THE PROBLEMS OF PHILOSOPHY

VAGUENESS

TIMOTHY WILLIAMSON

ROUTLEDGE

Vagueness

If you keep removing single grains of sand from a heap, when is it no longer a heap? This question, and many others like it, soon lead us to the problem of vagueness.

Timothy Williamson traces the history of the problem from discussions of the heap paradox in ancient Greece to modern formal approaches, such as fuzzy logic. He discusses the view that classical logic and formal semantics do not apply to vague languages and shows that none of the alternative approaches can give a satisfying account of vagueness without falling back on classical logic.

Against this historical and critical background, Williamson then develops his own epistemicist position. Vagueness, he argues, is an epistemic phenomenon, a kind of ignorance: there really is a specific grain of sand whose removal turns the heap into a non-heap, but we cannot know which one it is.

Williamson's argument has ramifications far beyond the study of vagueness. It reasserts the validity of classical logic and semantics; more generally, it makes the thoroughly realist point that even the truth about the boundaries of our concepts can be beyond our capacity to know it.

The approach throughout keeps technicalities to a minimum; this is partly to counter the illusion, encouraged by the emphasis on formal systems, that vagueness can be studied in a precise metalanguage. For the technically minded, an appendix shows how the epistemic view can be formalised within the framework of epistemic logic.

Timothy Williamson is Professor of Logic and Metaphysics at the University of Edinburgh. He is the author of *Identity and Discrimination.*

The Problems of Philosophy

Founding editor: Ted Honderich
Editors: Tim Crane and Jonathan Wolff, *University College London*

This series addresses the central problems of philosophy. Each book gives a fresh account of a particular philosophical theme by offering two perspectives on the subject: the historical context and the author's own distinctive and original contribution.

The books are written to be accessible to students of philosophy and related disciplines, while taking the debate to a new level.

DEMOCRACY*
Ross Harrison

THE EXISTENCE OF THE WORLD*
Reinhardt Grossman

NAMING AND REFERENCE
R. J. Nelson

EXPLAINING EXPLANATION*
David-Hillel Ruben

IF P, THEN Q*
David H. Sanford

SCEPTICISM
Christopher Hookway

HUMAN CONSCIOUSNESS
Alastair Hannay

THE IMPLICATIONS OF DETERMINISM
Roy Weatherford

THE INFINITE*
A. W. Moore

KNOWLEDGE AND BELIEF
Frederic F. Schmitt

KNOWLEDGE OF THE EXTERNAL WORLD
Bruce Aune

MORAL KNOWLEDGE*
Alan Goldman

MIND–BODY IDENTITY THEORIES*
Cynthia Macdonald

THE NATURE OF ART*
A. L. Cothey

PERSONAL IDENTITY
Harold W. Noonan

POLITICAL FREEDOM
George G. Brenkert

THE RATIONAL FOUNDATIONS OF ETHICS*
T. L. S. Sprigge

PRACTICAL REASONING*
Robert Audi

RATIONALITY*
Harold I. Brown

THOUGHT AND LANGUAGE*
J. M. Moravcsik

THE WEAKNESS OF THE WILL
Justine Gosling

THE MIND AND ITS WORLD*
Gregory McCulloch

PERCEPTION
Howard Robinson

THE NATURE OF GOD
Gerard Hughes

*Also available in paperback

Vagueness

Timothy Williamson

ROUTLEDGE

London and New York

First published 1994
by Routledge
11 New Fetter Lane, London EC4P 4E
29 West 35th Street, New York, NY 10001

This edition published in the Taylor & Francis e-Library, 2001.

Paperback edition 1996

British Library Cataloguing in Publication Data
A catalogue record for this book is available from the British Library

Library of Congress Cataloging-in-Publication Data
A catalogue record for this book is available from the Library of Congress

ISBN 0–415–03331–4 (hbk)
ISBN 0–415–13980–5 (pbk)
ISBN 0-203-01426-X Master e-book ISBN
ISBN 0-203-17453-4 (Glassbook Format)

For Nathan Isaacs and Jeff Williamson

Contents

Preface

This book originated in my attempts to refute its main thesis: that vagueness consists in our ignorance of the sharp boundaries of our concepts, and therefore requires no revision of standard logic. For years I took this epistemic view of vagueness to be obviously false, as most philosophers do. In 1988 Simon Blackburn, then editor of the journal *Mind*, asked me to review Roy Sorensen's intriguing book *Blindspots*, which includes a defence of the epistemic view. It did not persuade me; I could not see what makes us ignorant, and Sorensen offered no specific explanation. An alternative treatment of vagueness, supervaluationism, looked more or less adequate – unlike other popular alternatives, such as three-valued and fuzzy logic, which on technical grounds have always looked like blind alleys. However, I continued to think about the epistemic view, for the standard objections to it did not seem quite decisive. It was not clear that they did not assume a suspect connection between what is true and what we can verify. It then struck me that the notion of a margin for error could be used to give a specific explanation of ignorance of the sharp boundaries of our concepts, and the epistemic view began to look more plausible. A limited version of it was tentatively proposed in my book *Identity and Discrimination* (Oxford, Blackwell, 1990). The more closely the objections to it were analysed, the weaker they seemed. The next step was to focus on the fact that the meaning of vague expressions can be stated only in a language into which those expressions can be translated; it is a mistake to treat the language in which one theorizes about vagueness as though it were precise. Mark Sainsbury's inaugural lecture at King's College London, 'Concepts without Boundaries', helped to bring the significance of this point home to me, although we used it in quite different ways. It permits the formulation of arguments against a wide range of non-epistemic views, including the supervaluationism that had previously looked adequate (my objection to it, however, is not the one made in Sainsbury's lecture). The balance of arguments seemed to have moved firmly onto the side of the epistemic view. A book-length treatment was clearly needed. This is the result.

Some of the research for this book was carried out in late 1990 whilst I was a Visiting Fellow at the Research School of Social Sciences of the Australian National University in Canberra. Many people helped to make the visit a success; Philip and Eileen Pettit stand out. Gratitude is also due to University College Oxford and Oxford University for allowing me extra leave of absence in that academic year.

Ted Honderich kindly permitted me to substitute a volume on vagueness in this series for one planned on another subject. One result of working on the past of the problem of vagueness for which I am particularly grateful is a better sense of the richness of Stoic logic. In this connection, I thank David Sedley for permission to quote translations from the first volume of a work he edited with A.A. Long, *The Hellenistic Philosophers* (Cambridge, Cambridge University Press, 1987).

Parts of Chapters 7 and 8 are drawn from two previously published articles of mine: 'Vagueness and ignorance', *Aristotelian Society*, suppl. 66 (1992), 145–62, and 'Inexact knowledge', *Mind*, 101 (1992), 217–42. I am grateful to the Aristotelian Society and The Mind Association for permission to use this material.

For written comments on predecessors of parts of this book, many thanks go to Michael Bacharach, Justin Broackes, Myles Burnyeat, Peter Carruthers, Bill Child, Jack Copeland, Dorothy Edgington, Timothy Endicott, Graeme Forbes, Brian Garrett, Bill Hart, Dominic Hyde, Frank Jackson, Rosanna Keefe, Peter Lipton, Andrei Marmor, Gregory McCulloch, Karina and Angus McIntosh, David Over, Peter Pagin, Philip Percival, Philip Pettit, Mark Sainsbury, David Sedley, Jonathan Sutton, Charles Travis and David Wiggins. Peter Simons replied to 'Vagueness and ignorance' in an enjoyable symposium at Reading, chaired by Mark Sainsbury. More people than I can name helped with critical questions after talks on the epistemic view of vagueness, inexact knowledge and related

topics at the universities of Bradford, Bristol, Cambridge (the Moral Sciences Club), Dundee, Edinburgh, Heidelberg, Leeds, Lisbon, London (University College), New England (Armidale), Nottingham, Oslo, Oxford, Queensland, Stirling, Stockholm and Uppsala, the Australian National University and Monash University, and to a meeting of the Lisbon Philosophical Society in May 1991, an Anglo–Polish Symposium on the Philosophy of Logic and Language at Oriel College Oxford in September 1991, the Second Workshop on Knowledge, Belief and Strategic Interaction at Castiglioncello in June 1992 and the Joint Session of the Aristotelian Society and the Mind Association at Reading University in July 1992. Invidiously, I pick out George Bealer and Peter Menzies, because there is a particularly direct causal link between their questions and sections of the book. Early versions of several chapters were used in classes at Oxford, and were considerably improved as a result; Ron Chrisley, Michael Martin and Roger Teichmann were particularly persistent questioners. I have also been helped by conversations with Maria Baghramian, João Branquinho, John Campbell, David Charles, Kit Fine, Olav Gjelsvik and Peter Strawson (not to mention anyone previously mentioned). Juliane Kerkhecker guided me through Lorenzo Valla's Latin.

