi$$

then by a standard argument by cases and schemes (1) and (4), we get

$$\phi \vee \sim\phi$$

[1]The thought that there is an entailment here seems to require contraposition, which is not admitted as always valid by practitioners of gappy logic.

and by another argument by cases,

$$\phi \,\&\, \mathcal{I}\phi \vee \sim\phi \,\&\, \mathcal{I}\sim\phi$$

Hence we get that there is some proposition of the form $\psi \,\&\, \mathcal{I}\psi$ which is the case. But puzzling too: how could this possibly be the case?

3.2 The Fork In The Logical Road

The thesis here is that if we are adequately to describe the logic and semantics of a typical vague language then either classical logic fails in our language—the language we use to describe the object language—or we must accept that there are true propositions, expressible in our language, of the form $\phi \,\&\, \mathcal{I}\phi$.

The fundamental proposition the point depends upon is this:

The Fundamental Proposition The typical vague predicate F has the following property: there is vagueness about which things F is true of.

Consider again the series of 10,000 pairwise visually indistinguishable color tiles of (2.1). Suppose that the sentence "This tile is red." is written on the back of each tile. Now consider the question whether the red tiles have a sharp boundary. It is one of the central intuitions that make vagueness interesting that there is only an unsharp boundary between the red and the rest of the tiles. But consider the series of sentences on the backs of the tiles (taken in the same order we take the tiles in). Consider the question whether the true sentences have a sharp boundary. We seem to have just as much reason to believe that they do not as to believe that the red tiles do not have a sharp boundary.

Note that this is not to assume that a tile is red just in case the sentence on the back of it is true and that a tile is not red just in case its sentence is not true. Let it be that there are tiles whose sentences are not true but which do not satisfy "is not red", as is the case according to the gappy approaches to the logic of vagueness. Let that be, and the point remains.

The plausibility of the Fundamental Proposition does not flow from any special theoretical considerations. Rather, it flows from a direct consideration of the series of objects. One feels no relevant difference between the question whether there is a sharp boundary for the red tiles and the question whether there is a sharp boundary for the true sentences. The proposition that there is an unknown sharp boundary for the true sentences—an unknown number n such that it is a perfectly straightforward determinate fact that sentence n is true and sentence $n + 1$ is not—seems just as contrary to our sense of vagueness as does the proposition that there is an unknown sharp boundary between the red tiles and the non-red tiles.

There may be differences between "is red" and "is true"; perhaps it does not make sense that sentences earlier in the series are "truer than" later sentences, while it does seem to make sense that earlier tiles are redder than later ones. But whether there is such a comparative has no obvious impact on the issue whether there is a sharp boundary.

I have no *argument* for the Fundamental Proposition, just as I have no argument for the proposition that there is no sharp boundary for the red things. I do not think there could be a *convincing argument* for either claim of unsharpness; any valid argument for the unsharpness will have premises that are no more plausible than the conclusion. There may well be more convincing pieces of rhetoric than what I have given, but I will rest my case. Let us take the Fundamental Proposition for granted now, and see what follows from it.

3.2.1 The simplest form of the argument

It is fairly easy to see that if we want to talk about tiles and redness, we will have to either reject classically endorsed principles of reasoning (e.g. *reductio*) or accept that there are true instances of the scheme $\phi \,\&\, {\sim}\mathcal{D}\phi$ in our own language. For suppose that there is some tile t that is indeterminate with respect to redness. Thus it is not determinate that t is red, and it is not determinate that t is not red.[2]

But since it is, by assumption, indeterminate whether t is red, and indeterminate whether t is not red, either t is red and it is indeterminate whether t is red, or t is not red and it is indeterminate whether t is not red. Either way, it seems we have a true instance of a puzzling scheme, and hence we have reason to believe that there are true instances of a puzzling scheme.

Exactly parallel arguments can be given regarding the 10,000 sentences "This tile is red". Given that there is a lack of determinacy that flows from the unsharpness of the boundary of the true sentence, there is at least one sentence s of which it is indeterminate whether that sentence is true. We can argue that either s is true or s is not true, and we know that it is indeterminate whether s is true and whether s is not true.

To reject these arguments requires rejecting classically endorsed principles of reasoning. So if we are willing to talk about the truth of sentences, we are faced with either rejecting classical logic in our own reasoning about the

---

[2]It is possible to *argue* that either t is red or t is not red: if not, then if t were red, it would be the case that t is red or t is not red, and so it would not be the case that it is not the case that (either t is red or t is not red), so it thus must be that t is not red (since it cannot be, given our initial assumption, that t is red), and yet then either t is red or t is not red, and hence a contradiction follows from the assumption that it is not the case that either t is red or t is not red. Therefore it is not the case that it is not the case that either t is red or t is not red; hence either t is red or t is not red.

sentences, or accepting that there be true instances of a puzzling scheme in our own language.

3.2.2 An argument about models

Suppose that we undertake to "represent" with models the semantics of a language that contains the ten thousand sentences. We define some notion of a model and some notion of "truth at" a model. Suppose we now conjecture that there is some model $\mathfrak{M}$ such that for every sentence ϕ from the ten thousand, ϕ is true just in case $\mathfrak{M} \models \phi$. (The latter means that the sentence is "true at" the model.)

Now there is a short, classically endorsed path that leads to instances of a puzzling scheme. For let ϕ be one of the ten thousand sentences, one which is neither determinately true nor determinately not true. Now consider $\mathfrak{M}$, the model conjectured above. If $\mathfrak{M} \models \phi$, then by assumption ϕ is true, and yet it is not determinate that ϕ is true. And if $\mathfrak{M} \not\models \phi$ then ϕ is not true, yet it is not determinate that ϕ is not true.

So again, on any notion of model and "truth at" a model, if there is to be a model such that truth (for the ten thousand sentences) coincides with truth in the model, then either classically endorsed reasoning is incorrect, or we must accept that there are true instances of a puzzling scheme.

3.3 The difficulty of choosing a path

The puzzling schemes are obviously puzzling, and it is clear that accepting classical logic for a vague language is not a path free of mystery. But classical reasoning, in particular the steps of reasoning that lead to a conclusion of the form $\phi \vee \sim\phi$, is intuitively compelling. This author finds the idea that that reasoning is not completely reliable to be at least as mysterious as the idea that there are true instances of one of the puzzling schemes.

It is also worth pointing out a couple of things concerning models.

3.3.1 More about models

In 3.2.2 we saw that if there is a model $\mathfrak{M}$ such that

$$\mathfrak{M} \models \phi \Leftrightarrow \phi \text{ is true}$$

then either classical reasoning (in our language) is unsafe, or we get instances of the puzzling scheme $\phi \,\&\, \mathcal{I}\phi$. In this subsection, we will see that the absence of such a model is a serious matter for any proposal to represent with models the language containing the ten thousand sentences.

We will need the following assumptions about the models and the language:

Assumption 1 For any set of sentences Γ, and any sentence ϕ

$$\Gamma \models \phi \Leftrightarrow \phi \text{ is a consequence of } \Gamma.$$

Assumption 2 Consequence preserves truth.

$$\forall\Gamma\forall\phi(\ \phi \text{ is a consequence of } \Gamma \Rightarrow (\ (\forall x \in \Gamma\ x \text{ is true }) \Rightarrow \phi \text{ is true }))$$

Assumption 3 The language containing the ten thousand sentences can express the lack of truth of any of the ten thousand, and the models respect this, in the sense that if ϕ is one of the ten thousand, there is a sentence $n(\phi)$ such that both:

$$\phi \text{ is not true} \Rightarrow n(\phi) \text{ is true}$$

and

$$\forall\mathfrak{M}(\ \mathfrak{M} \models \phi \Rightarrow \mathfrak{M} \not\models n(\phi).$$

Given these three assumptions, and the fact that not every one of the ten thousand is true, there must be a model $\mathfrak{M}$ such that for every ϕ of the ten thousand, $\mathfrak{M} \models \phi$ just in case ϕ is true.

Proof. Suppose not. Then

$$\forall\mathfrak{M}\exists\phi(\ (\phi \text{ is true } \ \&\ \mathfrak{M} \not\models \phi)\ \vee\ (\phi \text{ is not true} \ \&\ \mathfrak{M} \models \phi)\ )$$

Now index the models; we will refer to them with $\mathfrak{M}_i$. For each $\mathfrak{M}_i$, there is at least one sentence ϕ that witnesses the above. Choose one, for each $\mathfrak{M}$, so that we have a ϕ_i for each model. Now let Γ be the set of sentences ψ_i defined as follows:

$$\psi_i = \begin{cases} \phi_i & \text{if } \phi_i \text{ witnesses the left disjunct,} \\ n(\phi_i) & \text{otherwise.} \end{cases}$$

Now since each ψ_i is *not* true at $\mathfrak{M}_i$, we have $\neg\exists\mathfrak{M}(\mathfrak{M} \models \Gamma)$. But then $\forall\phi\forall\mathfrak{M}(\mathfrak{M} \models \Gamma \Rightarrow \mathfrak{M} \models \phi)$. Then by Assumption 1, every ϕ is a consequence of Γ. But each $\psi_i \in \Gamma$ is true. So by Assumption 2, $\forall\phi\ \phi$ is true. But this contradicts the fact that there is a ϕ from among the ten thousand that is not true. □

Let us consider the assumptions. The first one would seem to be a *desideratum* of any model-theoretic representation of the language. It says that the model-theoretic "consequence" relation captures the real consequence relation among sentences of the language. To give up on this is to abandon a central goal of the project of giving a good representation of the language with models.

The second assumption seems uncontroversial.

The third assumption says that for each of our ten thousand sentences ϕ, there is another sentence in the language we are representing which "expresses" the lack of truth of ϕ, in a rather weak sense of "express". That there are such sentences in a language like English is quite obvious: if $\hat{n}$ is a name of the n^{th} sentence of the ten thousand, then the sentence that results from concatenating

$\hat{n}$ with " is not true" is an English sentence that meets the requirement. Of course there could be smaller languages which do not meet assumption three, but such languages are artificial.

3.3.2 Morals about models

We have seen that if we are to discuss the truth and non-truth of the ten thousand sentences, we are bound to either reject some classically endorsed reasoning in our own language, or to accept that there be true instances of the puzzling scheme "ϕ and it is not determinate that ϕ" in our own language.

If we hope to use models in our study of the logic of the vague language, we will want there to be a way of defining the models so that $\Gamma \models \phi$ just in case ϕ really is a consequence of Γ. But the arguments of (3.2.2) and (3.3.1) show that if our object language is as rich as a typical vague language like English, then we will be committed to the existence of a model $\mathfrak{M}$ such that $\mathfrak{M} \models \phi$ just in case ϕ is true, and this leads quickly to the unwanted disjunction of rejecting classical logic or accepting that there be true instances of puzzling schemes. The path of acceptance leads naturally to the Pure Qualification Approach to the logic of vagueness, outlined in section (3.4.1).

If we choose the path of rejecting classical logic within our own reasoning about the vague language, we must go into new and unfamiliar territory. In particular, it looks as though we will not be able to use any familiar mathematical constructions for our models and our notion of "truth" at a model. For such constructions generate a classical mathematical function t from pairs of models and sentences to $\{0, 1\}$ such that $t(\mathfrak{M}, \phi) = 1$ just in case $\mathfrak{M} \models \phi$, and $t(\mathfrak{M}, \phi) = 0$ otherwise. Armed with this function for a given model-theoretic apparatus, it is easy to get the arguments of this section going. In particular, if there is an $\mathfrak{M}$ such that $\mathfrak{M} \models \phi$ just in case ϕ is true, then it is extremely hard to avoid the existence of true instances of the puzzling schemes. But model-theoretic semantics using non-classical objects instead of classical mathematical objects is largely unexplored territory, and it is hard to see how to navigate it.

3.4 General approaches to the logic of vague language

In this section I introduce an approach to the logic of vague language which embraces classical logic and bivalence. This approach is closely associated with the Boundary Semantics of Chapter 4, and further discussion of the approach appears in that chapter. For now, I introduce it and contrast it with two rival approaches.

3.4.1 The Pure Qualification Approach

The Pure Qualification Approach (PQA) gets a philosophical propping in the idea that the *logical* effect of vagueness on a language is only upon those sentences that contain words which are explicitly about vagueness and its sister phenomena like determinacy, while the effect of vagueness (indeterminacy) on the truth values of sentences of the language is to make certain sentences fail to be *determinately* true or false rather than to make them neither true nor false. If a vague language had no words to express the notions of vagueness, indeterminacy, etc., then it would be logically indistinguishable from a precise language. (It is part of the PQA that the notions of truth and falsity are not supposed to be among the sister phenomena of vagueness. This is not to say that it is never indeterminate whether a sentence is true—that is expected—but that the central notions of vagueness, such as determinacy, are not definable in terms of truth and falsity alone.) Thus, since every sentence in a precise language is either true or false, so is every sentence in a vague language. The only contribution that vagueness makes to the logic of a language concerns expressions like 'it is indeterminate whether', and 'it is determinately the case that'.[3] All of the logical principles which govern precise

[3]And also of course "there is no sharp boundary for the F's." We will keep the focus on the determinacy operator and its cousin operators, as discussed in section 2.2.12.

language govern vague language as well. Thus all of the following hold for a vague language:

(Bivalence) Every sentence is either true or false.

(Classical Logic) Every theorem of classical logic is indeed true, and classical consequences are indeed consequences. Classical reasoning is safe.

(Basic Disquotation) The disquotational schemes

ϕ iff 'ϕ' is true

and

not-ϕ iff 'ϕ' is false

hold.

(Classical Compositional Disquotation) Such schemes as

'not-ϕ' is true iff 'ϕ' is not true

and

'ϕ or ψ' is true iff 'ϕ' is true or 'ψ' is true

hold.

The Pure Qualification Approach gets its name from the fact that according to it, one expresses vagueness, and the effect of vagueness on a language, only by using qualifications expressed with operators like 'it is vague whether' or 'it is determinate that' and related words. For example, if it is indeterminate whether ϕ, we can say things like:

- it is indeterminate whether ϕ;

- either ϕ or not ϕ, but it is indeterminate which;
- though the sentence ϕ is either true or false, it is neither determinately true nor determinately false.

Note the last example. On the qualification approach, it is not that vagueness has *no effect whatever* on the part of the language which is not explicitly about vagueness; rather, it makes for vagueness in the matter of which truth values some sentences have. Whatever it is that makes it indeterminate whether a man is bald transfers to the sentence that says the man is bald, and hence it becomes indeterminate whether the sentence is true or not. (And accordingly indeterminate whether it is false or not.)

Theorists who have taken the Pure Qualification Approach include the epistemicists and Vann McGee.[4] (See McGee [21] and [20].)

In the next chapter we will consider a formal semantics for vague languages which is a natural partner for the Pure Qualification Approach. As we will see, it is a natural partner for two reasons: first, it assigns truth values to formal sentences which represent the sentences of a vague language in a way that one would expect if the PQA is correct, generating a plausible candidate

[4]It should be added that toward the end of section 5 of [8], Fine entertains something like the Pure Qualification Approach and comes close to endorsing it, in the discussion of his "true$_T$", which "waxes and wanes" with the sentence it modifies.

for the logic of the vague language; second, at the end of the day one can suggest, as part of the PQA, that there is a Boundary Semantical model b of a vague language $\mathcal{L}$ such that a sentence of $\mathcal{L}$ is true (*simpliciter*) just in case it is formally true in b. This is so despite the problems about the relationship between models and genuine truth that were discussed in section 3.2.

3.4.2 The "Impure" Qualification Approach

This approach is the one traditionally associated with the application of supervaluational semantics to vagueness, and has been suggested by David Lewis and developed by Kit Fine and Hans Kamp, among others.[5] The main idea of the Impure Qualification Approach (IQA) is that truth for a sentence that contains vague terms is truth in every reasonable precisification of the original vague language which contains that sentence. Thus if a sentence says that Arnie is bald and "is bald" can be precisified in such a way to include Arnie, and also can be precisified in such a way as to exclude Arnie, then that sentence is neither true nor false. But the sentence "Arnie is bald or it is not the case that Arnie is bald" will be true, since no matter how you precisify "is bald", Arnie will either be included or excluded, and hence the disjunction will be true in the precisified language.

[5] See Lewis [14], [15], and [18], Fine [8], and Kamp [12].

My calling this approach the "Impure Qualification Approach" is perhaps a result of my own prejudices. The reason I do is that when we expand the object language to include words which describe the language's own vagueness, like the operator "it is determinate that", we end up saying things very much like what we would say on the Pure Qualification Approach. But the current approach is "impure" because of its non-classical truth-value gap.

The connectives, on this approach, are not truth-functional. (On the Pure Qualification Approach, they are truth-functional.) $\phi \vee \neg\phi$ is always true, even when both disjuncts are neither true nor false, but $\phi \vee \psi$ may not be true when both disjuncts are neither true nor false.[6]

Since precisifications of the vague language are classical languages, classical reasoning is (if indirectly) valid. But a vague language will differ from a precise language insofar as Bivalence will fail. Disquotation will probably fail too, for if it is indeterminate whether Arnie is bald, then

Arnie is bald if and only if 'Arnie is bald' is true.

will not straightforwardly hold. For the right hand side will be false, while the left hand side will be neither true nor false. (If we supervaluate such claims

[6]We could perhaps make the connectives value-functional if we re-made the system with the semantic values of sentences as sets of admissible precisifications and if we adjusted the connectives accordingly. A sentence would be called "true" if its value were the set of all admissible precisifications, "false" if its value were the null set. But it is not clear what would motivate such a version of this approach.

as that, then the approach begins to collapse into the Pure Qualification Approach.) And we can begin to cast some doubt on whether classical reasoning is *generally* upheld on the Impure Qualification Approach. For it appears, on this approach, that when the language we are addressing contains a truth predicate that can apply to its own sentences, given that

"Arnie is bald"

is true, it must be the case that

" 'Arnie is bald' is true"

is true. Yet it is *not* the case that given that

"It is not the case that 'Arnie is bald' is true"

is true,

"It is not the case that Arnie is bald"

must be true. In effect we have a ψ that is in a clear sense a consequence of a ϕ, yet $\neg\phi$ is not a consequence of $\neg\psi$, and hence a form of contraposition fails.[7]

[7] And as Fine notes, something like the Deduction Theorem fails too.

3.4.3 The Functional Gap Approach

Proponents of this approach to the logic of vagueness include those who advocate "fuzzy logic", with its multitude of non-standard truth-values, but the essence of the approach is clearly visible with a three-valued truth-functional logic.[8] The basic idea is that vague sentences have a truth-value somewhere "between" classical truth and falsity, and that the truth-values of compound sentences are determined by the truth-values of their constituents, as they are in a classical language.[9] Thus we get multi-valued truth-functional connectives, and multi-valued semantical quantifier rules.

To illustrate with the three-valued propositional logic, supposing that ϕ is vague, it is neither true nor false, but instead has the third value. But so then does its negation $\neg\phi$ and so does the disjunction $\phi \vee \neg\phi$. (In general, a disjunction is true if either disjunct is true, false if both are false, and neither true nor false otherwise.)[10] The truth-functionality of the connectives—rejected on

[8]The *locus classicus* of the fuzzy approach is Zadeh [32], though similar ideas can be found in earlier authors. The fuzzy approach has been developed extensively. Williamson [31] suggests that the first serious application of three-valued logic to vagueness can be found in Halldén [10].

[9]We may count as Functional Gap theorists some who would not want to think of this as a third *truth-value*, but would allow that it formally behaves just like one—the values of compounds are determined from the values of their constituents when we pretend that the lack of truth or falsity is a value in its own right.

[10]Though there are many possible truth-tables for non-bivalent connectives, the most common treatments of disjunction and negation give these results.

the Impure Qualification Approach—is retained, but classical reasoning within the vague language is not. For classically, $\phi \vee \neg\phi$ is a consequence of anything, but it is not a consequence of just anything on the Functional Gap Approach, though it is a consequence of ϕ.

3.5 In favor of the Pure Qualification Approach

This section gives some of the considerations that I think favor the PQA over alternative approaches to the logic of vague language.

3.5.1 Puzzling claims confined

Since classical logic is embraced in principle, so that it is held to be appropriate for the object language, the meta-language, and so forth, none of the puzzlement that arises from rejecting classical logic can arise on the PQA. The only puzzling facts which are tolerated are those that we discern when we consider such facts as conjunctions of the form "It is indeterminate whether ϕ and yet either ϕ or not-ϕ". Disjunctive reasoning tells us that either ϕ and it is not determinate that ϕ or not-ϕ and it is not determinate that not-ϕ. Thus there is a true instance of the puzzling scheme "ϕ and it is not determinate that ϕ". Though there is some comfort in the thought that no instance of this scheme could be *determinately* true, there is no denying its air of mystery.

On the other hand, really to abandon classical logic in a language that contains vague terms means rejecting classical logic in our own language, since it contains vague terms. And if the truth predicate of a vague language is itself vague, as was suggested in section 3.2, then a meta-language for our language will itself contain vague terms. Further, it appears that insofar as it makes sense to talk of a hierarchy of ever-higher "meta-languages", the truth predicate of every one built off of a typical vague language—the meta-language, the meta-meta-language, the meta-meta-meta-language, and so forth—will itself be vague. Thus if classical logic fails in vague languages, we, as a matter of practical fact, must give up classical logic. The failure of classical logic seems to me, however, to be even more puzzling than the truth of an instance of "ϕ and it is not determinate that ϕ". A revision of logic itself must be very forcefully motivated, and the avoidance of the other puzzling fact does not strike me as sufficient motivation. On the PQA, the puzzles do not concern logic itself, but only the quality of determinacy; this is part of why it is appropriately called the "Pure Qualification Approach".

This weighing of puzzlement is not meant to give an absolutely decisive argument for the PQA, for it may be that others simply do not weigh the degrees of puzzlement as I do. There is more to say, however.

3.5.2 Non-bivalence does not itself capture the lack of sharp boundary

In sections 2.1.2 and 3.2 we saw that the sense of a lack of sharp boundary for something like "the red tiles" is not exhausted by the categorization of tiles into three, as opposed to two, classes, as in the threesomes "truly red", "truly not red", and "neither truly red nor truly not red", and "determinately red", "determinately not red", and "neither determinately red nor determinately not red".

A simple three-valued semantics for a vague language is in no better position than a bivalent semantics when it comes to representing our sense that there is a lack of sharp boundary for a predicate in that language. The simple three-valued semantics assigns every sentence of the form "x is a red tile" either true or not true, and hence, speaking the meta-language, we can divide the tiles into 3 mutually exclusive and jointly exhaustive groups: the truly "red", the truly "not red", and the rest. There is nothing non-classical about the meta-language, and this partition into three groups is, from a logical point of view, no different in kind from the partition into two groups that a bivalent interpretation of "x is a red tile" requires. It is just a matter of three instead of two.

But this partition into three seems no better a representation of the lack of sharp boundary for the red tiles than does the bivalent two-way partition. *If* the bivalent partition is objectionably "sharp" then so is the tri-valent partition. It offends our sense of the lack of sharp boundary *just as much* that there should be a pair of adjacent tiles, the first of which is truly "red", the second of which is not, as that there should be a pair of adjacent tiles the first of which is red, the second of which is not.

In fact, the three-way partition is in a way a *worse* representation of our sense of the lack of sharp boundary. For it offers up *two* puzzling adjacent pairs instead of one. On one side, there is the step from truly "red" to not, and on the other, there is the step from not truly "not red" to truly "not red". This observation suggests that an increase in the complexity of the system of truth-values, from three to some larger number, will only make things worse. Nothing will be gained if we move to systems with a set of "truth-values" like the real interval $[0, 1]$ or a large boolean algebra.

What is emerging here is that with any formal semantics that involves the assigning of "truth-values" to sentences, one can define, using the meta-language, a network of terms that will provide a sort of "taxonomy" of the 10,000 tiles, based on their relationship to "is red". In a three-valued system, we get a three-way partition; in a system with truth-values like $[0, 1]$, we get the terms "satisfies 'is red' to degree n". Two points are relevant: First, there will be a meta-language with a classical logic in which the tiles divide up, producing many adjacent pairs of tiles that have what would seem to be important differences with respect to their "redness". There is thus some sense to the claim that a bivalent system for "is red" yields the *minimum* number of puzzling pairs, and thus does the *least* injustice, as a formal semantic system, to our sense of the lack of sharp boundary.

Second, in any reasonable formal semantics for "is red", there will be a "truth-value" that is meant to represent *truth*, however many other values there are. Call this value "1". Now there will be the existence of a pair of adjacent tiles such that the first satisfies "is red" (to degree 1) and the second does not. This pair represents the step from truly "red" to otherwise, and the existence of such a pair is guaranteed by the simple fact that the formal system has some special "truth-value" meant to represent "truth". The only way to avoid commitment to such a pair is to abandon the use of formal semantics of any familiar sort, and to either try to give a formal semantics *about which* one cannot reason classically, or to simply abandon the idea of using any formal semantics as an elucidation of vague language.

3.5.3 Internal coherence

The pure qualification approach deals with this problem in an entirely different way. Its allowance of the puzzling scheme $\phi \,\&\, {\sim}\mathcal{D}\phi$ as having true instances (though never determinately true instances) is not limited to the object language, but rather is seen as the central characteristic fact about indeterminacy and vagueness. Thus it allows the existence of the puzzling adjacent pair of tiles, the first of which is red, the second of which is not, but qualifies this fact as one that is not determinately witnessed by any adjacent pair: there is such a pair, but no pair is determinately like that. Similarly, on the PQA, we can say that any formal model of the original vague language is not determinately correct, since any such model will yield an adjacent pair like that. This distinctive internal coherence of the PQA is explored in more detail in section 4.5, after a formal semantics appropriate for the PQA has been fully laid out.

3.5.4 The psychological grip of classical logic

The argument above, in 3.5.2, and the arguments in section 3.2, suggest that unless we accept the possibility of true instances of the puzzling scheme $\phi \,\&\, {\sim}\mathcal{D}\phi$ in our own language, even the language with which we describe our formal semantics, we cannot have a plausible formal semantics for vague language unless we give up classical logic in our language. Our formal semantics would have to be an apparatus about which it is not correct to reason classically. Aside from the work of intuitionists in mathematics, we do not have much to go on in thinking about what a non-classical formal semantics would be like. But despite the efforts of some authors to assimilate the logic of vague language to intuitionistic logic, it does not appear very plausible that intuitionism or intuitionistic logic illuminates vague language.

In fact, it looks as though all attempts to capture the alleged non-classicality of vague language embrace classical logic at the explanatory level: they try to explain or show the coherence of the alleged failure of classical logic in the vague object language by using a completely classical meta-language. The standard three-valued formal semantics is an obvious example of this, as is the standard treatment of "fuzzy logic".[11]

The fact that it is so typical to conceive the non-classicality of an object language by embedding it in a classical context is not only testimony to the psychological attraction we have to classical logic, but suggests that the alleged failure of classical logic for vague language has never really been conceived in

---

[11]See any standard treatment of three-valued logic for vagueness, like [27], or a standard treatment of languages with the $[0, 1]$ truth-value set, like [32]. The models discussed in Parsons and Woodruff [23] are classical, which illustrates the theme. But later works of Parsons, and what I have learned from him in conversation, are on the track of rejecting classical logic in our own language.

a way that makes it more than superficial. If vagueness really requires the rejection of classical logic, it seems we have not yet met the requirement.

Chapter 4

Boundary Semantics

The subject of this chapter is the formal apparatus I call "Boundary Semantics". The name derives from Kit Fine's discussion in section 5 of [8]. The apparatus is a formal semantics for languages containing an operator meant to represent "it is determinately the case that". We will consider a treatment of a propositional language with the operator, one that is directly inspired by Fine's formal suggestion. We will examine the formal apparatus in some detail, and will show that its logic is quite rich. Then we will consider how the formal idea might be expanded to deal with quantifiers.

Before turning to the formal work, I will present an intuitive motivation for it. The formal idea can be seen as a formal counterpart to the Pure Qualification Approach to the logic of vagueness; the connection will be discussed after the formal apparatus has been laid out in full.

4.1 An intuitive motivation for the semantics

The leading idea

The leading idea is that something is determinately the case just when it is the case according to every admissible complete consistent description of the facts, where a complete consistent description is admissible just in case it does not conflict with anything which is determinately the case. The "something" can be thought of as a proposition or as a (possible) state of affairs, so that a proposition is *determinately true* if it is in every admissible complete description and a (possible) state of affairs *determinately obtains* if it obtains according to every admissible complete description.

The leading idea is circular. It is *not* intended as anything like an *analysis* of the notion of something's being determinately the case. But it is not free of power, especially when it is applied to complex propositions. Consider a proposition of the form $(P \vee \sim P)$. According to the leading idea, this proposition is determinately true just in case it is true according to every admissible, consistent, complete description of the facts. Every such description will include either $(P \vee \sim P)$ or $\sim(P \vee \sim P)$ and hence should (presumably) include

$P \vee \sim P$, since the alternative would (presumably) lead to inconsistency. Thus this disjunction is determinately true. An auxiliary justification of the inclusion of such a proposition is this: if it is determined that a given proposition could not be determined, then this alone is enough to determine the negation of that proposition. Since $\sim(P \vee \sim P)$ could not be determined—since it appears in no admissible, consistent, complete description—its negation is determinate.

One ought to wonder whether there is more than one admissible, consistent, complete description in the above defined sense. If not, the leading idea will indeed get us nowhere. But intuitively, if it is indeterminate whether a given tile is red, then there ought to be two admissible complete descriptions, one that includes the proposition that the tile is red, and one that includes its negation.

To express the idea another way: a sentence is determinately the case if it is the case for every way of "resolving" the vagueness of the language, where resolving the vagueness of something means arbitrarily choosing truth-values for any basic propositions with indeterminate truth-values that are among its components. But the choosing is not entirely arbitrary: it is meant to be coordinated, so the choice made for one indeterminate proposition makes sense given the choices made for others.