Those who know Elisabetta Perosino Williamson will guess how she helped in the writing of this book, and how much. It is dedicated to a great-uncle and an uncle, whose open-minded rationality (amongst other things) I rightly tried to imitate, with only mixed success.

Introduction

Logicians are often accused of treating language as though it were precise, and ignoring its vagueness. Their standards of valid and invalid reasoning are held to be good enough for artificial precise languages, but to break down when applied to the natural vague languages in which we actually reason about the world that we experience. A perfectly precise language for such reasoning is an idealization never to be realized. Although we can make our language less vague, we cannot make it perfectly precise. If we try to do so by stipulating what our words are to mean, our stipulations will themselves be made in less than perfectly precise terms, and the reformed language will inherit some of that vagueness.

The problem is not confined to logic. Attempts to describe the semantics of natural languages in formal terms are also frequently supposed to ignore vagueness, and therefore to misdescribe the meanings of ordinary expressions. Of course, a theory might ignore vagueness and remain a useful approximation for some purposes, but it is also legitimate to ask what changes of theory are needed to take vagueness into account.

At the core of classical (i.e. standard) logic and semantics is the principle of bivalence, according to which every statement is either true or false. This is the principle most obviously threatened by vagueness. When, for example, did Rembrandt become old? For each second of his life, one can consider the statement that he was old then. Some of those statements are false; others are true. If all of them are true or false, then there was a last second at which it was false to say that Rembrandt was old, immediately followed by a first second at which it was true to say that he was old. Which second was that? We have no way of knowing. Indeed, it is widely felt to be just silly to suppose that there was such a second. Our use of the word ‘old’ is conceived as too vague to single one out. On such grounds, the principle of bivalence has been rejected for vague languages. To reject bivalence is to reject classical logic or semantics.

At some times, it was unclear whether Rembrandt was old. He was neither clearly old nor clearly not old. The unclarity resulted from vagueness in the statement that Rembrandt was old. We can even use such examples to define the notion of vagueness. An expression or concept is vague if and only if it can result in unclarity of the kind just exemplified. Such a definition does not pretend to display the underlying nature of the phenomenon. In particular, it does not specify whether the unclarity results from the failure of the statement to be true or false, or simply from our inability to find out which. The definition is neutral on such points of theory. Just as we might agree to define the term ‘light’, or ‘poetry’, by examples, in order not to talk past each other when disagreeing about the nature of light, or poetry, so we can agree to define the term ‘vagueness’ by examples, in order not to talk past each other when disagreeing about the nature of vagueness.

The phenomenon of vagueness is broad. Most challenges to classical logic or semantics depend on special features of a subject matter: the future, the infinite, the quantum mechanical. For all such a challenge implies, classical logic and semantics apply to statements about other subject matters. Vagueness, in contrast, presents a ubiquitous challenge. It is hard to make a perfectly precise statement about anything. If classical logic and semantics apply only to perfectly precise languages, then they apply to no language that we can speak.

The phenomenon of vagueness is deep as well as broad. It would be shallow if it could be adequately described in precise terms. That is not generally possible. The difficulties presented by the question ‘When did Rembrandt become old?’ are also presented by the question ‘When did Rembrandt become clearly old?’. At some times, it was unclear whether it was unclear whether Rembrandt was old. The limits of vagueness are themselves vague. The same difficulties are presented by the question ‘When did Rembrandt become clearly clearly old?’; the point reiterates *ad infinitum.* This is the phenomenon of higher-order vagueness. It means that the meta-language in which we describe the vagueness of a vague language will itself be vague.

The use of non-classical systems of logic or semantics has been advocated for vague languages. New and increasingly complex systems continue to be invented. What none has so far given is a satisfying account of higher-order vagueness. In more or less subtle ways, the meta-language is treated as though it were precise. For example, classical logic is said to be invalid for vague languages, and is then used in the meta-language. Such proposals underestimate the depth of the problem.

The problem is not solved by the pessimistic idea that no system of logic or semantics, classical or non-classical, is adequate for a vague language. That idea still permits one to ask for perspicuous descriptions of vagueness in particular cases. No one has given a satisfying and perspicuous description of higher-order vagueness without use of classical logic. Of course, the nature of vagueness might be to defy perspicuous description, but that counsel of despair should prevail only if there is good evidence that it does not overestimate the depth of the problem.

The thesis of this book is that vagueness is an epistemic phenomenon. As such, it constitutes no objection to classical logic or semantics. In cases of unclarity, statements remain true or false, but speakers of the language have no way of knowing which. Higher-order vagueness consists in ignorance about ignorance.

At first sight, the epistemic view of vagueness is incredible. We may think that we cannot conceive how a vague statement could be true or false in an unclear case. For when we conceive that something is so, we tend to imagine finding out that it is so. We are uneasy with a fact on which we cannot attain such a first-personal perspective. We have no idea how we ever could have found out that the vague statement is true, or that it is false, in an unclear case; we are consequently unable to imagine finding out that it is true, or that it is false; we fallaciously conclude that it is inconceivable that it is true, and inconceivable that it is false. Such fallacies of the imagination must be laid aside before the epistemic view can be adequately assessed.

Most work on the problem of vagueness assumes that the epistemic view is false, without seriously arguing the point. If the epistemic view is true, that work is fundamentally mistaken. Even if the epistemic view is false, that work is ungrounded until cogent arguments against the view have been found. The assessment of the epistemic view is therefore one of the main tasks facing the study of vagueness. This book contributes to that task.

The assessment of the epistemic view has ramifications far beyond the study of vagueness. As already noted, classical logic and semantics are at stake. More generally, the epistemic view implies a form of realism, that even the truth about the boundaries of our concepts can outrun our capacity to know it. To deny the epistemic view of vagueness is therefore to impose limits on realism; to assert it is to endorse realism in a thoroughgoing form.

The first part of the book is historical. It traces the slow and intermittent recognition of vagueness as a distinct and problematic phenomenon, up to the origins of the theories of vagueness that have been popular over the last two decades. This part is also critical. It argues that none of the extant non-epistemic theories of vagueness is adequate. Not only do they abandon classical logic or semantics for alternatives of doubtful coherence; those sacrifices are not rewarded by adequate insight into the nature of vagueness. The second part of the book is constructive. It develops and applies an epistemic view of vagueness, finds the standard objections to it fallacious, and concludes that the epistemic view provides the best explanation of the phenomenon of vagueness.

The Greeks introduced the problem of vagueness into philosophy, in the guise of the original sorites paradox: if the removal of one grain from a heap always leaves a heap, then the successive removal of every grain still leaves a heap. Chapter 1 sketches the history of this paradox and its variants from their invention to the nineteenth century. Stoic logicians are interpreted as taking an epistemic view of sorites paradoxes.

What makes a sorites paradox paradoxical is the vagueness of its central term, e.g. ‘heap’. Historically, however, such paradoxes were identified by their form. Vagueness as such became a topic of philosophical discussion only at the start of the twentieth century, when it presented an obstacle to the ideal of a logically perfect language associated with the development of modern logic. Only with difficulty was the phenomenon of unclear boundaries separated from other phenomena, such as lack of specificity, to which the term ‘vagueness’ is also applied in everyday usage. Chapter 2 discusses three stages in the emergence of the philosophical concept of vagueness, in the work of Frege, Peirce and Russell.

As philosophical attention turned to ordinary language, vagueness acquired a more positive image. It was seen no longer as a deviation from an ideal norm, but as the real norm itself. As such, it was described by Black, Wittgenstein and Waismann. Their work is discussed in Chapter 3.

Formal treatments of vagueness have become common only in the last few decades. One main approach relies on many-valued logic, which replaces the dichotomy of truth and falsity by a manyfold classification. Chapter 4 follows the development of its application to the problem of vagueness, from the use of three-valued logic to the growth of 'fuzzy logic' and other logics based on infinitely many values, and then of more sophisticated accounts appealing to a qualitative conception of degrees of truth. These views are criticized on several grounds. None has adequately treated higher-order vagueness; degrees of truth are not connected with vagueness in the requisite way; the generalization from two-valued logic to many-valued logic has highly counter-intuitive consequences when applied to natural languages.