The central informal idea can usefully be compared to a certain way of thinking of "possible worlds" and the use of the formal possible-worlds semantics in the study of necessity and possibility. The idea is to reduce talk of possible worlds to talk about sets of propositions.[1] Let us say that a *world-story* is a maximal consistent set of propositions. Now we simply think of something's being the case in a possible world as something's being the case according to a world-story. Something is the case in every possible world if it is the case according to every world-story. Something is necessarily the case if it is the case according to every world-story. But there is no *reduction* of necessity to something involving possible worlds here. The notion of consistency used to define the very idea of a possible world is the driving notion, and it looks as though taking it as primitive is as good as taking necessity as primitive. (Though there are subtle issues here.) Similarly, our leading idea does not purport to analyze, explain, or reduce the notion of determinacy. On our approach, that notion is taken as primitive.

[1]Robert Adams gives a clear presentation and defense of the idea in section *VI* of [2].

4.1.1 Supervaluationism?

A number of philosophers have entertained related ideas about vague language, ideas associated with the name "supervaluationism". But there are at least four different themes associated with that term, only two of which will be embraced in this work. The themes are these:

1. An idea in the metaphysics of vague propositions and facts: the idea given above, that something is determinately the case if it is the case according to every admissible complete consistent description of the facts. Only a coincidence is asserted, not an explanation.

2. An idea in formal semantics: make a formal structure that includes a set of standard structures, and define a valuation function over the larger structure by "supervaluating" the valuation functions for the standard structures.

3. A substantial philosophical thesis about the nature of vague language: vagueness is semantic indecision, or, more precisely, the result of merely partial semantic decision. The vagueness of a term consists in its not being (completely) decided what the term means or expresses.

4. A related thesis about which sentences are true: a sentence is true just

in case it is true on every "admissible interpretation" of its terms (or of the language).

Note that the third theme is quite distinct in kind and in content from the other two themes. Note also that the fourth theme is not the same as the first theme: the fourth concerns which sentences are *true*, and the first concerns which propositions are *determinately true* or which states of affairs *determinately obtain*. Let us consider the themes is some detail.

The third and fourth themes are the historical starting point for supervaluations. They can be found in Mehlberg [22], Lewis [14], [17], and[18], and Fine [8] among other places. Mehlberg, generally regarded as the earliest advocate of "supervaluation" writes:

> The term "Toronto" is vague because there are several methods of tracing the geographical limits of the city designated by this name, all of them compatible with the way the name is used. It may be interpreted, for instance, either as including some particular tree on the outskirts of the city or as not including it. The two areas differing from each other with respect to the spot where this tree is growing are two distinct individual objects; the word "Toronto" may be interpreted as denoting either of these two objects and is for that reason vague.[2]

(We may pass over the apparent slip regarding just what it is whose limits are traced; Mehlberg does not intend to insinuate that there is an object—the city—whose limits can be traced in different ways, etc., an object presumably

[2]Mehlberg [22], p. 86 of [13].

distinct from any of the items traced out in such a process.) Mehlberg goes on to explain how the semantic indecision—the looseness in the use of the term that results in there being multiple tracings of the geographical limits—figures into the truth-values of sentences including the term.

> A statement including vague terms may nevertheless be either true or false if its truth-value is not affected by the multiplicity of their admissible interpretations. Such a statement is true (or false) under every admissible interpretation of the vague terms it contains. Thus, although both "Toronto" and "Canada" are vague terms, it is nevertheless true that Toronto is in Canada, because this statement remains true under any admissible interpretation of the two geographical terms it contains. Similarly the sentence "Toronto is in Europe" is false, because its falsehood is not altered by the choice of any admissible interpretation. The statement "The number of trees in Toronto is even" becomes true under some of the admissible interpretations of its subject, and false under the remaining interpretations; it is therefore neither true nor false.[3]

These two passages give nice illustrations of themes three and four above. It is easy to see that these ideas would lead philosophers to theme two, the generation of a formal counterpart for theme four. There have been various formal ideas, more or less elaborate, which illustrate theme two, including formal treatments given by van Fraassen in [29], by Lewis and Fine in the works cited above, and by Kamp in [12]. (van Fraassen is not concerned with vagueness, but is a pioneer in the development of the formal idea.) Fine is the only one to deal with "higher-order" vagueness in any detail at all, and it is

[3] Mehlberg [22], pp. 86–87 of [13].

his formal suggestion which we develop in this chapter.

It is important to notice that though, as a matter of history, themes three and four led to the development of theme two, theme two can be coupled with theme one and separated from its historical progenitors. That is just what I suggest.

The idea that a vague language has "precisifications"—that its vague sentences can be "resolved" in different ways without conflicting with determinate facts—should be divorced from the idea that vagueness results from indecision on our *part.* It could even be the "fault" of *the world* that we need to make an arbitrary choice if we want to assign a "truth value" to an indeterminate proposition.

Truth-value gaps and determinacy

Note now that the fourth theme suggests that a vague language will have truth-value gaps, while the first theme does not. For the fourth theme implies that if a sentence S can be given different interpretations, on one of which it is true, and on another of which it is false, then S is neither true nor false, as in Mehlberg's example "The number of trees in Toronto is even". But according to the first theme, if there are different admissible descriptions of the facts that give different truth values to S, all that follows is that it is not determinate that S is true and it is not determinate that S is false. In fact this first theme is entirely compatible with the Pure Qualification Approach to the logic of vagueness, on which vagueness does not make for truth-value gaps.

This has been noticed by another "supervaluationist": Van McGee. Consider this passage from his *'Kilimanjaro'*:

> The models we use are sharp. Within a model, every individual constant is assigned a unique individual and every formula is assigned an extension that is the complement of the anti-extension. Our thoughts and practices do not, however, pick out a unique referent for each singular term, nor do they pick out an exhaustive classification for each open sentence. The explanation is that our thoughts and practices do not pick out a unique model as the actual model. They pick out a class of models. The fundamental hypothesis of supervaluation theory is that the semantics of a vague language can be described by singling out an appropriate class of models, thus:
>
> > *Supervaluation Hypothesis* (first version): There is a class $\mathcal{K}$ of $\mathcal{U}$-models such that a variable assignment σ determinately satisfies a formula ϕ if and only if σ satisfies ϕ in every member of $\mathcal{K}$.
>
> An equivalent formulation of the hypothesis makes use of the following definition:
>
> > *Definition.* A $\mathcal{U}$-model $\mathfrak{A}$ is *acceptable* if and only if, for any variable assignment σ and open sentence ϕ, if σ determinately satisfies ϕ, then σ satisfies ϕ in $\mathfrak{A}$.
> >
> > *Supervaluation Hypothesis* (second version): A variable assignment σ determinately satisfies an open sentence ϕ if and only if σ satisfies ϕ in every acceptable $\mathcal{U}$-model.
>
> If $\mathfrak{A}$ is an acceptable model, then every sentence that is determinately true is true in $\mathfrak{A}$ ...

Notice that the passage does not end with McGee's saying that if a sentence is *true* then it is true in any acceptable model, but instead with the idea that if a sentence is *determinately true* then it is true in every acceptable model. McGee can, to this extent, be classified as an advocate of the Pure Qualification Approach: on his view, vagueness causes sentences to fail to be determinately true or determinately false, and it does not cause them to fail to be true or false. Notice that in his statements of the "Supervaluation Hypothesis", McGee uses an unanalyzed notion of determinacy, and indeed, the hypothesis concerns a condition of something's being determinately the case: an assignment σ, he asserts in the first version, determinately satisfies a formula ϕ just in case it satisfies ϕ in every member of a certain class. This is the content of the Hypothesis. And looking at the definition and the second version of the Supervaluation Hypothesis, one sees that McGee's thesis sounds very much like our leading informal idea, the first of the four themes.

McGee's embracing bivalence and rejecting theme four distances him from traditional supervaluational theses. But at the beginning of the passage, McGee expresses some theses that sound a lot like theme three. What he would say about "Toronto" seems to be very much like what Mehlberg says. McGee represents a sort of halfway-point between more traditional forms of supervaluationism and the approach to vague language advocated in this work.

Reduction vs. non-reduction

As noted above, McGee does not attempt to reduce the notion of determinacy; indeed he uses it to define his notion of an acceptable model, which is then used to express his fundamental hypothesis. But one feels that the drift of traditional supervaluationism—centered around theme three—is to reduce the notion of determinacy to something having to do with our practice or semantic indecision, or at least to explain determinacy in terms of our practice or indecision. In the case of Mehlberg's example, the picture would be roughly this:

- We have not, or our practice has not, decided which thing "Toronto" refers to.
- As a result, there are different "admissible" ways of interpreting "Toronto". In particular, "Toronto includes this tree" can be interpreted in such a way as to be true and also in such a way as to be false.
- As a result, "Toronto includes this tree" is neither true nor false.
- As a result, it is indeterminate whether Toronto includes this tree.

This is a chain of explanations; the basic line can be expanded to the reductive idea that what it is for it to be indeterminate whether Toronto includes this tree is for "Toronto includes this tree" to be neither true nor false, and what it is for that to be the case is for there to be different "admissible" ways of interpreting "Toronto"..., and so on, up the explanatory chain.

The suggested explanation and reduction are threatened if the notion of "admissibility" turns out to be expressible in terms of the notion of determinacy in question. For suppose we went with a parallel of McGee's definition of an acceptable model. We would say that it is admissible that "Toronto" refers to α (which includes the tree) and also that "Toronto" refers to β (which does not). But this would mean that it is not determinate that it is not the case that "Toronto" refers to α (and similarly for β). Now if we aim to understand *all* determinacy and indeterminacy as reducible or explainable in the way exemplified in the explanation above, the fact that it is not determinate that "Toronto" does not refer to α will itself be explained by there being an admissible interpretation of the sentence

(S^1) "Toronto" refers to α.

on which it is true. But the fact that there is such an admissible interpretation will be explained by its not being determinate that it is not the case that sentence S^1 is true. This in turn will be explained by there being an admissible interpretation of

(S^2) Sentence S^1 is true.

on which *it* is true. And so on. But the explanations would never come to an end.

Thus on the traditional versions of supervaluation, we must, if we are to avoid explanatory regress, take the notion of the *admissibility* of an interpretation as basic, and not to be defined by, reduced to, explained in terms of, or even on a par with (definable in terms of) the notion of determinacy. This should help emphasize the extent to which McGee's defining "acceptability" with the notion of determinacy is a break from the tradition, and a step in the direction of my approach, on which determinacy and admissibility are inter-definable primitives like necessity and possibility.

4.1.2 Iterations of "it is determinate that"

We turn now to the question of how to use our informal leading idea in connection with a formal apparatus. Among other things, we want to deal with the fact that it is not determinate which things are determinately the case. We flesh out now an informal idea that will lead directly to the formal idea. This idea is directly inspired by remarks of Kit Fine in section 5 of [8]; I have fleshed out his idea somewhat, and divorced it from its connection to traditional supervaluationism (themes three and four above).

Suppose that one is given the task of choosing a classical model $\mathfrak{m}$ for the formal language (without its $\mathcal{D}$ operator) in a way which meets the following constraints:

1 each propositional letter is to be taken to correspond to an "atomic" sentence of a vague natural language;

2 if the natural language correspondent of a propositional letter ϕ is determinately true (determinately false), then $\mathfrak{m}(\phi) = 1 (= 0)$;

3 the assignment must be globally acceptable in the sense that it respects connections among even indeterminate sentences. For example, if x is redder than y, it should not assign 1 as the value for the formal correlate of "y is red" and assign 0 to the formal correlate of "x is red".

Since there will be "atomic" sentences of the natural language that are neither determinately true nor determinately false, there will be more than one way to complete this task. Suppose we make all the relevant choices, thus completing the first task. The result—an assignment of formal truth-values to each propositional letter—we call a zero-order boundary. After choosing one of these, we proceed to the next task: to provide a set of acceptable ways of completing the last task; that is, to provide a set of classical models which meet the constraints given—a set of admissible zero-order boundaries. Since there were different equally acceptable ways to complete the first task, there will be more than one model in our set. To complete our second task we must choose some set of classical models with the idea that we are producing an acceptable choice to play the role of the set of all acceptable choices according to the criteria of the first task. But since there is probably vagueness about exactly which set should count as the set of all acceptable ways of completing the first task, we have a third task: to provide a set of sets of zero-order boundaries, or a set of first-order boundaries. The process has no natural stopping point, and we can imagine iterating it indefinitely.

Given an omega-sequence of boundaries, the first of which is a zero-order boundary, the second a first-order boundary, that includes the first, the third a second-order boundary that includes the second, and so on, a natural way of doing the semantics emerges. If P represents "x is red" and it is indeterminate whether x is red, and if $\mathcal{I}$ represents "it is indeterminate whether..." then we want $\mathcal{I}P$ to be true. What is it in the omega-order boundary which represents the indeterminacy of the natural sentence "x is red"? It is the fact that in the first-order boundary—the chosen *set* of classical models conforming to the constraints of the first task—there is a model which assigns P 1 and also one which assigns P 0. And it goes similarly for the sentence $\mathcal{I}\mathcal{I}P$. If it is indeterminate whether it is indeterminate whether x is red, then there will be a first-order boundary in the second-order boundary which includes two contrary valuations of P and one which does not. $\mathcal{I}P$ would be true at one choice for the first-order boundary, but not for another. And so it goes for arbitrary iterations of $\mathcal{I}$.

The ω-order boundary semantics is to answer to the idea that something is *determinately* the case just when it is the case for every way of resolving its vagueness; for sentences that do not explicitly involve a determinacy operator, this means looking at the first-order boundary—the set of admissible ways of assigning truth-values to sentences which do not mention vagueness. When a sentence includes a determinacy operator, the situation grows more complicated; we may have to look at ways to resolve the vagueness of admissibility (or determinacy) itself in order to resolve the sentence's vagueness. Hence we may need to look at sets of admissible ways of choosing a set of admissible ways of resolving sentences which do not mention vagueness, and so on. This gives us a hint of how the formal semantics of the $\mathcal{D}$ operator will go; immediately below we will turn to a full formulation of a valuation rule for $\mathcal{D}$.

4.2 Propositional ω-order boundaries

We turn now to the formal apparatus. Let us consider first a propositional language $\mathcal{L}$ consisting of propositional letters $P_1, P_2, \ldots$, parentheses, and the connectives $\sim$, $\&$, and $\mathcal{D}$ with the usual grammar. (We define $\mathcal{A}$ and $\mathcal{I}$ for admissibility and indeterminacy as in 2.2.1, and we define $\rightarrow$ and $\vee$ in the usual ways. We may sometimes refer to the propositional letters as P, Q, R, and so forth.)

Our formal apparatus will be constructed out of classical models for the propositional letters of $\mathcal{L}$. A **classical propositional model** is simply a function from the propositional letters to the values {0,1}.

A **propositional ω-order boundary** is now defined as an ω-sequence of objects $b_0, b_1, b_2, \ldots$ such that:

1. b_0 is a classical propositional model, b_1 is set of classical propositional models, b_2 is a set of sets of classical propositional models, and so on. In general, b_n is a set of the sort of object that b_{n-1} is;
2. $b_0 \in b_1$, $b_1 \in b_2$, and so on. In general $b_n \in b_{n+1}$.

ω-order boundaries are the main elements of the boundary semantics. We will define a valuation function so that there is a notion of truth at an ω-order boundary for *all* the sentences of $\mathcal{L}$. We will need first the notion of an *accessibility* relation R on the ω-order boundaries. It is defined as follows: for ω-order boundaries $b = b_0, b_1, \ldots$ and $c = c_0, c_1, \ldots$

$$bRc \Leftrightarrow \forall n \geq 0, c_n \in b_{n+1}$$

The valuation function val_b for an ω-order boundary $b = b_0, b_1, \ldots$ is defined as follows:

1. If ϕ is a propositional letter, $val_b(\phi) = b_0(\phi)$.

2. If $\phi = {\sim}\psi$, $val_b(\phi) = \begin{cases} 1 & \text{if } val_b(\psi) = 0, \\ 0 & \text{otherwise.} \end{cases}$

3. If $\phi = (\psi \,\&\, \chi)$, $val_b(\phi) = \begin{cases} 1 & \text{if both } val_b(\psi) = 1 \text{ and } val_b(\chi) = 1, \\ 0 & \text{otherwise.} \end{cases}$

4. If $\phi = \mathcal{D}\psi$, $val_b(\phi) = \begin{cases} 1 & \text{if } \forall c\ (\ bRc \Rightarrow val_c(\psi) = 1\ ), \\ 0 & \text{otherwise.} \end{cases}$

A sentence ϕ is "true at" an ω-order boundary b ($b \models \phi$) just in case $val_b(\phi) = 1$; otherwise ϕ is "false at" b.

4.3 Some formal features of boundary semantics

4.3.1 Elementary facts

The $\mathcal{D}$ operator of the boundary semantics is like the "necessity" operator in modal logic: $\mathcal{D}\phi$ is true at a boundary just in case ϕ is true at all accessible boundaries. It is easy to see that the (inter-definable) $\mathcal{A}$ operator is semantically similar to the "possibility" operator of modal logic: $\mathcal{A}\phi$ is true at a boundary just in case ϕ is true at some accessible boundary. And the $\mathcal{I}$ operator is similar to the modal "contingency" operator. $\mathcal{I}\phi$ is true at a boundary just in case ϕ is true at some accessible boundary, and ϕ is false at some accessible boundary.

Not only are all classically valid sentences true at all boundaries, but if ϕ is classically valid then $\mathcal{D}\phi$ is true at all boundaries as well. These schemes hold:

$$b \models (\phi \vee \psi) \Rightarrow b \models \phi \text{ or } b \models \psi$$

$$b \models (\phi \rightarrow \psi) \Rightarrow \text{ if } b \models \phi \text{ then } b \models \psi$$

as do their counterparts for $\sim$ and $\&$. Further, note that for arbitrary ϕ and ψ,

$$\phi \models \psi \quad \Leftrightarrow \quad \forall b, b \models (\phi \rightarrow \psi)$$

(for any boundary b).

The connection to standard possible-worlds semantics

Boundaries can be mapped to possible-worlds models in a natural way. Let a "possible-worlds model" $\mathfrak{m}$ consist of a set of points ("worlds"), with one designated point w_0 (the "actual" world) an "accessibility" relation $R_{\mathfrak{m}}$ on the worlds, and a map associating each world with exactly one classical model. (We have in mind here classical models for the language $\mathcal{L}$ without the $\mathcal{D}$ operator, and hence possible-worlds models for $\mathcal{L}$.) Truth at a world w in a possible-worlds model $\mathfrak{m}$ is as usual: If ϕ is atomic, $w \models \phi$ just in case the classical model associated with w is a model of ϕ; the usual sorts of clauses hold for the truth-functional operators; and $w \models \mathcal{D}\phi$ just in case $\forall v(wR_{\mathfrak{m}}v \Rightarrow v \models \phi)$. Finally, let it be that $\mathfrak{m} \models \phi$ just in case $w_0 \models \phi$.

We now give a map $g()$ from ω-order boundaries to possible-worlds models such that for each boundary b, $b \models \phi \Leftrightarrow g(b) \models \phi$. Let the set of worlds of $g(b)$ be the set of all ω-order boundaries that are either accessible from b, accessible from something accessible from b, etc. (I.e., the set of all "ancestors" of b under the accessibility relation on boundaries.) Let $R_{g(b)}$ be the restriction of the accessibility relation on boundaries to this set of boundaries. And let us associate with each possible world the classical model which is the 0-level element of the boundary which is that possible world. Finally, let the w_0 of $g(b)$ be b. Then it is easy to see that for all boundaries b, $b \models \phi \Leftrightarrow g(b) \models \phi$. This is best seen by an induction argument on the structure of ϕ. The base case is trivial, and the induction steps are straightforward.[4]

---

[4]It is worth mentioning that it is not at all obvious how to do something interesting that goes the other way around, i.e., how to define, just in terms of the intrinsic features of possible-worlds models, a class C of possible-worlds models and a map $g()$ from C to the set of all boundaries such that for each possible-worlds model $\mathfrak{m} \in C$, for all sentences ϕ, $\mathfrak{m} \models \phi \Leftrightarrow g(\mathfrak{m}) \models \phi$, especially when we add the restriction that every boundary should be in the range of $g()$. The rich structure of the accessibility relation built in to the boundaries, which emerges immediately below, makes this task difficult.

What is the logic?

Consider now the question what the logic of boundary semantics is. First, note that it conforms to the Minimal Assumptions of (2.2.1). Second, note that every sentence with the form of a classical tautology is true at every ω-order boundary. Third, as Kit Fine observes, the accessibility relation R defined above is reflexive, and hence, in one way, the correct logic for boundary semantics would seem to be the modal logic **T**. (Thinking of boundary-semantical $\mathcal{D}$ as modal-semantical L or $\Box$.) But in fact there are sentences of the language which will be true at every ω-order boundary ("boundary" for short) and which are not part of **T**. As we will see below, given that ϕ contains no occurrences of $\mathcal{D}$, the formula

$$\mathcal{A}\phi \rightarrow (\psi \rightarrow \mathcal{A}(\phi \,\&\, \mathcal{A}\psi))$$

is true at every boundary. But note that if ϕ contains $\mathcal{D}$, then the above formula may not be true at a boundary. Hence if the logic for boundary semantics is the set of sentences such that every one of them is true at every boundary, then the logic is richer than **T**, but if the logic is the set of sentences such that every substitution instance of each is true at every boundary, then the logic is **T**.[5] Which is the correct notion of the "logic" seems to me probably

[5]Fine asserts that the logic of his structures is **T**.

a merely verbal dispute, and it is of some interest what the correct "logic" in the first sense is. The answer to this question is, as far as I know, open. But we can safely say that every theorem of **T** is a validity of the boundary semantics, whichever way we think of validity.

4.3.2 Divergences from familiar modal logics

There is a fairly simple fact about the boundaries which can be used to show that there are various ways in which the "logic" of boundary semantics deviates from that of familiar modal logics like **T**. That fact is this:

Fact 3. $\forall b \forall c (bRc \Rightarrow \exists d (d_0 = c_0 \text{ and } bRd \text{ and } dRb))$

Proof. Suppose that bRc. Let $d_0 = c_0, d_n = b_n (n \geq 1)$. Then bRd since $d_0 (= c_0) \in b_1$, and bRc. And since b is a boundary, it is clear that $d_n (= b_n) \in b_{n+1} (n \geq 1)$. And dRb since clearly $b_n \in d_{n+1} (n \geq 0)$. □

Thus the "accessibility" relation on the boundaries is constrained in such a way as to force a richer network of connections than those that are forced by any obvious stipulations about the accessibility relation in standard possible-worlds semantics.

Now it is obvious that if two boundaries have the same 0-level member, then they agree on all sentences that do not involve $\mathcal{D}$, since the truth conditions for such sentences at a boundary do not concern anything other than the 0-level member of the boundary. Thus, considering the above fact, it is fairly easy to see that if ϕ does not contain $\mathcal{D}$

$$\mathcal{A}\phi \to (\psi \to \mathcal{A}(\phi \,\&\, \mathcal{A}\psi))$$

will be true at any boundary. For

$$\begin{aligned}
b \models (\mathcal{A}\phi \,\&\, \psi) \quad &\Rightarrow \quad \exists c(bRc \text{ and } c \models \phi) \\
&\Rightarrow \quad \exists d(d_0 = c_0 \text{ and } bRd \text{ and } dRb) \quad \text{(by Fact 3)} \\
&\Rightarrow \quad d \models \phi \text{ and } dRb \\
&\Rightarrow \quad d \models \phi \,\&\, \mathcal{A}\psi \qquad \text{(since } b \models \psi) \\
&\Rightarrow \quad b \models \mathcal{A}(\phi \,\&\, \mathcal{A}\psi) \qquad \text{(since } bRd)
\end{aligned}$$

Similarly, note that if $b_n = c_n (n \geq 1)$, then bRd if and only if cRd, and hence $b \models \mathcal{D}\phi$ if and only if $c \models \mathcal{D}\phi$. Now suppose that ϕ does not contain $\mathcal{D}$.

Then

$$
\begin{aligned}
b \models \mathcal{A}\phi \quad & \Rightarrow \quad \exists c \quad c \models \phi \text{ and } bRc \\
& \Rightarrow \quad d(= c_0, b_1, b_2, \ldots) \models \phi \\
& \Rightarrow \quad (b \models \mathcal{D}\psi \;\Rightarrow\; d \models \mathcal{D}\psi) \\
& \Rightarrow \quad (b \models \mathcal{D}\psi \;\Rightarrow b \models \mathcal{A}(\phi \,\&\, \mathcal{D}\psi)\,) \quad (\text{since } bRd) \\
& \Rightarrow \quad b \models \mathcal{D}\psi \rightarrow \mathcal{A}(\phi \,\&\, \mathcal{D}\psi))
\end{aligned}
$$

Thus

$$\models \quad (\mathcal{A}\phi \,\&\, \mathcal{D}\psi) \rightarrow \mathcal{A}(\phi \,\&\, \mathcal{D}\psi)$$

yet not every such sentence is a theorem of **T**.

Fact 3 generalizes, showing that the set of sentences true at all boundaries will not be easy to describe. To state the generalization, we will need the following definition. We will write (for $n \geq 2$)

$$bR^n c$$

to mean

$$\exists x^1 \exists x^2 \ldots \exists x^{n-1}(bRx^1 \text{ and } x^1Rx^2 \text{ and } \ldots \text{ and } x^{n-1}Rc)$$

Fact 4. If boundaries b and c differ in at most one member then bRc and cRb.

Proof. Suppose b and c are boundaries and $b_n \neq c_n$, but for $i \neq n$, $b_i = c_i$. Then to show that bRc we need to show that for all j, $c_j \in b_{j+1}$. This is obvious for $j < (n-1)$ and for $j > n$. Now $c_{n-1} = b_{n-1}$, and b is a boundary, so $c_{n-1} \in b_n$. And $c_n \in c_{n+1}$ since c is a boundary, but $c_{n+1} = b_{n+1}$, and so $c_n \in b_{n+1}$. Thus bRc. Swap b and c and the same reasoning gives us that cRb. □

Now we are set to observe some structural facts about the interconnections among boundaries. Let us say that boundary b *neighbors* boundary c if b and c differ in at most one member. We saw above that if bRc then there is a boundary whose 0-level member is c_0 but is otherwise identical with b and that this boundary is accessible from b. This boundary was a neighbor of b. We can show that if $bR^n c$, then there is a series of boundaries, each of which neighbors the last, the first of which is b and the last of which shares its 0-level member with c.

Let $f()$ be the function defined recursively as

$$f(1) = 1; f(n+1) = f(n) + n + 1.$$

Fact 5. If $bR^n c$ then there is a boundary d such that $bR^{f(n)}d$ and $dR^{f(n)}b$ and $d_0 = c_0$.

Proof. Suppose $bR^n c$. Then there are $x^1, \ldots, x^{n-1}$ such that bRx^1 and x^1Rx^2

$\ldots$ and $x^{n-1}Rc$. Now we construct $y^0 \ldots y^{f(n)}$ as follows:

$$
\begin{aligned}
y^0 &= b \\
y^1_0 &= x^1_0 && \text{and otherwise } y^1 = y^0 \\
y^2_1 &= x^1_1 && \text{and otherwise } y^2 = y^1 \\
y^3_0 &= x^2_0 && \text{and otherwise } y^3 = y^2 \\
y^4_2 &= x^1_2 && \text{and otherwise } y^4 = y^3 \\
y^5_1 &= x^2_1 && \text{and otherwise } y^5 = y^4 \\
y^6_0 &= x^3_0 && \text{and otherwise } y^6 = y^5 \\
y^7_3 &= x^1_3 && \text{and otherwise } y^7 = y^6 \\
y^8_2 &= x^2_2 && \text{and otherwise } y^8 = y^7 \\
y^9_1 &= x^3_1 && \text{and otherwise } y^9 = y^8 \\
y^{10}_0 &= x^4_0 && \text{and otherwise } y^{10} = y^9 \\
&\vdots && \vdots \\
y^{f(n-1)}_1 &= x^{n-1}_1 && \text{and otherwise } y^{f(n-1)} = y^{f(n-2)} \\
y^{f(n)}_0 &= c_0 && \text{and otherwise } y^{f(n)} = y^{f(n-1)}
\end{aligned}
$$

I trust that the pattern will become evident upon examination. It is obvious that y^i neighbors y^{i+1}. It may not be immediately evident that each y^i is a boundary. But this too will become evident when one reflects that not only is each x^i a boundary (so $x^i_j \in x^i_{j+1}$), but also $x^i R x^{i+1}$ (and so $x^i_j \in x^{i-1}_{j+1}$). Let the desired boundary d be $y^{f(n)}$ and we are done. □

Now note that

$$(bR^n c \text{ and } c \models \phi) \Leftrightarrow b \models \mathcal{A}^n \phi$$

(This is pretty obvious if $n = 1$, and the induction to the general case is easy.)

Combining this with Fact 5, we get

Fact 6. If ϕ does not contain $\mathcal{D}$ and if $b \models \psi \,\&\, \mathcal{A}^n\phi$, then $b \models \mathcal{A}^{f(n)}(\phi \,\&\, \mathcal{A}^{f(n)}\psi)$.

Thus if ϕ does not contain $\mathcal{D}$,

$$(\psi \,\&\, \mathcal{A}^n\phi) \to \mathcal{A}^{f(n)}(\phi \,\&\, \mathcal{A}^{f(n)}\psi)$$

is true at every boundary.

4.4 Boundary Semantics with Quantifiers

4.4.1 The straightforward addition: a static domain

There is a straightforward way to extend the formal apparatus of boundary semantics to treat existential and universal quantifiers. Let us add to our language $\mathcal{L}$ an existential quantifier $\exists$, n-ary predicate symbols $F_1^n, F_2^n, \ldots$ (for $n \geq 1$), and variables $x_1, x_2, \ldots$ with the usual first-order grammar. (We will talk of symbols $\to$, $\vee$, and $\exists$ with the usual abbreviations in mind. We may refer to a propositional letter as a 0-ary predicate symbol, and we may pretend that the variables are letters like x, y, z, and w.) Next we must adjust the notion of an ω-order boundary. Again an ω-order boundary will be constructed out of "classical models" appropriate for the language with the determinacy operator taken out. But now the classical models must be more complex; each one will include a domain—a set of objects—and many functions from parts of the language to structures made of the objects in the domain. For the current system, each ω-order boundary will be built of classical models all with the same domain.