A technically subtler approach to vagueness is supervaluationism, with which Chapter 5 is concerned. It preserves almost all of classical logic, at the expense of classical semantics, by giving a non-standard account of truth. It also treats higher-order vagueness in a promising way. However, it is argued that the treatment of higher-order vagueness undermines the non-standard account of truth, making supervaluationism as a whole unmotivated.

On a more pessimistic view, vagueness is a form of incoherence. If this view is taken globally, Chapter 6 suggests, all rational discourse is subverted, for vagueness is ubiquitous. However, local forms of nihilism might be coherent. They have been defended in the special case of concepts used to describe perceptual appearances, on the grounds that such concepts cannot differentiate between perceptually indiscriminable items, yet perceptually discriminable items can be linked by a sorites series of which each member is perceptually indiscriminable from its neighbours. However, careful attention to the structure of the relevant concepts shows that the paradoxical arguments are unsound. In particular, they falsely assume that appearances are just what they appear to be.

Chapter 7 defends the epistemic view of vagueness. First, it argues that it is incoherent to deny the principle of bivalence for vague statements in unclear cases. It then questions our ability to think through coherently the consequences of a non-epistemic view of vagueness. Obvious objections to the epistemic view are analysed and shown to be fallacious. A picture of linguistic understanding is sketched, on which we can know that a word has a given meaning without knowing what the boundaries of that meaning are in conceptual space.

Chapter 8 develops the epistemological background to the epistemic view. It gives independent justification for principles about knowledge on which the ignorance postulated by the view was only to be expected as a special case of a much wider phenomenon, inexact knowledge. Nevertheless, the case is special, for the source of the inexactness is distinctive in being conceptual. Higher-order vagueness has a central place in this account, for a central feature of inexact knowledge is that one can know something without being in a position to know that one knows it; when the inexactness takes the form of vagueness, this becomes unclarity about unclarity. The epistemology of inexact knowledge is then used to analyse in greater depth the phenomena of indiscriminability to which the nihilist appeals.

It is controversial whether in any sense the world itself, as opposed to our representations of it, can be vague. Chapter 9 examines the issue. It argues that the epistemic view permits objects to be vague in a modest sense, for the impossibility of knowing their boundaries may be independent of the way in which the objects are represented.

The Appendix identifies the formal system uniquely appropriate for the logic of clarity and unclarity on the epistemic view of vagueness.

Most of the book keeps technicalities to a minimum. The gain in intelligibility will, it is hoped, outweigh the loss in rigour. There is also a philosophical reason for minimizing technicality. The emphasis on formal systems has encouraged the illusion that vagueness can be studied in a precise meta-language. It has therefore caused the significance of higher-order vagueness to be underestimated. Indeed, to use a supposedly precise meta-language in studying vague terms is to use a language into which, by hypothesis, they cannot be translated. Since vague terms are meaningful, this is an expressive limitation on the meta-language. It is not an innocent one. The argument in Chapter 7 for the incoherence of denials of bivalence in unclear cases can be stated only in a language into which the relevant vague terms can be translated. To deny bivalence is in the end to treat vague utterances as though they said nothing. Vagueness can be understood only from within.

Chapter 1

The early history of sorites paradoxes

1.1 THE FIRST SORITES[1]

The logician Eubulides of Miletus, a contemporary of Aristotle, was famous for seven puzzles. One was the Liar: if a man says that he is lying, is he telling the truth? Another was the Hooded Man: how can you know your brother when you do not know that hooded man, who is in fact your brother? The Electra turned on the delusion of Orestes in his madness, who took his sister Electra for a Fury. The Elusive Man will appear later. There was also the Horned Man: since you still have what you have not lost, and you have not lost horns, you still have them (hence the horns of a dilemma). The remaining puzzles were the Bald Man and (accompanying five men and one woman) the Heap. In antiquity they were usually formulated as series of questions.[2]

Does one grain of wheat make a heap? Do two grains of wheat make a heap? Do three grains of wheat make a heap? . . . Do ten thousand grains of wheat make a heap? It is to be understood that the grains are properly piled up, and that a heap must contain reasonably many grains. If you admit that one grain does not make a heap, and are unwilling to make a fuss about the addition of any single grain, you are eventually forced to admit that ten thousand grains do not make a heap.

Is a man with one hair on his head bald? Is a man with two hairs on his head bald? Is a man with three hairs on his head bald? . . . Is a man with ten thousand hairs on his head bald? It is to be understood that the hairs are properly distributed, and that a man with reasonably few hairs is bald. If you admit that a man with one hair is bald, and are unwilling to make a fuss about the addition of any single hair, you are eventually forced to admit that a man with ten thousand hairs is bald.

The standard ancient terms for the Heap and the Bald Man were '*sorites*' and '*phalakros*' respectively. The Greek adjective '*sorites*' comes from the noun '*soros*', for 'heap', and means literally *heaper*, as in 'the heaper paradox'. Perhaps Eubulides coined the word himself. The primary reference was to the content of the puzzle, although there may also have been a secondary allusion to the heaping up of questions in its form. The term was extended to similar puzzles, such as the Bald Man. They were also known as little-by-little arguments.[3]

So far as is known, Eubulides invented the Heap and the Bald Man himself. The Heap may have been inspired by a different puzzle, the Millet Seed, propounded by Zeno of Elea a century earlier. If the fall of a seed to the ground were completely silent, so would be the fall of a bushel of seed, which it is not; thus each seed must make a noise when it falls to the ground. That puzzle can certainly be adapted to sorites form. Does one seed make a noise when it falls to the ground? Do two seeds? Three? . . . However, Zeno seems to have based his puzzle on something much more specific: a principle that the noise made when some grains fall to the ground is proportional to their weight.[4]Eubulides may well have been the first to focus on the general difficulty that sorites questioning presents; he is unlikely to have overlooked the common form of the Heap and the Bald Man, however much their difference in content caused them to be listed as distinct puzzles.

It is not known what Eubulides used sorites puzzles for – fun, troublemaking or some graver purpose. Many philosophical doctrines have been suggested as the target he intended them to destroy: the coherence of empirical concepts (such as 'bald' and 'heap'), the law of non-contradiction ('Not both *P* and not *P*'), the law of excluded middle ('Either *P* or not *P*'), pluralism (the existence of more than one thing), Aristotle's theory of infinity (as potential rather than actual), Aristotle's theory of the mean (as the place of virtue between vicious extremes). The evidence gives little support to any of these suggestions. Eubulides is indeed said to have attacked Aristotle, but in slanderous terms; the sources do not connect the dispute with any of the puzzles. Aristotle betrays no clear awareness of sorites reasoning in any of his extant works. Some later commentators did consider its use against Aristotle's theory of the mean, but without suggesting that either Eubulides or Aristotle had done so. Eubulides' interests were described as purely logical; if he had a specific target in mind, it is likely to have been a logical one.[5] But if he had no specific target in mind, it does not follow that his interest in sorites puzzles was intellectually frivolous, for one can treat any absurdity apparently licensed by accepted standards of argument as a serious challenge to those standards. A puzzle threatens the whole system of thought within which it is formulated, and for that very reason refutes no element of it in particular until a positive diagnosis is given, fixing the source of the trouble. There is no evidence that Eubulides had such a diagnosis.

The sorites puzzles can be traced conjecturally forward from Eubulides in the mid fourth century BCE through a chain of teacher–pupil links. One pupil of Eubulides was Apollonius Cronus; one of his pupils was Diodorus Cronus. David Sedley argues that it was Diodorus who gave Eubulides' puzzles wide circulation and intellectual status. According to Sedley, Diodorus inherited from Eubulides 'his sophistical leanings, his flamboyancy and his love of showmanship'; he could present a puzzle so that it not only caught the imagination but emerged as a serious challenge to philosophical theory.[6] Diodorus taught dialectic, 'the science of discoursing correctly on arguments in question and answer form', the standard medium of philosophical argument and that in which the sorites puzzles were formulated.

No explicit mention of the sorites by Diodorus survives. However, he almost certainly discussed it; his intellectual ancestors and descendants did, and it is just the kind of puzzle he liked. Moreover, he propounded an argument about motion with very strong affinities to the Heap. Its details as they survive are so unconvincing that one must hope them to have been garbled. It begins with the promising general principle that a predicate which can apply 'absolutely' can also apply 'by predominance'; for example, a man can be grey-headed because all the hairs on his head are grey, but he can also be grey-headed because a majority of them are (the Bald Man involves a different form of aging). The idea is then to suppose that a body of three particles is moving by predominance, and to add stationary particles until a body of ten thousand particles would also be moving by predominance when only two particles are. Heaps are indeed mentioned in the course of the argument.[7] Although a sorites puzzle could be constructed out of such material, this is not it; the words 'a majority' would need to be replaced by something like ‘most’ in the definition of ‘predominance’. The puzzle would then turn on ‘most’ rather than ‘moving’. It is tempting to connect the sorites with some of Diodorus’ other interests too. For example, he argued that the present truth of a past tense statement such as ‘It has moved’ or ‘Helen had three husbands’ does not require the past truth of a present tense statement such as ‘It is moving’ or ‘Helen has three husbands’.[8] Might the question ‘When?’ be similarly out of place with respect to the creation or destruction of a heap? Again, Diodorus studied the minima of perceptibility; some later sorites paradoxes depend on the fact that imperceptible differences can add up to a perceptible one. Unfortunately, there is no evidence to verify these speculative connections.

Diodorus is unreliably reported to have died of shame when he failed to solve a puzzle set him by another philosopher in front of Ptolemy Soter, King of Egypt. After him, sorites puzzles became standard weapons in disputes between two rival schools of philosophy, the Stoa and the Academy. He influenced both. In about 300 his ex-pupil Zeno of Citium began to teach in a public hall in Athens, the Stoa Poikile, from which his pupils were known as Stoics. In about 273 Arcesilaus became head of Plato’s Academy and achieved a remarkable *coup* by converting it to scepticism; Arcesilaus was influenced by Diodorus’ logic.[9] The sceptics of the Academy then attacked the Stoic ‘dogmatists’. Their chief weapons included sorites arguments, against which they took the Stoics to have no non-sceptical defence. A sceptic does not feel obliged to answer any of the sorites questions; he can simply plead ignorance. If a Stoic is obliged to answer each question ‘Yes’ or ‘No’, he will find himself in an embarrassing position.

The Stoic theory of knowledge was an obvious focus for sceptical attack, and sorites reasoning a particularly apt means. The central concept of Stoic epistemology was that of a cognitive impression, something like a Cartesian clear and distinct idea. Cognitive impressions have a propositional content, and are essentially such as to derive from and represent real objects with complete accuracy. They are the foundation of the Stoic account of the possibility of knowledge. The sceptics proceeded to construct sorites series from cognitive to non-cognitive impressions, replacing each impression by a virtually indistinguishable one, and took themselves to have undermined Stoic claims to knowledge.[10]

The Stoics needed a defence against sceptical sorites attacks. But they must have discussed sorites arguments prior to the institution of scepticism in the Academy, for Zeno emphasized the development of techniques for dealing with just such puzzles as part of a training in dialectical techniques. He was succeeded as head of the Stoa by Cleanthes in 262, and Cleanthes by Chrysippus (*c*.280–*c*.206) in about 232. It was Chrysippus who was responsible for the systematic construction of Stoic logical theory. In antiquity he was reputed the greatest logician of all (Aristotle being the greatest scientist, and Plato the greatest philosopher).[11] He also devised strategies to handle attacks by the sceptics, and in particular their use of sorites arguments. Unfortunately, the two volumes (i.e. scrolls of papyrus) he is known to have written *On the Little-by-Little Argument* do not survive; nor do his three volumes *On Sorites Arguments Against Words*.[12]There are only fragments and reports. What follows is a speculative reconstruction.

1.2 CHRYSIPPAN SILENCE

The Stoics firmly accepted the principle of bivalence: every proposition is either true or false. Unlike Aristotle, they did not reject it for future contingencies; it is true or false that there will be a sea-fight tomorrow. According to Cicero, Chrysippus 'strains every nerve to persuade us that every *axioma* [proposition] is either true or false'.[13] Thus for every proposition P there is one right answer to the question 'P?'; it is 'Yes' if P is true and 'No' if P is false. Consequently, for every sequence of propositions $P_1, \ldots, P_n$ there is one sequence of right answers to the questions 'P_1?', . . . , 'P_n?', each member of which is either 'Yes' or 'No'.

In a sorites sequence, the right answers to the first and last questions are obvious and opposite. The Stoics used 'Are i few?' as the schematic form of the ith question; it will be convenient to follow their example.[14] Thus the right answers to the first and last questions are 'Yes' and 'No' respectively. This fits the Bald Man rather than the Heap, but the difference is not significant. Now the usual sorites sequences are monotonic, in the sense that a question rightly answerable 'No' never comes between two questions rightly answerable 'Yes', nor vice versa. If i are few then, *a fortiori*, fewer than i are few; if i are not few then, *a fortiori*, more than i are not few. Thus the sequence of right answers to the questions 'Is one few?', . . . , 'Are ten thousand few?' consists of 'Yes' a certain number of times followed by 'No' a certain number of times. In particular, there is a last question 'Are i few?' rightly answerable 'Yes', immediately followed by a first question 'Are $i + 1$ few?' rightly answerable 'No'. i are few and $i + 1$ are not few; i is a sharp cut-off point for fewness.

The argument from bivalence to the existence of a sharp cut-off point assumes that the sentences 'One is few', . . . , 'Ten thousand are few' do express propositions. The Stoics themselves distinguished the proposition asserted from the sentence by means of which it is asserted. However, someone who utters 'i are few' with the sense 'A man with i hairs on his head is bald' does assert something, which on the Stoic view requires the sentence to express a proposition. The assumption gave no escape from the argument to a sharp cut-off point.

There is independent evidence that the Stoics accepted sharp cut-off points.[15] First and most tenuous, in Cicero's account Chrysippus compares himself to a clever charioteer who pulls up his horses before he comes to a precipice; what is the sharp drop if not from truth to falsity? Second, in other cases which look susceptible to sorites reasoning the Stoics insisted on sharp cut-off points. For example, they denied that there are degrees of virtue, holding that one is either vicious or perfectly virtuous. An analogy was drawn with a drowning man as he rises to the surface; he is coming closer to not drowning but he is drowning to no less a degree until he breaks the surface, when he is suddenly not drowning at all. Third, in rebutting the sorites argument against cognitive impressions, Chrysippus dealt explicitly with the case 'when the last cognitive impression lies next to the first non-cognitive one'; cognitiveness has a sharp cut-off point. The Stoics were prepared to apply bivalence to sorites reasoning and swallow the consequences.

For the Stoics, there are sharp cut-off points. The difficulty in answering the sorites questions must come not from the non-existence of right answers but from our ignorance of what they are. The sorites is a puzzle in epistemology. This book is a defence of that Stoic view. The immediate need, however, is not to defend the view but to explore its consequences.

One might answer the questions 'Is one few?', . . . , 'Are i few?' 'Yes' and the questions 'Are $i + 1$ few?', . . . , 'Are ten thousand few?' 'No': but one would be guessing. No one has such knowlege of cut-off points; no one knows both that i are few and that $i + 1$ are not few. Such a pattern of answers is forbidden by the principle that one should give an answer only if one knows it to be correct. If they were to respect that principle, the Stoics needed an alternative pattern of answers.

The problem concerns what one should believe, not just what one should say. If one believes that i are few and that $i + 1$ are not few, one violates the principle that one should believe only what one knows to be correct. The Stoics were committed to this principle in a sense, but the elucidation of that sense requires a little more background in Stoic epistemology.[16]

The Stoics made a threefold distinction between mere opinion (*doxa*), cognition (*katalepsis*) and scientific knowledge (*episteme*). The ordinary man cognizes by assenting to cognitive impressions. He believes their propositional content, as in normal perception. Cognitive impressions tend to cause us to assent to them. We are strongly but not irresistibly inclined to take what is clearly the case to be the case. However, the ordinary man tends to assent indiscriminately to all impressions. When he does not realize that conditions are abnormal, he may land himself with a mere opinion by assenting to a non-cognitive impression. The wise man, the Stoic ideal, does not err in that way (or any other). He has trained himself to assent only to cognitive impressions. This is possible because there is always some qualitative difference, however slight, between a cognitive and a non-cognitive impression, given the Stoic principle that no two entities are absolutely indiscernible.[17] In case of doubt, the wise man suspends judgement. Where certainty is not to be had, he may act on plausible assumptions, but without believing them; what he believes is just that they are plausible.[18] Since cognitive impressions cannot be false, the wise man has no false beliefs. He is not omniscient, but he is infallible.[19] His beliefs, and only his, are sure to withstand attempted refutations. For example, he will not be tricked by the sorites questioner into denying the proposition that ten thousand grains make a heap, to which he had previously assented on the basis of a cognitive impression. Only he has scientific knowledge.