A **classical quantificational model with domain** D is an ordered pair $(D, predval())$, where D is a set of objects, and *predval* is the union of a sequence of functions $(predval^0(), predval^1(), \ldots)$ such that $predval^0()$ is a function from the propositional letters to $\{0, 1\}$, $predval^1()$ is a function from the unary predicate letters to the set of all subsets of the domain D, $predval^2()$ is a function from the 2-ary relation symbols to the set of all sets of ordered pairs of elements of the domain, and in general $predval^n()$ $(n \geq 1)$ is a function from the n-ary predicate symbols to the set of all sets of n-tuples of elements of the domain. Thus $predval(F_i^n)$ will be $predval^n(F_i^n)$. If b_0 is a classical quantificational model, then we will use "$b_0()$" to refer to its $predval()$ function, so that $b_0(F) = predval(F)$.

A **static-domain ω-order boundary with domain** D is a pair $(D, (b_0, b_1, \ldots))$ where D is a set of objects, and the ω-sequence of objects $(b_0, b_1, b_2, \ldots)$ is such that:

1. b_0 is a classical quantificational model with domain D, b_1 is set of classical quantificational models with domain D, b_2 is a set of sets of classical quantificational models with domain D, and so on. In general, b_n is a set of the sort of object that b_{n-1} is;

2. $b_n \in b_{n+1}$ for $n \geq 0$.

If $b = (D, (b_0, b_1, \ldots))$, we will refer to $(b_0, b_1, \ldots)$ as the *content* of b.

The accessibility relation R is much as before: for ω-order boundaries b and c with content $(b_0, b_1, \ldots)$ and $(c_0, c_1, \ldots)$ respectively,

$$bRc \Leftrightarrow \forall n \geq 0, c_n \in b_{n+1}.$$

Note that if bRc, then the domain of c_0 (and of every member of c_1, and of every member of every member of c_2, and so on) is also D.

We will now give a recursive definition of truth at a boundary model $b = (D, (b_0, b_1, \ldots))$ relative to an *assignment*. An assignment is a function $a()$ from the variables $x_1, x_2, \ldots$ of the language to the domain D. We define $a[n/o]$ as the assignment that assigns the same values to the variables as does a, *except* that it assigns the object o to the variable x_n. We refer to the function *predval*() in the classical model b_0 as $b_0()$. The definition of the truth-value of a sentence ϕ at an ω-order boundary b relative to an assignment a, notated as $val_{b,a}(\phi)$, is as follows:

1. If ϕ is a propositional letter, $val_{b,a}(\phi) = b_0(\phi)$.
2. If ϕ is of the form $F^n(x_1, \ldots, x_n)$, $val_{b,a}(\phi) =$ $\begin{cases} 1 & \text{if } (a(x_1), \ldots, a(x_n)) \in b_0(F^n), \\ 0 & \text{otherwise.} \end{cases}$
3. If $\phi = {\sim}\psi$, $val_{b,a}(\phi) =$ $\begin{cases} 1 & \text{if } val_{b,a}(\psi) = 0, \\ 0 & \text{otherwise.} \end{cases}$

4. If $\phi = (\psi \,\&\, \chi)$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if both } val_{b,a}(\psi) = 1 \text{ and } val_{b,a}(\chi) = 1, \\ 0 & \text{otherwise.} \end{cases}$

5. If $\phi = \forall x_n \psi$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if for every object } o \text{ in } D,\ val_{b,a[n/o]}(\psi) = 1 \\ 0 & \text{otherwise.} \end{cases}$

6. If $\phi = \mathcal{D}\psi$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if for every } c\ (\ bRc \Rightarrow val_{c,a}(\phi) = 1\ ), \\ 0 & \text{otherwise.} \end{cases}$

We will say that a sentence ϕ is "true" at a boundary b ($b \models \phi$) just in case for every assignment a, $val_{b,a}(\phi) = 1$.

4.4.2 The logic of the static-domain models

There are many straightforward observations we can make about the set of all sentences true at every boundary model (the set of "validities" of static-domain boundary models). It is easy to see that any sentence that is true at every propositional boundary model is true at every static-domain boundary model. And it is easy to see that if a sentence does not contain $\mathcal{D}$, then it is true at every static-domain boundary model just in case it is a validity of the classical first-order predicate calculus. Further, letting $\models \phi$ mean that every static-domain boundary model is a model of ϕ, it is easy to see that

$$\models \phi \quad \Rightarrow \quad \models \mathcal{D}\phi$$

$$\models \phi \quad \Rightarrow \quad \models \forall x\phi.$$

Next we can note that if ϕ is a theorem of classical propositional logic, then if ψ arises from ϕ by uniform substitution of sentences of the language of static-domain boundary models for propositional letters in ϕ, then $\models \psi$.

One of the more interesting validities of static-domain boundary models is the (analogue of the) *Barcan Formula.* Schematically,

$$\forall x \mathcal{D} \phi \rightarrow \mathcal{D} \forall x \phi.$$

To show that it is a validity, we will show that

$$\forall x \mathcal{D} \phi \models \mathcal{D} \forall x \phi$$

and infer the desired result. (The fact that if $\phi \models \psi$ then $\models \phi \rightarrow \psi$ is, as for propositional boundary semantics, easy to see.) Now we reason as follows:

$$
\begin{aligned}
b \models \forall x \mathcal{D}\phi \Rightarrow\ & \text{For every object } o \text{ and every assignment } a \\
& val_{b,a[x/o]}(\mathcal{D}\phi) = 1 \\
\Rightarrow\ & \text{For every object } o \text{ and every assignment } a\text{, and every } c \text{ such that } bRc \\
& val_{c,a[x/o]}(\phi) = 1 \\
\Rightarrow\ & \text{For every object } o \text{ and every assignment } a\text{, and every } c \text{ such that } bRc \\
& val_{c,a[x/o]}(\forall x\phi) = 1 \\
\Rightarrow\ & \text{For every object } o \text{ and every assignment } a \\
& val_{b,a[x/o]}(\mathcal{D}\forall x\phi) = 1 \\
\Rightarrow\ & b \models \mathcal{D}\forall x\phi
\end{aligned}
$$

An interesting non-validity of static-domain boundary models is

$$\mathcal{D}\exists x\phi \rightarrow \exists x\mathcal{D}\phi.$$

Here is a counterexample. Let b's domain D be $\{\alpha, \beta\}$. Let b_0 be a classical quantificational model whose *predval*() function $b_0()$ assigns F $\{\alpha\}$, and let c_0 be one with $c_0(F) = \{\beta\}$. Let b_1 be $\{b_0, c_0\}$. Then every c such that bRc makes $\exists xFx$ true, and hence $b \models \mathcal{D}\exists xFx$. But there is no object o in D such that for every c such that bRc, $o \in c_0(F)$, and hence no assignment a such that

$val_{b,a}(\mathcal{D}Fx) = 1$, and thus $b \not\models \exists x\mathcal{D}Fx$. In fact, $b \models {\sim}\exists x\mathcal{D}Fx$. Therefore $\not\models \mathcal{D}\exists x\phi \rightarrow \exists x\mathcal{D}\phi$.

That this is not valid might at first seem surprising. For, one might think, if it is determinate that there is something that is a certain way, shouldn't it be that there is some thing that is determinately that way? If nothing is determinately that way, then what could make it the case that it is determinate that there is something that is that way? But these phenomena are like the characteristic possibility embraced on our approach, that there be a determinately true disjunction neither disjunct of which is determinately true.

4.4.3 Adding constants

Constants can be straightforwardly added to the apparatus of the static-domain models. The association between constant symbols and objects of the domain D is to be as static as the domain itself. Thus it makes sense to think of the association as part of the entire boundary model. So we add to our conception of a static-domain ω-order boundary with domain D a constant assignment function $referent()$ from the constants $\{c_1, c_2, \ldots\}$ of the language to D. We need to modify the second clause of our recursive definition of the truth-value of a sentence at a boundary with domain D relative to an assignment a so that it reads as follows:

If ϕ is of the form $F^n(t_1, \ldots, t_n)$, where $t_1, \ldots, t_n$ are either variables are constants, then let $object(t_i)$ be $a(t_i)$ if t_i is a variable and let $object(t_i)$ be $referent(t_i)$ if t_i is a constant. Now $val_{b,a}(\phi) =$
$$\begin{cases} 1 & \text{if } (object(t_1), \ldots, object(t_n)) \in b_0(F^n), \\ 0 & \text{otherwise.} \end{cases}$$

4.4.4 Identity in static-domain models

Treating identity in static-domain models is straightforward. We add the two-place relational predicate '=' to the language, and add the following clause to our recursive definition of the truth-value of a sentence at a boundary with domain D relative to an assignment a.

If ϕ is of the form $s = t$, where s and t are variables or constants, then let $object(s)$ be either $referent(s)$ (if s is a constant) or $a(s)$ (if s is a variable), and similarly for t. Then $val_{b,a}(\phi) = \begin{cases} 1 & \text{if } object(s) = object(t) \\ 0 & \text{otherwise.} \end{cases}$

Since the domain is static, nothing especially complicated happens. Note that

$$\forall x \forall y (x = y \rightarrow \mathcal{D}(x = y))$$

is true at every boundary. So this simple treatment of identity will not help to model the idea that identity could be indeterminate. Further, note that

$$\models \forall x \forall y (x = y \rightarrow \phi(x) \leftrightarrow \phi(y))$$

(where ϕ is an arbitrary formula and $\phi(y)$ arises from $\phi(x)$ by substitution of one or more occurrences of x by y). Similarly, for any constants c_i and c_j,

$$c_i = c_j \models \phi(c_i) \leftrightarrow \phi(c_j).$$

Thus classically validated principles of substitution will be validated.

4.4.5 Variable-domain boundary models

Identity can be given a quite different treatment, one which is aimed at modeling the idea that identity can fail to be determinate. The intuitive idea behind it goes along with the motivation for the general boundary-semantical approach discussed in 4.1. In the propositional setting, the classical model b_0 at the "bottom" of the ω-order boundary b is meant to represent a complete admissible description of the facts: a way of selecting truth-values for all propositions, vague or not, that does not conflict with any determinate facts. Thus it can be said to represent a "precisification" of the facts. When we introduce quantifiers, constants and the identity relation, we can think of the precisification as precisifying not only the propositions, but, in a sense, the objects themselves. At any rate, when we precisify we must settle any indeterminate identity propositions. Suppose that for some α and β, it is indeterminate whether α is identical with β. Then when we precisify the indeterminate proposition that $\alpha = \beta$, we may settle it either way, as long as we coordinate this bit of precisifying with the way we precisify other propositions. And when we shift to building b_1, which represents (a precisification of) the collection of all legitimate ways of precisifying the facts, we will want to include both precisifications.

The formal technique we will use is to have for each ω-order boundary b a domain D which we will call the "master domain" of b. D is to be thought of as the collection of all objects, whether indeterminate or not. Now at the classical model b_0, there is a (possibly distinct) domain D_p which represents a precisification of the collection of all objects. We need to include a function which associates the two domains, telling us, in effect, for each object in D, which object of D_p represents it. Thus we modify the "classical models" out of which an ω-order boundary is constructed. Each will include its own domain, and a mapping from predicates to constructions out of that domain. Each will also include a mapping from D to the objects of its domain.

Here is the formal apparatus for our first-order language $\mathcal{L}$ with constants and identity.

A c**-model appropriate for master domain** D is a quadruplet $(D, D_p, precisification(), predval())$, where

1. D is a set of objects (the master domain);
2. D_p is a set of objects (the "local domain");
3. $precisification()$ is a function from D to D_p;
4. $predval()$ is the union of a sequence of functions $(predval^0(), predval^1(), \ldots)$ such that $predval^0()$ is a function from the propositional letters of $\mathcal{L}$ to $\{0, 1\}$, $predval^1()$ is a function from the unary predicate letters to the set of all subsets of the domain D_p, $predval^2()$ is a function from the 2-ary relation symbols to the set of all sets of ordered pairs of elements

of the domain, and in general $predval^n()$ $(n \geq 1)$ is a function from the n-ary predicate symbols to the set of all sets of n-tuples of elements of the domain D_p. Again, $predval(F^i)$ is $predval^i(F^i)$.

If b_0 is a c-model appropriate for D, we will refer to its $predval()$ and $precisification()$ functions as simply "$b_0()$", so that for a predicate F, $b_0(F) = predval(F)$ and for an object $\alpha \in D$, $b_0(\alpha) = precisification(\alpha)$.

A **variable-domain ω-order boundary with master domain** D will be a triplet $(D, referent(), (b_0, b_1, \ldots))$, where

1. D is a set of objects;
2. $referent()$ is a function from the constants of $\mathcal{L}$ to D;
3. b_0 is a c-model appropriate for D, b_1 is set of such models, b_2 is a set of sets of such models, and so on. In general, b_n is a set of the sort of object that b_{n-1} is;
4. $b_n \in b_{n+1}$ for $n \geq 0$.

The accessibility relation R is much as before: for ω-order boundaries $b = (D, referent(), (b_0, b_1, \ldots))$ and $c = (D, referent(), (c_0, c_1, \ldots))$

$$bRc \Leftrightarrow \forall n \geq 0, c_n \in b_{n+1};$$

b and c must have the same domain and the same $referent()$ function.

As before, we will need a notion of an assignment to the variables $x_1, x_2, \ldots$. An assignment a will be a function from the variables to D. As before we define $a[n/o]$ as the assignment that assigns the same values to the variables as does o, *except* that it assigns the object o to the variable x_n. Recall that we refer to the $predval()$ function of the c-model b_0 as $b_0()$, and we refer to the $precisification()$ function for c-model b_0 as $b_0()$. The definition of the truth-value of a sentence ϕ at a variable-domain ω-order boundary b, with domain D, relative to an assignment a, notated as $val_{b,a}(\phi)$, is as follows:

1. If ϕ is a propositional letter, $val_{b,a}(\phi) = b_0(\phi)$.
2. If ϕ is of the form $F^n(t_1, \ldots, t_n)$, where $t_1, \ldots, t_n$ are either variables or constants, then let $object(t_i)$ be $a(t_i)$ if t_i is a variable and let $object(t_i)$ be $referent(t_i)$ if t_i is a constant. Now $val_{b,a}(\phi) =$ $\begin{cases} 1 & \text{if } (b_0(object(t_1)), \ldots, b_0(object(t_n))) \in b_0(F^n), \\ 0 & \text{otherwise.} \end{cases}$
3. If ϕ is of the form $t_1 = t_2$, where t_1 and t_2 are either variables are constants, then let $object(t_1)$ be $a(t_1)$ if t_1 is a variable and let $object(t_1)$ be $referent(t_1)$ if t_1 is a constant. (Similarly for t_2.) Now $val_{b,a}(\phi) =$ $\begin{cases} 1 & \text{if } b_0(object(t_1)) = b_0(object(t_2)) \\ 0 & \text{otherwise.} \end{cases}$
4. If $\phi = \sim\psi$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if } val_{b,a}(\psi) = 0, \\ 0 & \text{otherwise.} \end{cases}$
5. If $\phi = (\psi \,\&\, \chi)$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if both } val_{b,a}(\psi) = 1 \text{ and } val_{b,a}(\chi) = 1, \\ 0 & \text{otherwise.} \end{cases}$
6. If $\phi = \forall x_n \psi$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if for every object } o \text{ in } D,\ val_{b,a[n/o]}(\psi) = 1 \\ 0 & \text{otherwise.} \end{cases}$

7. If $\phi = \mathcal{D}\psi$, $val_{b,a}(\phi) = \begin{cases} 1 & \text{if for every } c \text{ (} bRc \Rightarrow val_{c,a}(\phi) = 1 \text{),} \\ 0 & \text{otherwise.} \end{cases}$

A characteristic fact about variable-domain boundaries that distinguishes them from static-domain boundaries with identity is this:

$$\not\models \forall x \forall y (x = y \rightarrow \mathcal{D}(x = y))$$

Thus our models do not validate a sentence which represents the idea that identity is always determinate. To see this, consider a boundary b with a master domain $\{\alpha, \beta\}$. Let b_0 have only the single object α in its local domain D_p, and let there be a c-model c_0 appropriate for D in b_1 whose local domain is $\{\alpha, \beta\}$. Let $b_0(\alpha) = b_0(\beta) = \alpha$ and let $c_0(\alpha) = \alpha$ and $c_0(\beta) = \beta$. Let a be an assignment such that $a(x) = \alpha$ and $a(y) = \beta$. Then $val_{b,a}(x = y) = 1$. Yet there is a c such that bRc with c_0 as its 0-level c-model, and $val_{c,a}(x = y) = 0$. Thus $val_{b,a}(\mathcal{D}(x = y)) = 0$, and so $val_{b,a}(x = y \rightarrow \mathcal{D}(x = y)) = 0$, and finally $val_{b,a}(\forall x \forall y (x = y \rightarrow \mathcal{D}(x = y))) = 0$.

Fully general principles of substitution (in the manner of *Leibniz' Law*) fail. First note that

$$\models c_i = c_i$$

and (hence)

$$\models \mathcal{D}(c_i = c_i)$$

Now, continuing with the example above, if $referent(c_1)$ in b is α and $referent(c_2)$ in b is β, then $b \models \mathcal{D}(c_1 = c_1)$, while $b \models {\sim}\mathcal{D}(c_1 = c_2)$. Yet $b \models (c_1 = c_2)$, and so we see that in general, the schemes

$$c_1 = c_2 \rightarrow (\phi(c_1) \leftrightarrow \phi(c_2))$$

and

$$\forall x \forall y(x = y \rightarrow (\phi(x) \leftrightarrow (\phi(y))))$$

are not valid: they have invalid instances. Nevertheless *If ϕ does not contain $\mathcal{D}$*, the resulting instances of the schemes *are* valid. This is because the truth-value of a sentence that does not contain $\mathcal{D}$ at a boundary b does not depend on anything outside of the c-model b_0, and c-models are effectively classical quantificational models with their local domains. In general, if ϕ does not contain $\mathcal{D}$, then for every n,

$$\models \forall x \forall y(\mathcal{D}^n(x = y) \rightarrow \mathcal{D}^n(\phi(x) \leftrightarrow (\phi(y))))$$

since the truth of $\mathcal{D}^n(\psi)$ relative to an assignment requires that ψ be true relative to that assignment at a certain set of boundaries, and if ψ does not contain $\mathcal{D}$, as neither ϕ nor $x = y$ do, then its truth conditions depend only on the classical models at the bottoms of these boundaries.

4.5 The internal coherence of the Pure Qualification Approach

In section 3.2 we saw that any approach to the logic of a typical vague language with a determinacy operator is faced with the prospect of accepting that there are true sentences of the form "P and it is not determinate that P" in our very own language. And we also saw in 3.2.2 and 3.3.1 that something like this holds if we want to have a model-theoretic semantics for the object language. We now consider how the Pure Qualification Approach to the logic of such a language can give an internally coherent picture of the relationship between a vague language with a determinacy operator and boundary models for it.

Consider again the 10,000 "This tile is red" sentences of 3.2. Let us symbolize them as $P_1, \ldots, P_{10,000}$. Any boundary model for a language including these 10,000 sentence letters will determinately assign each of them one of two values. Thus there will not be an unsharp boundary for the sentences which are assigned 1 (made true at the model).

Considering the 10,000 colored tiles, we want to say that there is no sharp boundary for the red tiles. Thus we may initially want to say that there is no adjacent pair of tiles such that the first is red and the second is not red. But this leads to trouble, for if there is no such pair, then it looks as if the following holds: if a tile n is red, then tile $n + 1$ is red. But then it looks like given that the first tile is red, we'll be able to conclude that the 10,000$^{\text{th}}$ tile is red. This is a classic *sorites* paradox.

We can instead say that there is no adjacent pair of tiles such that the first is determinately red and the second is determinately not red. This does not land us in the same sort of difficulty. But it does mean that if there is a pair of tiles such that the first is red and the second is not, then either there is a tile that is red and not determinately red, or there is a tile that is not red and is not determinately not red. Thus (unless we give up on the classical reasoning used here) we must conclude that there is a true instance of the form "S and it is not determinate that S".

Thus we will have to accept that there are true instances of this puzzling scheme, unless we want to give up on classical reasoning. We will have to regard this puzzling fact as one of the distinctive marks of vagueness. But if we can accept this then we will be able to say some traditional sounding things about the relationship between a vague language and models for it.

Consider first a simple "classical model" m for the 10,000 sentences $s_1, \ldots, s_{10{,}000}$ "This tile is red" written on the backs of the 10,000 tiles. It assigns a value (1 or 0) to each sentence. Now we want to say that there is no sharp line between the sentences that are true and the sentences that are not true. Can we consistently say this and that for each sentence s, s is true if and only if the simple classical model m assigns s to 1? We can if we accept that the lack of sharp boundary here manifests itself as (among other things) the existence of some sentences that are neither determinately true nor determinately not true. So even if model m assigns s_{5000} the value one, and so the sentence is true, it may not be determinately true.[6]

Similarly, suppose we accept that for all $n > 1$, if tile n is red, then tile $n-1$ is red. It is inconsistent with classical logic to hold this and that tile 1 is red and tile 10,000 is not red *and* that there is no last red tile in the series. Thus we may have to accept that there is a last red tile. But we can still register the lack of sharp boundary of the red tiles by adding that there is no *determinately* last red tile. We can say that there is no tile n of which it is determinate that it is red and it is determinate that $n+1$ is not red. Thus though (classical) logic forces us to accept that there is a last red tile, we can still register our sense that there is no unknown "magic tile" which marks a sharp boundary between the red and the non-red by saying that there is no *determinately* last red tile. And we may similarly say that though (classical) logic and semantics force us to accept that there is a classical model m such that each sentence s_n of the 10,000 is true just in case m assigns s_n 1, there is

---

[6]We should expect that m makes true $\sim\mathcal{D}s$.

no unknown “magic” model m such that it is *determinate* that each sentence s_n is true just in case m assigns it 1.

A boundary model for a language including the 10,000 sentences will make each one either “true” or “false”. Now let F represent “is red”, and let R represent the two-place relation that holds between o and p just in case o and p are two adjacent tiles from among the 10,000. Then a boundary model can make true

$$\sim\exists x \exists y (\mathcal{D}Fx \,\&\, Rxy \,\&\, \mathcal{D}{\sim}Fy)$$

even if it also makes true

$$\exists x \exists y (Rxy \,\&\, Fx \,\&\, {\sim}Fy)$$

These correspond to what we said above. And we may similarly say that though there is a boundary model b such that for each sentence s in the language, s is true if and only if b assigns s 1, there is no boundary model b such that it is determinate that each sentence s is true if and only if b assigns s one. Thus the boundary models can represent as true sentences about tiles which one would expect to be true if the Pure Qualification Approach is correct. And sentences of the same form as those represented as true can be used to express the extent to which boundary models can accurately represent a natural vague language: in particular, $\exists x\phi(x) \,\&\, {\sim}\exists x\mathcal{D}\phi(x)$. There is a “correct” boundary model, but not a determinately correct one.

It is also worth adding that though the boundary semantics gives a model for the existence of propositions that are true but not determinately true, for states of affairs that obtain but do not determinately obtain, there is a scheme that the semantics validates which helps remove the sense of mystery that there should be such things. For any ϕ,

$$\models \sim\mathcal{D}(\phi \,\&\, {\sim}\mathcal{D}\phi).$$

This corresponds to the plausible thought that nothing could be *determinately* both the case and not determinately the case. If only what is determinate is properly assertible, than this helps to explain why no proposition of the form $\phi \,\&\, {\sim}\mathcal{D}\phi$ could be assertible, even if it is assertible that there be such propositions that are true.

Chapter 5

Indeterminacy of Identity

This chapter considers some of the issues that surround indeterminacy of identity: why it might be real, the distinctive and serious philosophical issues that it raises, and the related "problem of the many". Some suggestions are made about dealing with these issues, but many of them are provisional and tentative.

5.1 Why there might be indeterminacy of identity

There are a number of sorts of plausible examples of situations in which there are indeterminate identity propositions. Two main sorts are cases of fission and fusion. In a case of fission, one thing enters a change, and two come out. The reproductive splitting of an amoeba is a common actual example of fission. A hypothetical example considered frequently in the philosophical literature on personal identity is the fission of a person: a person enters the mad-scientist's laboratory, and two people come out, each intimately metaphysically tied to the one that entered. There are many variations on this theme that differ over the details of the process that occurs in the laboratory. Fusion is the opposite of fission: two become one somehow.

We will not labor over the details of the variety of cases that lead one to consider the possibility of indeterminacy of identity. We may focus upon two different sorts: indeterminacy of survival and fission. The sort of thing I mean to call "indeterminacy of survival" is very inclusive: many examples of fission would count as cases of indeterminacy of survival. The basic scheme is just this: an object undergoes a change that leaves us in doubt about whether the object survived the change. But this is the form of doubt characteristic of vagueness, and not typical ignorance. One can even construct *sorites* series to show that there are such cases: at one end, imagine a process that will obviously leave the object intact; at the other, imagine a process that will obviously destroy the object. And make these choices in such a way that there lie, in between the two of them in the space of possible processes, a series of many pairwise very similar processes. For a concrete example, let the object be a wooden baseball bat. Our *sorites* series of processes is this: at one end, the process of removing the outermost 1/10,000$^{\text{th}}$ of the bat and incinerating the removed wood; at the other end, the process of removing the outermost 10,000/10,000$^{\text{th}}$ of the bat—the whole bat—and incinerating the removed wood. Term n of the 10,000 term series of processes is the removal and incineration of the outermost $n/10,000^{\text{th}}$ of the bat. The classic *sorites* paradox can be used for another such example: let term n of the series be the removal of n grains of sand from the heap.

These series suggest that there are objects and processes such that when the processes occur, it is indeterminate whether the object, which existed before the process occurred, still exists. Further, it may well seem that there exists, after the process occurred, *something*, such that it is indeterminate whether the thing that entered the process is identical with this second thing, which exists after the process is complete. Suppose that α enters the process, and β exists afterwards. It is indeterminate whether α still exists, and it may seem reasonable to think that it is also indeterminate whether $\alpha = \beta$. If α is the baseball bat, β would be a wooden object that exists after some amount of the wood of the bat has been removed and incinerated. If α is a person who undergoes a radical "treatment" in the mad-scientist's laboratory, β would be the person who walks out.

Cases of fission, in which an object α splits into two objects β_1 and β_2, may well be cases of indeterminacy of survival. And if it is indeterminate whether α survived the process, it may seem to be indeterminate whether $\alpha = \beta_1$ and whether $\alpha = \beta_2$. But cases of fission raise special difficulties: presumably it is determinate that $\beta_1 \neq \beta_2$, and thus it would seem determinate that it is not the case that both $\alpha = \beta_1$ and $\alpha = \beta_2$.

5.2 Difficulties making sense of indeterminacy of identity

5.2.1 Evans on vagueness in the world

The *locus classicus* of attacks on the coherence of the possibility of indeterminate identity is Gareth Evans' one page article "Can there be vague objects?"[7] Here are the formal steps of the argument.

Let "a" and "b" be singular terms such that the sentence "$a = b$" is of indeterminate truth value. We will use "$\mathcal{I}$" to represent "It is indeterminate whether"; Evans used "∇".

1. Suppose $\mathcal{I}(a = b)$.

2. Therefore $\hat{x}[\mathcal{I}(x = a)]b$.

   ("$\hat{x}[\mathcal{I}(x = a)]$" represents a property; Evans says that this step is valid since the first premise "reports a fact about b" and premise two "expresses" this fact too.)

3. $\sim\mathcal{I}(a = a)$

4. Hence $\sim\hat{x}[\mathcal{I}(x = a)]a$.

5. But from the second and fourth premises, together with Leibniz' law, we can get $\sim(a = b)$.

Evans says that $\sim(a = b)$ "contradict[s] the assumption, with which we began, that the identity statement "$a = b$" is of indeterminate truth value." It is not clear just what the contradiction is supposed to be, for

$$\sim a = b \,\&\, \mathcal{I}a = b$$

is not a formal contradiction. (As we saw in 4.4.5, it can be given a formal model which with some plausibility is appropriate for modeling the idea of indeterminate identity.) Regardless, it is true that it seems *unacceptable* that $\sim(a = b)$ should be a consequence of $\mathcal{I}(a = b)$ (and that worldly indeterminacy of identity should be possible). Thus if Evans has shown that this implication holds, he has indeed cast serious doubt upon the coherence of worldly indeterminacy of identity.

Just what is "worldly indeterminacy of identity", in Evans' view? He is explicit about the idea whose coherence he is attempting to challenge:

> It is sometimes said that the world might itself *be* vague. Rather than vagueness being a deficiency in our mode of describing the world, it would then be a necessary feature of any true description of it. It is also said that amongst the statements which may not have a determinate truth value as a result of their vagueness are identity statements. Combining these two views we would arrive at the idea that the world might contain certain objects about which it is a *fact* that they have fuzzy boundaries. But is this idea coherent?

Evans is too loose in this passage. It is not at all obvious that when we combine the two views described in the first three sentences, we get something about *fuzzy boundaries.* It is coherent that the world itself is vague, and that there be objects about which it is a fact that they have fuzzy boundaries, and that there are no *identity* sentences which are indeterminate in truth value due to worldly vagueness. This is because for objects to have fuzzy boundaries requires no more than this: it can be a fact about an object o that there exist another object p such that it is indeterminate whether p is part of o. More broadly, there may be no sharp line for the things that are parts of o. In such a situation, there need be no indeterminate identity sentences.[1]

[1]This is true even if there is an object that might be called "o minus p". For that may well be a vastly different kind of object from o, and hence determinately distinct from o.

This is a central point about Evans' argument that should be quite clear: worldly vagueness or indeterminacy does not require that there be identity sentences that are indeterminate as a result of vagueness. That there be objects with "fuzzy" or unsharp boundaries does not obviously require that there be such possibly problematic identity sentences.