Understandably, the Stoics were not sure that there had been any wise men. Perhaps Socrates was one, but the number was very small, and the Stoics did not themselves claim to be wise. Thus they could not claim to say or believe only what they scientifically knew. However, they did aspire to wisdom. In particular, they sought to train themselves to assent only to cognitive impressions, to what is clear. The term ‘knowledge’, unqualified by ‘scientific’, will be used for assent to what is clear, i.e. to truths guaranteed as such by cognitive impressions; it covers both cognition and scientific knowledge. In circumstances for which they had trained, ordinary Stoics might indeed say and believe only what they knew in this looser sense. Moreover, they could not hope to deal successfully with the sorites interrogation if they assented to non-cognitive impressions in the course of it. Thus any good strategy for responding to the interrogation would involve one in saying and believing only what one knows. This principle will help to explain the Stoic treatment of the puzzle.

Chrysippus recommended that at some point in the sorites interrogation one should fall silent and withhold assent.[20] The wise man would and the ordinary Stoic should suspend judgement, making no statement and forming no belief either way, thereby avoiding error in both thought and speech.

If the source of the puzzle is just that one does not know whether ‘Yes’ or ‘No’ is the right answer to some of the questions, it turns out to be on a level with other matters of which one is simply ignorant. For example, the Stoics readily admitted that they did not know the right answer to the question ‘Is the number of stars even?’. If no more is involved, the Stoic could confidently face a sorites interrogation armed only with the three possible answers ‘Yes’, ‘No’ and ‘I don’t know’. If he knew *i* to be few he would answer ‘Yes’; if he knew *i* not to be few he would answer ‘No’; in every other case he would answer ‘I don’t know’. Why should such an honest admission of ignorance not completely dissolve the puzzle?

The Stoics did not classify the sorites interrogation merely as a list of questions some of whose answers were unknown; they classified it as a sophism. The question about the stars makes it very easy to say only what one knows to be true. The sorites makes it very hard not to say what one knows to be false. At first sight, the epistemic approach seems to lose the difficulty of the puzzle. What follows is an attempt to think through the epistemology of sorites series more carefully on Stoic lines. The results are subtler than one might expect, and help to explain some otherwise puzzling features of the Stoic strategy.

Recall that Chrysippus did not say that one should admit ignorance; he said that one should fall silent. Under interrogation, saying ‘I don’t know’ is quite a different policy from saying nothing. In the former case but not the latter one denies knowledge. The Stoic is supposed not to make a statement unless he knows it to be correct. Now to say ‘I don’t know’ in answer to the question ‘Are *i* few?’ is in effect to make the statement ‘I neither know that *i* are few nor know that *i* are not few’, just as ‘Yes’ is tantamount to the statement ‘*i* are few’ and ‘No’ to the statement ‘*i* are not few’. Thus the Stoic is supposed to answer ‘I don’t know’ only if he *knows* that he neither knows that *i* are few nor knows that *i* are not few. The ‘Yes’/‘No’/‘Don’t know’ strategy requires the Stoic to answer ‘I don’t know’ whenever he does not know. The strategy is therefore available on Stoic terms only if for each *i*:

(1) If one neither knows that *i* are few nor knows that *i* are not few, then one knows that one neither knows that *i* are few nor knows that *i* are not few.

Knowledge was defined above as assent to what is clear. Thus what is known is clear. Conversely, since cognitive impressions strongly incline us to assent, what is clear tends to be known. The ‘Yes’/‘No’/‘Don’t know’ strategy is therefore very close to the following: if *i* are clearly few, say ‘Yes’; if *i* are clearly not few, say ‘No’; otherwise, say ‘Unclear’. It will be simplest to begin with the latter strategy. Once its flaws have emerged, the former can be reconsidered.

The ‘Yes’/‘No’/‘Unclear’ strategy is available on Stoic terms only if the third answer is clearly correct when neither the first nor the second is. (1) corresponds to:

(2) If *i* are neither clearly few nor clearly not few, then *i* are clearly neither clearly few nor clearly not few.

This is equivalent to the conjunction of two simpler principles:[21]

(3a) If i are not clearly few, then i are clearly few.
(3b) If i are not clearly not few, then i are clearly not clearly not few.

(3a) and (3b) are two instances of what is called the S5 principle for clarity: if it is not clear that P, then it is clear that it is not clear that P. Thus the 'Yes'/'No'/'Unclear' strategy is available only if clarity satisfies the S5 principle.

On a Stoic view, clarity cannot in general satisfy the S5 principle. For example, not knowing that the light is abnormal, I assent to the non-cognitive impression that x is yellow, when in fact x is white. Thus x is not clearly yellow, but it is not clearly not clearly yellow.[22] If it were, I should have a cognitive impression to the effect that x is not clearly yellow; such an impression would strongly incline me to judge that x is not clearly yellow. But I have no such inclination at all; it does not occur to me that the situation is other than a normal case of seeing something to be yellow. My problems come from glib assent to a non-cognitive impression, not from resistance to a cognitive one. Of course, the Stoic view is that, if I had been more attentive, I could have avoided the mistake. My attentiveness might have enabled me to have a cognitive impression to the effect that my perceptual impression was non-cognitive, but it does not follow that I have such a second-order cognitive impression in my actual state. I have no such impression.

The S5 principle seems equally vulnerable in the case of (3a) and (3b). Just as fewness is sorites-susceptible, so is clear fewness. One is clearly few; ten thousand are not clearly few (for they are not few). By Stoic logic, there is a sharp cut-off point for clear fewness: for some number i, $i - 1$ are clearly few and i are not clearly few. One is in no better a position to say what that number is than to locate the cut-off point for fewness itself. The cognitive impression that $i - 1$ are few is too relevantly similar to the non-cognitive impression that i are few for one to judge reliably that the latter is indeed a non-cognitive impression. One cannot discriminate so finely. Although i are not clearly few, they are not clearly not clearly few. Some instance of (3a) is false. By a parallel argument, so is some instance of (3b). Thus (2), which entails them, is false too. The 'Yes'/'No'/'Unclear' strategy fails on Stoic terms, for at some points in the series none of those three answers is clearly correct, so none is known to be correct. One cannot recognize the last of the clear cases.

Jonathan Barnes has argued to the contrary that 'it seems relatively easy to show that *if* there is a last clear case then we can *recognise* it as such'. Transposed to present notation, his reasoning runs:

> Consider any relevant case: are *i* clearly few or not? One may answer 'No', 'Yes' or 'I don't know'; but if one doesn't know whether *i* are clearly few, then *i* are not clearly clearly few and hence not clearly few. Hence one can always answer 'No' or 'Yes' to the question 'Are *i* clearly few?': hence the last clear case, if it exists, is recognizable.[23]

The reasoning is perfectly general. If it were valid, it would show that clarity satisfies the S5 principle everywhere. In particular, it would apply to my mistake about the colour of *x*. But then the fallacy is clear. If, in my glib state, I am asked whether *x* is clearly yellow, I shall answer 'Yes', but I shall be wrong.[24]

One might try to revive Barnes's argument by adding an extra premise. It was noted above that the Stoic can expect to survive the sorites interrogation intact only if he has trained himself not to assent to any relevant non-cognitive impression. One might therefore add the premise that the Stoic's answers are clearly correct. This excludes counterexamples of the foregoing kind. If one answers 'Yes' to the question 'Are *i* clearly few?', then *i* are clearly clearly few; if one answers 'No', they are clearly not clearly few. However, it still needs to be shown that if neither of these answers is available then one clearly doesn't know whether *i* are clearly few, for otherwise the answer 'I don't know' is unavailable too.

One might hope to make the answer 'I don't know' available just by suspending judgement as to whether *i* are clearly few whenever one did not say 'Yes' or 'No'. For knowledge requires assent; if one assents neither to the proposition that *i* are clearly few nor to the proposition that they are not clearly few, one knows neither proposition. Moreover, one can suspend judgement clearly, to make one's ignorance clear. One then clearly fails to know that *i* are clearly few. However, what needs to be shown at this point in the argument is that *i* are not clearly clearly few. But this does not follow, for one's failure to know is guaranteed only by one's failure to assent. If one can refuse assent to what is clear, one might refuse assent to the proposition that *i* are clearly few even when *i* are clearly clearly few. Thus the argument seems to require clarity to be sufficient as well as necessary for the assent of a well-trained Stoic. This is tantamount to the assumption that such a person can discriminate with perfect accuracy between cognitive and non-cognitive impressions. But that is in effect what the argument was supposed to show.

One cannot reliably answer the question 'Are *i* clearly few?' just by following the advice: if you hesitate to say 'Yes', say 'No'. If that policy worked, one would unhesitatingly judge that *i* were clearly few if and only if *i* were clearly few; whatever one thought was right would be right. But clarity is an independently desirable epistemic feature about whose application one can be wrong as well as right. Unless one is reasonably cautious, one will often unhesitatingly judge that *i* are clearly few when they are not in fact clearly few (or even few). On the other hand, if one is reasonably cautious, one will often hesitate over what turns out to be genuinely clear; one suspects a hidden catch and finds there is none. The story is told of a mathematician lecturing who began a sentence 'It is clear that . . .', was seized by sudden doubt, spent several minutes in agonized thought, and then resumed 'It *is* clear that . . .'. The point of the story is that he may have been right.

There is no universal algorithm for detecting clarity. Chrysippus seems to have held that not even the wise man can discriminate with perfect accuracy between cognitive and non-cognitive impressions. Sextus Empiricus writes:

> For since in the Sorites the last cognitive impression is adjacent to the first non-cognitive impression and virtually indistinguishable from it, the school of Chrysippus say that in the case of impressions which differ so slightly the wise man will stop and become quiescent, while in the cases where a more substantial difference strikes him he will assent to one of the impressions as true.[25]

Not even the wise man can locate the last clear case with perfect accuracy. Practice has improved his powers of discrimination but not made them perfect. In order to avoid the risk of assent to a non-cognitive impression, he refrains from assenting to impressions that he cannot discriminate from non-cognitive ones. He will therefore sometimes refrain from assenting to what is in fact a cognitive impression. When he refrains from assenting to an impression, he does not always classify it as unclear, for if he did he would sometimes be mistaken. He does not say ‘Unclear’; he falls silent. The ordinary Stoic, who takes him as a model, should do the same.

If one answers the simple sorites question ‘Are *i* few?’ with ‘Yes’ whenever that answer is clearly correct, on Stoic assumptions one stops answering ‘Yes’ either at the point when it ceases to be clearly correct or at some later stage. In the former case one has located the cut-off point for clarity with perfect accuracy; in the latter one has violated the constraint that all one’s answers should be clearly correct. Since one cannot reliably locate the cut-off point for clarity with perfect accuracy, one will reliably satisfy the constraint only if one stops answering ‘Yes’ before it has ceased to be the clearly correct answer. One must undershoot in order to avoid the risk of overshooting.

The Stoic will fall silent *before* the end of the clear cases. Chrysippus is reported as advising just that.[26] Barnes’s claim that the end of the clear cases can easily be shown to be recognizable as such makes him suggest that Chrysippus may have been misreported, and merely suggested that one should stop *at* the end of the clear cases. What has emerged is that there is no reason to reject the report, for the advice it attributes to Chrysippus is good advice on Stoic terms.

There remains a trivial sense in which the ‘Yes’/‘No’/‘Don’t know’ strategy is feasible, unlike the ‘Yes’/‘No’/‘Unclear’ strategy. For if one’s refusal of assent is very clear, one can recognize it and thereby come to know that one does not know. A trained Stoic may satisfy (1) throughout a sorites series, if he assents to no relevant non-cognitive impression. However, ‘I don’t know’ now in effect reports his refusal to assent, a psychological episode, not the state of the evidence available to him; what he does not know may nevertheless be clear. Such a report would be of little interest to the questioner. Silence remains intelligible.

The moral that one should stop before the end of the clear cases can be generalized. One would like to obey two injunctions:

(4a) If ‘Yes’ is a good answer, say ‘Yes’.
(4b) If ‘Yes’ is not a good answer, do not say ‘Yes’.

The goodness of an answer is some truth-related property of it, and does not simply consist in its being given. There is play between the antecedents and consequents of (4a) and (4b); in an imperfect world they will sometimes come apart. In such a case, one either fails to say 'Yes' when 'Yes' is a good answer, violating (4a), or says 'Yes' when 'Yes' is not a good answer, violating (4b). If one regards violations of (4a) and (4b) as equally serious, one may simply aim to say 'Yes' when and only when it is a good answer. Other things being equal, one's misses are as likely to fall on one side of the target as on the other, and no matter. But one might regard a violation of (4b) as worse than a violation of (4a); given the choice, one would rather err by omission, not saying 'Yes' when it is a good answer, than by commission, saying 'Yes' when it is not a good answer. For example, one may prefer failing to make true or warranted statements to making false or unwarranted ones. In that case, one will follow a policy of saying nothing when in doubt. One decreases the risk of more serious violations by increasing the risk of less serious ones. At the limit, the price of never violating (4b) is sometimes violating (4a). That is the choice the Stoic made in falling silent before the end of the clear cases; here clarity is goodness. It was worse to say 'Yes' in an unclear case than not to say it in a clear one. Those who take the opposite view should fall silent after the end of the clear cases.

The Chrysippan strategy results from two levels of precaution. At the first level, goodness in (4a) and (4b) is simply truth. The Stoics were not alone in holding it to be worse to give a false answer than to fail to give a true one. For truth, (4a) rather than (4b) is to be violated. This preference motivates the constraint that one should give an answer only if it is clear. But then clarity takes on a life of its own as a cognitive end, and again the Stoic takes the cautious option. (4a) rather than (4b) is to be violated for clarity too.

The Chrysippan strategy is incomplete if it gives no clue as to where amongst the clear cases one should fall silent. If the advice is to fall silent a little before the end of the clear cases, it is very vague, and also presupposes that the respondent can predict the future course of the questioning – an impossible task in the case of non-numerical sorites, such as those about cognitive impressions, and others considered below. A sceptic might suspend judgement until the questioning was over, but that was not what Chrysippus recommended. Perhaps one is supposed to be silent for all the cases one cannot discriminate from unclear cases. For reasons of a kind already given, there is no algorithm for discriminability from unclear cases; one will sometimes be wrong about it. However, if goodness is taken to be discriminability from unclear cases, one might regard violation of (4b) as no worse than violation of (4a) at this third level. That would bring the regress to an end.

1.3 SORITES ARGUMENTS AND STOIC LOGIC

A paradox may be defined as an apparently valid argument with apparently true premises and an apparently false conclusion. One often speaks of a sorites paradox, and there was mention above of sorites arguments. Yet the Heap and the Bald Man have been presented, as they usually were in antiquity, as series of questions, not as arguments with premises and conclusions. According to Barnes, 'we can – and the ancients did – see a logical structure behind that dialectical façade'.[27]

Consider the following argument, with premises above the line and conclusion below:

1 is few
If 1 is few then 2 are few
If 2 are few then 3 are few
.
.
.
If 9,999 are few then 10,000 are few
___
10,000 are few

The argument appears to be valid; if its premises are true, its conclusion will be true too. The relevant rule of inference is *modus ponens*, which allows one to argue from '*P*' and 'If *P* then *Q*' to '*Q*'; it appears impossible for its premises to be true and its conclusion false. By *modus ponens*, '1 is few' and 'If 1 is few then 2 are few' entail '2 are few'. In the same way, '2 are few' and 'If 2 are few then 3 are few' entail '3 are few'. After 9,999 applications of *modus ponens*, one finally reaches the conclusion '10,000 are few'. The premise '1 is few' is apparently true and the conclusion

‘10,000 are few’ apparently false. The gradualness of the sorites series makes each of the conditional premises appear true. Thus the apparently valid argument has apparently true premises and an apparently false conclusion. At least one of these appearances is misleading, for the conclusion cannot be both true and false.