Now Evans' argument does directly address the coherence of the idea that actually does result from combining the two views he mentions. If we entertain the idea that vagueness is (sometimes) directly the result of worldly indeterminacy, *and* sometimes identity sentences are indeterminate in truth value as a result of worldly indeterminacy, then we must consider Evans' argument.

5.2.2 Other issues reminiscent of Evans'

The indeterminacy of identity raises some other problems that echo Evans' argument. Consider again a process that α undergoes and that issues β, while it is indeterminate whether α is β. Presumably it is determinate that α existed before the process occurred, and it is indeterminate whether α exists afterward. Yet it is *indeterminate* whether β existed before the process occurred, and *determinate* that β exists afterward. On the face of it, this suggests that α and β have different properties: α has, while β lacks, the property of determinately existing before the process occurred. Moreover, suppose that it is determinate that β is in the room. Since it is indeterminate whether α exists, it is plausible that (at best) it is indeterminate whether α is in the room. Again, it seems that α and β have different properties: β has, and α lacks, the property of being determinately in the room. These observations, like Evans' argument, cast doubt on the coherence of the indeterminacy of identity.

If there is indeed a property of being determinately identical with o, and o has this property, and p does not, then it is extremely difficult to avoid the conclusion that o is not identical with p. But perhaps there is no property of being determinately identical with o. We might think of the situation with indeterminately identical objects o and p like this: *since* it is indeterminate whether o is identical with p, it must be indeterminate whether o and p have the same properties. And so we would deny that there is a property had by o but lacked by p called "being determinately identical with o"; and we would make a similar denial of a property called "being determinately in the room before the change".

5.3 Clearing up the possibility of indeterminacy of identity

There are at least two tasks that we can undertake that may help to clear up the possibility of indeterminacy of identity. The easier task is the task of providing a formal model that may help us to understand the situation, and what might be wrong with the arguments. The harder task is to give a deep metaphysical explanation of the situation.

5.3.1 The formal representation of indeterminate identity

The formal task can be tackled in different ways. One way involves the paradigm of treating vague language with a three-valued or gappy logic. Terry Parsons and Peter Woodruff have explored this route in depth in [23]. Another way, original with this dissertation, is suggested by the variable-domain boundary models of section 4.4.5. The logical paradigm here is the Pure Qualification Approach, with its embrace of classical logic.[2] The general shape of

[2]The commitment to classical logic is in one way not complete, for the logic of identity sentences might be called "non-classical", given the facts discussed at the end of 4.4.5. The issue is delicate, for when sentences (even involving identity) do not contain $\mathcal{D}$ they logically relate to one another just as they classically should.

the formal treatment can be appreciated with the following extremely simple model. Let us suppose that there are just two relevant properties, and that the following occurs: object α determinately exists, a process occurs, and afterward, object β determinately exists, but it is indeterminate whether $\alpha = \beta$. We will consider time to be "built-in" to properties, so that sentences are simply true true or false, not true or false at a time. Our quantifiers range over all objects at all times, not just over presently existing objects, and propositions need not be indexed to times. Let us further suppose that it is determinate that α is F (think of F as "in the room before the change") and it is determinate that β is G ("in the room after the change").

On our formal treatment, there will be two objects, representing α and β, in the master domain D of an appropriate variable-domain boundary model. (For simplicity, we will talk about them as if they were α and β themselves. But they are not; they are merely representatives of α and β. It is indeterminate whether α is β, but it is determinate that the representative in the model of α is distinct from the representative of β.) The other relevant facts about an appropriate model b are these: its first-level element b_1 will contain classical models appropriate for D, one of which (c_0) will contain a single object in its local domain, and the other of which (d_0) will contain two objects. One can think of these two classical models as each appropriate for one of the two possible "precisifications" of the facts about α and β. One possible precisification has it that $\alpha = \beta$, and the other that $\alpha \neq \beta$; the other facts adjust accordingly. Thus α and β are both mapped to the single object γ in the domain of c_0, and are mapped to the two objects ϵ and δ in d_0. To fill things out appropriately, γ will satisfy both F and G in c_0, and only ϵ will satisfy F in d_0, while only δ will satisfy G in d_0. Let us use o as a constant that refers to α in b and p as a constant that refers to β in b. The following sentences will be true in b:

$$\mathcal{I}o = p$$

$$\mathcal{D}(Fo \,\&\, Gp)$$

$$\mathcal{I}Go \,\&\, \mathcal{I}Fp$$

Also true will be

$$\mathcal{D}(Fp \rightarrow p = o)$$

That this is true begins to show the difference of the current formal approach from the formal approaches that go with truth-functional three-valued or gappy logics. For on those approaches, the conditional $\phi \rightarrow \psi$ is often given a treatment that gives it the same truth value as the disjunction $\sim\phi \vee \psi$. On this treatment, given that both the antecedent and the consequent of $Fp \rightarrow p = o$ are neither true nor false, the conditional itself is neither true nor false, and hence prefixing it with $\mathcal{D}$ certainly cannot yield a truth. Other important three-valued treatments of the conditional make it true when the antecedent and consequent are both neither true nor false. On such treatments, $Fp \rightarrow p = o$ will come out true. But on these treatments of the conditional, it looks as though

$$\mathcal{D}(Fp \rightarrow p \neq o)$$

will come out with the same truth value as $\mathcal{D}(Fp \rightarrow p = o)$. It is unintuitive that these sentences be equally true, and our approach in fact makes the one true and the other false.

Another sentence to consider is

$$\mathcal{D}\forall x(x = o \rightarrow Fx)$$

Under an assignment that assigns object p to the variable x, both the antecedent and the consequent of the embedded conditional are neither true nor false (on those approaches).[3] This raises issues like those raised for $\mathcal{D}(Fp \rightarrow p = o)$. I believe that these sentences represent propositions that intuitively should be true in the imagined circumstances. $\mathcal{D}(Fp \rightarrow p = o)$ says that (it is determinate that) if p is F—if α was in the room before the

[3]Of course this depends upon how quantification and identity are handled in a gappy system.

change—then $p = o$—then α is identical with β. The second says that (it is determinate that) for each thing, if that thing is o, then it is F—if a thing is α, then it was in the room before the change.

If there are no other objects in the master domain whose representatives satisfy F or satisfy G in any of the classical models in b_1, then

$$\mathcal{D}(\exists x Gx \,\&\, \forall y(Gy \rightarrow x = y)))$$

is true. This is intuitive: it says that it is determinate that there is exactly one G—exactly one thing in the room after the change. Yet it too will not come out true on the gappy approaches, for both Go and $o = p$ should be neither true nor false, as there are substitution instances of $Gy \rightarrow x = y$ which are neither true nor false.

5.3.2 Indeterminacy of identity and formal semantics

There is an issue that arises for any formal semantical treatment of vague language. Suppose we want to give a formal treatment of a language which can discuss the facts about the objects α and β, as described above. The sentence $\mathcal{D}o = o$ that expresses the proposition that it is determinate that α is identical with α should have a different truth value from the sentence $\mathcal{D}o = p$ that expresses the proposition that it is indeterminate whether α is identical with β. Since the only syntactic difference between these two sentences is a single symbol (o or p), it follows that the symbols o and p must be semantically different. But if the semantic value of a constant symbol is just on object, then the two symbols must have different objects as their semantic values.

Suppose that the sole semantic value of a genuine name is an object. If a formal semantics is to represent this fact by assigning to each constant symbol nothing more than an object, then if it is to represent $\mathcal{D}o = o$ as true and $\mathcal{D}o = p$ as false, then it must assign different objects to o and p. This suggests that though the sentence $\mathcal{I}o = p$ might be "true" in the formal semantics, something is nevertheless fishy about the representation, for the semantic values of the constant symbols will have to be distinct. Further, if it is *determinate* what the formal "truth-values" of the sentences are, then there would not be indeterminacy about the fact that the object assigned to o is distinct from the object assigned to p. In this situation, it cannot be that both the semantic value of o is α and the semantic value of p is β, for then it would be indeterminate whether the semantic value of o was the same as the semantic value of p, and so it would not be determinate that $\mathcal{D}o = o$ differs in truth value from $\mathcal{D}o = p$.

Setting aside systems in which the formal semantic value of a constant symbol is other than a mere object, we can say that no matter what the details are, the formal semantics will diverge from reality itself in this regard: in reality, it is indeterminate whether the names of α and β co-refer, while it is determinate that o and p do not co-refer. That o and p have determinately distinct semantic values seems an unavoidable "artifact" of any formal semantics on which the sentences $\mathcal{D}o = o$ and $\mathcal{D}o = p$ determinately have different truth values.

That this is the case is in one way disappointing, for it shows that we will not find a formal model-theoretic semantic system in which there is some model $\mathfrak{A}$ such that reference in vague language $\mathcal{L}$ is nothing but "reference" in $\mathfrak{A}$. But the unavoidable artifact just identified is reminiscent of the facts discussed in section 3.2.2 and elsewhere in section 3.2, which show that "truth" in models will differ from real truth in a vague language. Thus we might accept the "artifact" as another reflection of the fact that the world is not a formal system, a fact that comes out particularly clearly when we consider vagueness.

But the present case is rather subtle, and there are avenues that need further consideration. For, setting formal semantics aside, and letting 'α' and 'β' name the objects whose identity is indeterminate, we still have to deal with the fact that there would seem to be sentences of our language which differ only in the presence of one occurrence of these names and which (determinately) have different truth values. This should give pause to those who think that the semantic value of a typical name is nothing but an object. But there is also some doubt that this plausible doctrine about names is consistent with the facts about sentences of the form "Pierre believes that S" (and others). Whether this doubt can be allayed by a better understanding of the semantics of "believes that" sentences may not yet be clear, but it may well be that the doctrine about reference is correct. And perhaps a better understanding of "it is determinate that" will show the doctrine to be consistent with the possibility of indeterminacy of identity.

5.3.3 The metaphysics of indeterminacy and indeterminate identity

In a case in which it is indeterminate whether α is identical with β, and yet α determinately has some property ϕ (like being in the room before the change), and it is not determinate that β has ϕ, it appears at first that α and β must *differ*, since the one is *determinately* ϕ while the other is not. The only way to avoid this conclusion, it seems, is to deny that there is a property, in addition to ϕ, called "being determinately ϕ".[4]

It is true that it is painfully difficult to explain or even to grasp the dif-

[4]This line of approach is considered in [23].

ference between being red and being determinately red, and just as difficult to explain the difference between not being red and not being determinately red.[5] Yet from β's not being determinately in the room it should not follow that β is not in the room.

This is a very difficult subject, and my best attempts at dealing with it have seemed to me too unsatisfactory to report in any detail. The direction I suspect we must go is, very roughly, this: Perhaps propositions about determinacy (and indeterminacy) are not directly about the facts in the way that propositions which do not involve determinacy are. Determinacy and indeterminacy may concern "aspects" or "qualifications" on facts, and may not generate higher "levels" of facts, or may generate facts of an entirely different kind.

5.4 McGee on indeterminate identity

We turn now to consider one last argument against indeterminate identity. It is in effect an extension of Evans' argument by Vann McGee. McGee's discussion

[5] The advocate of gappy logic might suggest here that at least there is clearly a difference between not being red and not satisfying "is red". But to identify being determinately red with satisfying "is red" is just as clearly unsatisfactory. When we say that there is indeterminacy about what is red, we do not mean to talk about language.

is relevant to this work because it is especially clear, is closely associated with the ideas called "supervaluationism", some of which are similar to, but distinct from, ideas in this dissertation, and because it will lead naturally to a brief consideration of Unger's "Problem of the Many".

McGee's argument, in *'Kilimanjaro'*[21], begins with a discussion of Evans' argument, and then goes on to explain how McGee's own version of supervaluationism can be brought to bear both on Evans' formal argument and on the issue of the indeterminacy of identity generally. Broadly, his argument runs like this: if 'Kilimanjaro' names an object with an indeterminate identity—a vague object—then it can be instantiated inside of the context of the "determinately" operator; but such instantiation will lead to absurdities, and hence 'Kilimanjaro' does not name a vague object. As I take him, he intends that 'Kilimanjaro' would name a vague object if there were any vague objects, and the name can stand in for any purported name of a vague object in his arguments.

Let us turn to the details. First, let us consider McGee's treatment of the truth conditions for the "determinately" operator, which he symbolizes as '$\Box$'. He writes:

> The Supervaluation Hypothesis enables us to apply the model theory of modal logic to the study of the 'determinately' operator, since we can treat the acceptable $\mathcal{U}$-models like possible world. As long as we are not interested in nested '$\Box$'s, we can give the

satisfaction condition for $\ulcorner \Box\phi \urcorner$ (read 'determinately ϕ') as follows:

> σ satisfies $\ulcorner \Box\phi \urcorner$ in $\mathfrak{A}$ if and only if σ satisfies ϕ in every acceptable $\mathcal{U}$-model.[6]

A problem with this satisfaction condition is that it alone leads to the truth of a sentence which expresses an unacceptable principle, given that we are trying to give a coherent account of indeterminate identity sentences. With this satisfaction condition, the sentence

$$\forall x \forall y(x = y \leftrightarrow \Box x = y)$$

is validated. Thus so is

$$\forall x \forall y(\sim\Box(x = y) \leftrightarrow \sim x = y)$$

This principle is closely related to Evans' argument, which suggests that for arbitrary names a and b, $\mathcal{I}a = b$ entails $\sim a = b$. The principle is unacceptable because it says, in effect, that if there are x and y such that it is not determinate that x is identical with y, then x is not identical with y. If there can really be indeterminate identity statements, we should not be able to infer from the indeterminacy of x's being identical with y that x is not y. It can be regarded as an unacceptable artifact of McGee's satisfaction condition that this principle is validated. The formal system of 4.4.5 deals better with the possibility of

[6]McGee, [21] pp. 156–157.

indeterminate identity sentences, as we just saw. Thus McGee's construal of the determinacy operator seems to stack the deck against vague objects at the outset.

Proceeding in his argument against vague objects, McGee distinguishes between what he calls *rigid* and *non-rigid* singular terms. A singular term is rigid just in case it is assigned the same object in all acceptable $\mathcal{U}$-models. (Recall that a $\mathcal{U}$-model is a model whose domain is $\mathcal{U}$. McGee supposes all along that there is a domain $\mathcal{U}$ such that there is a class of $\mathcal{U}$-models such that a sentence is determinately satisfied just in case it is satisfied in every $\mathcal{U}$-model.) Now there are two possibilities for a name like 'Kilimanjaro': it is rigid or not. McGee takes his technical notion of rigidity to coincide with the idea that a name be a "precise term denoting a vague object", while a non-rigid name is a name about which it is indeterminate what the name refers to.

Next McGee suggests that there is some very large (but finite) list of objects such that anything of the kind he takes Kilimanjaro to be of—land mass in Tanzania—is on the list.

> Let us assume, for the moment, that a land mass is fully specified by determining what its constituent molecules are. That is, for any collection of molecules, there is uniquely determined at most one land mass that has those molecules as its parts. Let us also assume that for any given molecule, our language provides a description in the language that is true of that molecule and no

> other. This is so because molecules are spatially discrete, so that, to uniquely pick out a molecule, it is enough to find a region that contains that molecule and no others. (There is fuzziness about what counts as a molecule, and there are problems about the exact specification of spatial locations, but let's not worry about those now.) We also assume that there isn't any land mass that contains part but not all of a given molecule.
>
> Given these assumptions, every land mass can be precisely specified by stating what its constituent molecules are. Thus there is a (very long) list $a_1, a_2, \ldots, a_N$ of precise names such that every land mass in Tanzania is sure to be named somewhere on the list.[7]

It is the last proposition of this passage that is at the core of McGee's argument against vague objects. The preceding text is meant to warm us up to this central proposition; he will later consider what happens if he gives up on some of the assumptions in that paragraph, and argues that, in effect anyway, the complications that cast doubt on the assumptions of that paragraph do not cast doubt on the central point.

Let us see how the central point leads to the rejection of vague objects. The argument, extracted from pages 159–160 of [21] is this:

1. There is a list $a_1, a_2, \ldots, a_N$ such that every land mass in Tanzania is sure to be named somewhere on the list.

2. The following sentence is determinately true

$$(\forall x)(\text{x is a land mass in Tanzania} \leftrightarrow \bigvee_{i \leq N} x = a_i).$$

[7] McGee [21], p. 159.

3. Kilimanjaro is a land mass in Tanzania.

4. So we can derive

$$\bigvee_{i \leq N} \text{Kilimanjaro } = a_i \, .$$

5. We also have

$$\forall x \forall y (x = y \leftrightarrow \Box x = y).$$

6. Since each a_i is a rigid designator (precise term), we have

$$(\forall x)(x = a_i \leftrightarrow \Box x = a_i).$$

7. If ‘Kilimanjaro’ is a precise term (denoting a vague object), then we can use it to instantiate into ‘determinately’ contexts, in which case we get

$$(\text{Kilimanjaro} = a_i \leftrightarrow \Box\text{Kilimanjaro} = a_i)$$

and also

$$\bigvee_{i \leq N} \Box\text{Kilimanjaro} = a_i \, .$$

8. But this last formula is absurd.

9. Hence ‘Kilimanjaro’ is not a precise term denoting a vague object.

The reason that the last formula is absurd is that each of the a_i’s is supposed to be a name of a *precise* land mass; that was the point of the passage quoted above.

I do not challenge the validity of this argument. I have already suggested that the fourth premise is dubitable.[8] But the more currently relevant objection to the argument is that, taken as I stated it, the first premise must be strengthened to something implausible in order to guarantee that the last formula is indeed absurd. Consider instead of the current first premise, this stronger one:

- There is a list *of precise objects* $a_1, a_2, \ldots, a_N$ such that every land mass in Tanzania is sure to be named somewhere on the list.

Without this stronger premise, the last formula (see premise eight) is not absurd, for, as defenders of vague objects, we could simply say that Kilimanjaro, the vague object, should go on the list, and then it is no shocker that

$$\bigvee_{i \leq N} \Box \text{Kilimanjaro} = a_i$$

is true. But when the premise is strengthened as above, the argument obviously seems question-begging: It is exactly because Kilimanjaro is a vague object that there is no list of precise objects such that every land mass in

[8]The formal system of 4.4.5 does not validate $\forall x \forall y (x = y \leftrightarrow \Box x = y)$ and so suggests an alternative response to McGee's argument. We might accept that it is determinate that Kilimanjaro is identical with one of the a_is, but reject the assertion that this entails that there is an a_i such that it is determinate that Kilimanjaro is identical with a_i. This option seems to me, however, not to take seriously the fact that Kilimanjaro, a mountain, is a different kind of thing from any of the precisely delineated "land masses", as I will discuss in 5.5.

Tanzania is sure to be on the list. (Either that, or Kilimanjaro is not a "land mass". If we read "land mass" as "mereological sum of molecules" then the strengthened premise one is fine, but premise two fails, for it is quite implausible that Kilimanjaro is a mereological sum of molecules. For some molecules are part of Kilimanjaro now, but not last year, yet every part of every mereological sum of molecules is forever a part of that sum of molecules.)

McGee is aware of this sort of objection, and respects the fact that it is implausible that a mountain is a mereological sum of molecules. But, he thinks, if we replace the a_i's with sums of *time-world slices of elementary particles*, the resulting argument will be sound. (We cannot use sums of time slices of elementary particles, for those have modal properties which differ from the modal properties of mountains.) Setting aside for the moment the exact nature of sums of time-world slices of elementary particles, we can see that the point is, in effect, to justify the strengthened premise one. But the strengthened premise one simply begs the question against the existence of "vague objects" which are land masses in Tanzania. Granting premise two, premise one is question begging; denying premise two, the argument is unsound.

McGee suggests that the lesson of his argument (coupled with what he has borrowed from Gareth Evans' argument about the indeterminacy of identity) is this:

> What we get from Evans's argument is the conditional: If there are, indeed, precise objects in the great profusion mereology postulates, then there are not vague objects as well. If mereology is right, there are so many precise objects that they crowd all the available space, leaving no room for vague objects too. In view of the great credibility mereology commands, this is a good, though not an inescapable argument against vague objects.[9]

(McGee sees his argument as a sort of extension of Evans'.)

I think that even the conditional has not been shown, for the mere existence of the many precise objects does not at all show that they are the only candidate objects for the reference of 'Kilimanjaro'. To say that they are is in effect to assert the strengthened premise one, and no special grounds have been given for this assertion.

Setting the technical aspects of McGee's argument aside and looking at it's metaphysical core, it can usefully be seen as a variation of Unger's "Problem of the Many" arguments. Like Unger, McGee is suggesting that there are many (precise) things that are equally good and maximally good candidates for "playing the role of" a macroscopic object like Kilimanjaro. McGee takes this thought and interweaves it with the supervaluational formal apparatus, concluding that the sentence that says, in effect, "Kilimanjaro is identical with one of those (precise) objects" is determinately true. The many of Unger's argument become the many (and for some reason, the only) candidates for the

[9] McGee, [21], p. 163.

reference of the name; in McGee's picture, they are the admissible precisifications of the name.

5.5 The problem of the many

The *locus classicus* of the "problem of the many" is Peter Unger's paper of that name. [28] David Lewis' "Many, but almost one" [18] addresses the problem crisply, and I will borrow from his presentation.

There is one cat on the mat, we would normally say. But look closer and you see that there are some hairs that are not clearly attached and not clearly not attached to the body of the cat. Look much closer and you'll see very many little particles at and near the cat's surface, and you'll find that it is unclear which ones are parts of the cat's matter and which are not. This suggests (according to Lewis) that there are very many precisely delineable objects—*p-cats*—which differ very slightly from one another, and which are on the mat.[10]

All these p-cats have an equal "claim" to being *the* cat on the mat. Unger

[10] What exactly these "objects" are is not clear. Lewis seems to have in mind something like the material content of a precisely delineated region of space-time or the sum of qualitative tropes in the region. But it seems that such a thing is not a cat, since a cat could have occupied a very different region of space-time than the one it happens to occupy, whereas the material content of a region of space-time could not have.

would go on to argue roughly as follows: since if there are cats there must be just one on the mat, and since there are very many equally qualified candidates for the role of being the unique cat on the mat, there is not *really* just one cat on the mat, and hence there are no cats. Lewis attempts to save the phenomena, however, and offers supervaluations and a notion of *partial* identity as a means to making ordinarily believed sentences like "There is just one cat on the mat" come out true.[11]

Unger is straightforwardly revisionary: our ordinary beliefs are just false. Lewis attempts to save our ordinary beliefs by interpreting them. I will not directly address Unger or Lewis here; Lewis' position in particular is subtle, and a full discussion would take too much space. Instead I will offer my own response to the problem, and relate it to the larger issue of this section of this paper.

What to do about it

My attitude is that at bottom, the problem of the many is a problem that needs to be dissolved rather than solved. Insofar as it is granted, for whatever reason, that there are typically very many things on the mat that are equally

[11]As discussed in 4.1.1, "supervaluation", in the literature covers both an informal idea about the nature of vagueness in language and a formal idea that develops the informal one. Lewis employs both, while I advocate the formal idea as divorced from the informal idea.

good “candidates” for the “role” of the cat on the mat, and thus no one thing that best plays this role, to that extent it is granted that our ordinary beliefs are radically mistaken. “Interpret” them as you will, make them “come out true” by clever machinations; none of these maneuvers really rescue us from radical error in our ordinary beliefs—if what has been granted is so.[12]

The reasons for granting such a thing must be extremely strong. It seems to me that the sensible thing to do when presented with reasons to believe such a thing is to doubt whatever propositions are offered. The presumption should be if some propositions together entail that there are at best many close competitors for the role of the cat on the mat, then those propositions are pretty surely not all true. It it is then the philosopher’s job to figure out why they are not all true, rather than to announce the discovery that there aren’t cats, or that there are millions of cats on the mat, or that our ordinary beliefs cannot be taken at face value. If there are very many p-cats on the mat, this just shows that whatever p-cats are, they are a very different kind of thing from cats.

---

[12]There may be a way to interpret the belief that the region of space taken up by a cat is pretty much filled with matter so that it comes out true. Yet there is a surprising fact, discovered by chemists and physicists, that this belief, *taken at face value*, turns out to be *false*, even if the cat has just eaten a huge meal. Imagine (if you can) a biologist announcing the discovery that there are actually very many life forms, each about the size of a cat, sharing the space on the mat.

The foregoing is more attitude than substantive argument; it sets the task of arguing against purported reasons to believe that there are really many equally good “choices” for the role of the cat on the mat. I believe that even if I am not up to the task of finding the problems with such purported reasons, someone will be.

Concluding Remarks

There is exactly one cat on the mat, with fuzzy boundaries, and there is no other thing of a similar kind on the mat, with which the cat must compete for our attention. We have come full circle, back to the idea that some of the vagueness in our language is a result of worldly indeterminacy. We have seen, along the way, that though the notion of the lack of sharp boundary is connected with indeterminacy, the notion of the lack of sharpness is not easy to reduce to the notion of indeterminacy. We have also seen that when its depth is appreciated, indeterminacy inevitably leads to puzzles. It has been argued that the real heart of the puzzle is the fact that something can be the case without being determinately the case, and that acceptance of this fact leads to the smoothest conception of the effect of vagueness upon logic, the Pure Qualification Approach. We may always qualify, and we may add that nothing can be *determinately* both the case and not determinatly the case. Boundary Semantics gives us a formalization of the PQA, and provides us with a bare formal system for indeterminate identity. Though arguments against the possibility of indeterminacy of identity can be answered up to a point, there are special metaphysical problems that need further examination, and it is clear that fuzzy boundaries do not logically require indeterminacy of identity.

Bibliography

[1] The paradox of the heap. In U. Monnich, editor, *Aspects of Philosophical Logic*, pages 225–77. Reidel, Dordrecht, 1981.

[2] Robert Merrihew Adams. Theories of actuality. *Noûs*, 8:211–231, 1974.

[3] Michael Ayers. *Locke: Epistemology and Ontology*. Routledge, 1991. Two Volumes.

[4] P. Bosch. 'Vagueness' is context-dependence: A solution to the sorites paradox. In T. Ballmer and M. Pinkal, editors, *Approaching Vagueness*, North-Holland Linguistic series, pages 189–210. Elsevier Science Publishers, Amsterdam, 1983.

[5] Daniel C. Dennett. Real patterns. *Journal of Philosophy*, 88(1):27–51, January 1991.

[6] Daniel C. Dennett. *Darwin's Dangerous Idea*. Simon & Schuster, New York, 1995.

[7] Gareth Evans. Can there be vague objects? *Analysis*, 38:208, 1978. Reprinted in [13], 317.

[8] Kit Fine. Vagueness, truth, and logic. *Synthese*, 30:265–300, 1975. Reprinted in [13], 119–50.

[9] Delia Graff. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics*. Forthcoming.

[10] Sören Halldén. *The Logic of Nonsense*. Uppsala Universitets Arsskrift, 1949.

[11] Terry Horgan, editor. *Vagueness*. 1994. Proceedings of the Spindel Conference 1994. Southern Journal of Philosophy 33, Supplement.

[12] H. Kamp. Two theories about adjectives. In E.L. Keenan, editor, *Formal Semantics of Natural Language*, pages 123–155. Cambridge University Press, Cambridge, 1975.

[13] Rosanna Keefe and Peter Smith, editors. *Vagueness: A Reader*. The MIT Press, Cambridge and London, 1996.

[14] David Lewis. General semantics. *Synthese*, 22:18–67, 1970. Reprinted in [16], 189-229.

[15] David Lewis. Scorekeeping in a language game. *Journal of Philosophical Logic*, 8:339–59, 1979. Reprinted in [16], 233-49.

[16] David Lewis. *Philosophical Papers*, volume 1. Oxford Univeristy Press, New York and Oxford, 1983.

[17] David Lewis. *On the plurality of worlds*. Basil Blackwell, Oxford, 1986.

[18] David Lewis. Many, but almost one. In Keith Campbell, John Bacon, and Lloyd Reinhardt, editors, *Ontology, Causality, and Mind: Essays on the Philosophy of D.M. Armstrong*. Cambridge University Press, Cambridge, 1993. Reprinted in [19] 164–182.

[19] David Lewis. *Papers in Metaphysics and Epistemology*. Cambridge Univeristy Press, Cambridge, 1999.

[20] Vann McGee. *Truth, Vagueness, and Paradox*. Hackett, Indianapolis, 1991.

[21] Vann McGee. 'Kilimanjaro'. In Ali Kazmi, editor, *Meaning and Reference*, Canadian Journal of Philosophy Supplementary Volume 23. University of Calgary Press, Calgary, 1997.

[22] Henryk Mehlberg. *Truth and Vagueness*. University of Toronto Press, 1958. Reprinted in [13], 85–88. Originally § 29 of *The Reach of Science*.

[23] Terence Parsons and Peter Woodruff. Indeterminacy of identity of objects and sets. In *Mind, Causation, and World*, number 11 in series *Philosophical Perspectives*, pages 321–348. 1997.

[24] Warren Quinn. Abortion: Identity and loss. *Philosophy and Public Affairs*, 13:24–54, 1984.

[25] Bertrand Russell. Vagueness. *Australasian Journal of Philosophy and Psychology*, 1:84–92, 1923. Reprinted in [13], 61–68.

[26] R. M. Sainsbury. Concepts without boundaries, 1990. Inaugural lecture published by the King's College London Department of Philosophy. Reprinted in [13], 251–64.

[27] M. Tye. Vague objects. *Mind*, 99:535–557, 1990.

[28] Peter Unger. The problem of the many. In *Midwest Studies in Philosophy*, number 5, pages 411–467. 1980.

[29] B.C. van Fraassen. Singular terms, truth-value gaps, and free logic. *Journal of Philosophy*, 63:481–95, 1968.

[30] Timothy Williamson. Definiteness and knowability. 1994. In [11].

[31] Timothy Williamson. *Vagueness*. Routledge, London and New York, 1994.

[32] L.A. Zadeh. Fuzzy sets. *Information and Control*, 8:407–428, 1965.