The argument is valid by the standards of orthodox modern logic; it cannot have true premises and a false conclusion. It is also valid by the standards of Stoic logic.[28] Two logical principles are at stake. One is *modus ponens*; it was the first indemonstrable (basic) form of argument in Stoic logic: ‘If the first, then the second; but the first; therefore the second’. The other is the principle, sometimes known as Cut, that valid arguments can be chained together: for example, the valid argument from ‘1 is few’ and ‘If 1 is few then 2 are few’ to ‘2 are few’ can be chained together with the valid argument from ‘2 are few and ‘If 2 are few then 3 are few’ to ‘3 are few’, giving a valid argument from ‘1 is few’, ‘If 1 is few then 2 are few’ and ‘If 2 are few then 3 are few’ to ‘3 are few’. The third Stoic ground-rule for the analysis of complex arguments is the relevant form of Cut: when from two propositions a third is deduced, and extra propositions are found from which one of those two can be deduced, then the same conclusion can be deduced from the other of the two plus those extra propositions. Suppose that one can deduce ‘n are few’ from ‘1 is few’, ‘If 1 is few then 2 are few’, . . . , ‘If $n - 1$ are few then n are few’. By *modus ponens*, from the two propositions ‘n are few’ and ‘If n are few then $n + 1$ are few’ one deduces the conclusion ‘$n + 1$ are few’. The ground-rule then permits one to deduce the conclusion ‘$n + 1$ are few’ from the premises ‘1 is few’, ‘If 1 is few then 2 are few’, . . . , ‘If n are few then $n + 1$ are few’. By continuing in this way, one eventually reaches a complete Stoic analysis of the sorites argument above into basic inferences.[29]

On Stoic terms, the argument is valid, its first premise is true and its conclusion false. Thus not all the conditional premises are true. By the Stoic principle of bivalence, at least one of them is false. Yet the gradualness of the sorites makes them all appear true. How can one of the conditionals ‘If i are few then $i + 1$ are few’ be false?

At this point there is a complication. The truth-conditions of conditionals were the subject of a fierce controversy that went back to Diodorus and his contemporary Philo, and was taken up by the Stoics. In

Alexandria the poet Callimachus wrote ‘Even the crows on the roof tops are cawing about the question which conditionals are true’. Philo treated the truth-value of the conditional as a function of the truth-values of its components. ‘If *P* then *Q*’ is true in three cases: ‘*P*’ and ‘*Q*’ are true; ‘*P*’ is false and ‘*Q*’ is true; ‘*P*’ and ‘*Q*’ are false. It is false in case ‘*P*’ is true and ‘*Q*’ is false. The Philonian conditional is the weakest construction to obey *modus ponens,* and therefore the weakest kind of conditional; it is true if any conditional with antecedent ‘*P*’ and consequent ‘*Q*’ is true. The Philonian ‘If *P* then *Q*’ is equivalent to ‘Not: *P* and not *Q*’. In contrast, Diodorus held ‘If *P* then *Q*’ to be at least as strong as ‘Not ever: *P* and not *Q*’. Chrysippus went still further; for him, a conditional is true if and only if its antecedent is incompatible with the negation of its consequent. Thus ‘If *P* then *Q*’ becomes equivalent to ‘Not possible: *P* and not *Q*’. In modern terms, Philo’s conditional is material implication and Chrysippus’ is strict implication. Later Stoics tended to follow Chrysippus.

In the sorites argument, some conditional premise ‘If i are few then i + 1 are few’ is supposed to be false. If the conditional is Chrysippan, it is false if and only if ‘i are few’ is compatible with ‘i + 1 are not few’. However, this conclusion looks banal; who ever thought them incompatible? Chrysippus might cheerfully allow that all the conditional premises, so taken, are false. To know the falsity of such a conditional is not to identify a cut-off point; it is merely to know that a certain point is not logically debarred from being the cut-off. Some modern philosophers would disagree, holding that sorites puzzles arise because vague concepts are subject to tolerance principles which do rule out the possibility of cut-off points. For them, ‘i are few’ does threaten to be incompatible with ‘i + 1 are not few’, making the Chrysippan conditional ‘If i are few then i + 1 are few’ true. But the Stoics did not take that view, and may not have regarded the argument with Chrysippan conditional premises as genuinely challenging.[30]

The most challenging form of the sorites argument uses Philonian conditionals. Its premises claim the least, and are therefore the hardest to deny. Since the Philonian reading of the conditional was not the standard one, the premises had to be formulated explicitly as negated conjunctions. Just that was done in standard Stoic accounts of the argument:

> Not: 2 are few, but not 3 as well. Not: the latter but not 4 as well. And so on up to 10. But 2 are few. Therefore 10 are few as well.[31]

The manoeuvre does nothing to *solve* the paradox. For example, it cannot remove any semantic or conceptual pressure against cut-off points. If the argument with Chrysippan conditionals were in any respect more paradoxical than the argument with negated conjunctions, it would be mere evasion to replace the former by the latter. The point is the reverse: to confront the paradox in its most telling form. As already noted, the argument with Chrysippan conditionals may not have been regarded as paradoxical at all. In any case, the Chrysippan conditionals entail the negated conjunctions, so anything that generates the argument with Chrysippan conditionals will also generate the argument with Philonian conditionals.

Once the explicit conditional has been eliminated, *modus ponens* can no longer be used to drive the argument forward. But Stoic logic still obliges. Its third indemonstrable principle is 'Not both the first and the second; but the first; therefore not the second'. Thus from 'i are few' and 'Not: i are few and $i + 1$ are not few' one can derive 'Not: $i + 1$ are not few', and Stoic principles allowed the double negation to be eliminated.[32] Such arguments can be chained together by the third ground-rule, just as before. Thus any argument of this form is valid on Stoic terms:

P_1
Not: P_1 and not P_2
Not: P_2 and not P_3
.
.
.
Not: P_n-_1 and not P_n

P_n

Indeed, Chrysippus sometimes supported his own doctrines by arguments with exactly the same form as the sorites.[33] Another Stoic example is:

> It is not the case that while fate is of this kind, destiny does not exist; nor that while destiny exists, apportionment does not; nor that while

> apportionment exists, retribution does not; nor that while retribution exists, law does not; nor that while law exists, there does not exist right reason enjoining what must be done and prohibiting what must not be done. But it is wrong actions that are prohibited, and right actions that are enjoined. Therefore it is not the case that while fate is of this kind wrong and right actions do not exist.[34]

The sorites argument with negated conjunctions is valid, its first premise is true and its conclusion false. Thus some premise of the form 'Not: i are few and $i + 1$ are not few' is false. Hence i are few and $i + 1$ are not few. Thus i is a sharp cut-off point for fewness. Since one cannot identify such a point, one is in no position to deny any of the premises. One can only suspend judgement. The challenge 'Which premise is false?' is unfair, for one may be unable to find out even though one knows that at least one premise is false.

What has been gained by presenting the sorites as an argument with premises and conclusion? Its logical structure was not the heart of the problem, for the argument is formally valid according to those whom it threatens, the Stoics. They used arguments with that structure themselves. As for the sceptics, they could suspend judgement on its logical status; it was enough for their purposes that their opponents took such arguments to be valid. The logical structure provides a convenient way of laying out the problem, but so far nothing more.

It is tempting to argue for a dialectical structure behind the logical façade. First, the use of conditionals in the sorites argument is a distraction, for the sorites interrogation shows that one can set the puzzle going in a language whose only resources are 'Yes', 'No' and simple sentences (without logical connectives such as 'if', 'and' and 'not') in the interrogative mood. Second, the argument has been persuasive so far not because its premises commanded assent, but because they forbad dissent. The problem was not that one could say 'Not: i are few and $i + 1$ are not few', but that one could not say 'i are few and $i + 1$ are not few'. One is not presumed to believe the premises of the sorites argument. The point of the questions is to force one to take up attitudes for or against the individual propositions, for any pattern of such attitudes leads one into trouble; hence the power of the interrogative form. Since the premises of the sorites argument seem compelling only when one is interrogated on their components, the question form takes primacy.

The situation is transformed if the premises of the sorites argument can be given positive support. If they can, the argument form takes primacy: the question form leaves too much unsaid. Moreover, Chrysippan silence is no longer an adequate response, for it does not undermine the positive support for the premises. One strand in later developments was the attempt to provide that support.

1.4 THE SORITES IN LATER ANTIQUITY

Chrysippus' strategy may have satisfied later Stoics. There is no sign of attempts to take the investigation further. The puzzles became a standard and perhaps stale item in the Stoic logical curriculum.[35]

The sceptics were not satisfied with Chrysippus' silence. It was most notably attacked half a century after his death by Carneades, who renewed the scepticism of the Academy. 'For all I care you can snore, not just become quiescent. But what's the point? In time there'll be someone to wake you up and question you in the same fashion.' Chrysippus was dialectically no better off than he would have been had he fallen asleep. 'Why should your pursuer care whether he traps you silent or speaking?'[36]Carneades' attitude was that of a chess-player with what he takes to be a winning strategy, whose opponent simply refuses to make a move (in a game without time-limits).

Suspension of judgement was the sceptical attitude, and Carneades fastened on the extent to which Chrysippus' strategy allowed it to spread. If Chrysippus suspended judgement in clear cases, on what basis did he object to the sceptic's suspension of judgement? The question does not reduce the strategy to immediate incoherence, for some sort of reply is open to Chrysippus: do not suspend judgement when the case is discriminable from the unclear cases. Nevertheless, the Stoic is in a very delicate position. He stops at clear cases, avoiding the risk of giving answers that are not clearly correct only at the cost of failing to give answers that are clearly correct (in the Stoic sense). The Stoics' epistemological caution enlarged the concessions to scepticism that their bivalent semantics forced them to make under sorites questioning. The concessions did not amount to surrender, for cases remained in which they could still claim knowledge, but these cases were marked off by a disputed no man's land rather than a compelling principle. Perhaps that is just what knowledge is like.

Carneades made particular use of sorites reasoning against the Stoic theory of the gods:

> If Zeus is a god, . . . , Posidon too, being his brother, will be a god. But if Posidon [the sea] is a god, the [river] Achelous too will be a god. And if the Achelous is, so is the Nile. If the Nile is, so are all rivers. If all rivers are, streams too would be gods. If streams were, torrents would be. But streams are not. Therefore Zeus is not a god either. But if there were gods, Zeus too would be a god. Therefore there are no gods.[37]

Unlike standard sorites series, this one has varying relations between successive members. Posidon is the brother of Zeus, but a stream is a smaller version of a river. Nevertheless, it is clearly a little-by-little argument. The moral is not intended to be that there are no gods; rather, it is that Stoic rational theology fails, because its attempts to demarcate divinity by non-arbitrary general principles provide it with no way to resist any one step. Carneades supports each premise by separate considerations whose cogency the Stoic is supposed to grant. Mere silence would not be an adequate defence against this argument.[38]

If the screw is to be tightened in the usual sorites arguments, their premises need support. Awareness of this is shown by Galen (CE *c*. 129–*c*. 199):

> If you do not say with respect to any of the numbers, as in the case of the 100 grains of wheat for example, that it now constitutes a heap, but afterwards when a grain is added to it, you say that a heap has now been formed, consequently this quantity of corn becomes a heap by the addition of the single grain of wheat, and if the grain is taken away the heap is eliminated. And I know of nothing worse and more absurd than that the being and non-being of a heap is determined by a grain of corn.[39]

Chrysippus could not suspend judgement on the general claim that one grain does not make the difference between a heap and a non-heap. He must deny it, for it contradicts the existence even of an unknown cut-off point. For him, the addition of one grain can turn a non-heap into a heap.

Galen's interest in sorites puzzles was connected with a long-running dispute between Empirical and Dogmatic (one might say Rationalist) Doctors. The Empirical Doctors based their medical knowledge on inductive inferences, holding it to be reliable only if derived from sufficiently many observations; their opponents applied sorites reasoning against the notion 'sufficiently many'. The Empirical Doctors replied that the argument proved too much; if it destroyed the notion 'sufficiently many', it would by parity of reasoning destroy much of the common sense on which all must rely. They gave the examples of a mountain, strong love, a row, a strong wind, a city, a wave, the open sea, a flock of sheep and a herd of cattle, the nation and the crowd, boyhood and adolescence, the seasons: none would exist if sorites arguments were sound. The Empirical Doctors could reasonably claim that sorites arguments were unsound, without being able to say exactly where the flaw lay. Not even Chrysippus could say which premise in negated conjunction form was false.[40]

There are also signs of a rather different Empiricist point. The sorites questioner is compared to someone who asks a shoemaker what last will shoe everyone: the question has no answer, for different feet require different lasts. The idea may be that the required number of observations depends on the circumstances of the particular case. There is no general answer to the question 'Are fifty observations enough?'.[41] The point has been repeated by modern philosophers, and is correct as far as it goes, but that is not very far. For the questions can be asked about a particular case, and the Empiricist still cannot plausibly claim to know all the answers. Similarly, fifty grains may make a heap in one arrangement and not in another, but in any particular process of heaping up grains one by one there will come a point at which the right answer to the question 'Is this a heap?' is unknown.

As already noted, the Empiricists knew that susceptibility to sorites puzzles is widespread. They had no interest in logic, and did not to attempt to demarcate the phenomenon in logical terms. The logicians themselves regarded sorites arguments as instantiating a logically valid form, the trouble lying in the premises. For both, sorites susceptibility depends on content, but how? The Empiricists attributed it to everything which has 'a measure or multitude'. Burnyeat complains that this characterization is too narrow, since it excludes puzzles based on qualitative rather than quantitative variation, such as a series of shades of colour from red to orange.[42] However, the Empirical Doctors' own example of strong love suggests that they had a broader notion in mind, for the strength of love is intensive rather than extensive. Perhaps they meant that anything which comes in degrees can give rise to sorites puzzles.

It was known that for every sorites series which proceeded by adding (as Eubulides' original series seem to have done), a reverse sorites series proceeded by subtracting. Thus examples tend to come in pairs of opposites: rich and poor, famous and obscure, many and few, great and small, long and short, broad and narrow.[43] The awareness of reversibility no doubt helped to check the tendency to think of a sorites puzzle as showing its conclusion to be strange but true, for the conclusion of one sorites argument contradicts the first premise of the reverse argument.

Some modern writers have argued for a special connection between sorites puzzles and concepts applied on the basis of observation. The preceding examples show no sign of this. Some followers of Aristotle did make a connection, but not quite the modern one. In commentaries on Aristotle's *Nicomachean Ethics*, both Aspasius (CE *c*. 100–150) and an anonymous writer took the puzzles to show that empirical concepts must be applied case by case. General rules are no substitute for observation and good judgement. Where the modern writers take observation to pose the puzzles, the Aristotelian commentators took it to solve them. But can even a man of good judgement always tell by looking whether something is a heap, as the grains pile up one by one?[44]

The Dogmatic Doctors presumably hoped that their use of sorites arguments against rivals would carry conviction with potential clients, who might use the debate to choose their doctor. It would not have been profitable to use arguments that everyone knew were unsound. Nevertheless, sorites paradoxes were well known outside philosophy. Horace playfully used a sorites series for 'old' against the belief that old poets are best, and Persius one for 'rich' against the desire to become rich.[45]The general reader was expected to catch allusions to heaps. The notoriety of slippery slope arguments no more prevented their frequent employment than it does today.

One late development may be mentioned. The fourth-century rhetorician Marius Victorinus applied the term 'sorites syllogism' to the iterated syllogisms already discussed by Aristotle:[46]

All F_1s are F_2s
All F_2s are F_3s
.
.
.
All F_n–1 are F_ns

All F_1s are F_ns

This should be compared with the Stoic forms of argument discussed above. They are not quite the same. The Stoics iterated rules of inference involving complex premises built out of simpler propositions; in 'If the first then the second; but the first; therefore the second', 'the first' and 'the second' are variables for propositions. What Victorinus iterated was the Aristotelian syllogistic form with premises 'All *F*s are *G*s' and 'All *G*s are *H*s' and conclusion 'All *F*s are *H*s', where the variables stand in for sub-propositional terms such as 'horse' and 'tree'. In modern terms, the Stoics were doing propositional logic, the Aristotelians predicate logic. The point of analogy is the chaining together of arguments, not the nature of the arguments so chained together.

His combination of Stoic and Aristotelian terminology aside, Victorinus was not an innovator in logic. However, his writings had some influence in the Christian world of later antiquity, for he was a Christian convert and taught St Jerome. That made him an acceptable channel for the transmission of logical doctrine, and 'sorites syllogism' would later become standard terminology.[47] But it did so simply as a name for compound syllogisms, conceived as unproblematic. The connection with the sorites as a puzzle had faded out.

1.5 THE SORITES AFTER ANTIQUITY

A thousand years may be passed over in a few words. Sorites puzzles formed no part of the medieval logic curriculum, no doubt because they were absent from the works of Aristotle and his followers.[48] Their revival had to wait for what was in other respects the corruption and decline of logic in the Renaissance.

Lorenzo Valla (1407–57) was one of the chief instigators of a shift in the curriculum from the formal rigour of scholastic logic to the more literary pursuit of humanist rhetoric. An ordinary language philosopher of the Renaissance, he insisted that argument should be conducted in idiomatic Latin. The achievements of medieval logic were dismissed as solecisms of grammar, for they depended on the artificial regimentation of Latin as a semi-formal language. Valla taught the art of arguing, but as an essentially informal and practical skill. In this context he discussed the sorites, of which he knew from Cicero's account. The aim was to inculcate the ability to distinguish between sound and sophistical arguments in particular cases.

Valla uses the term '*coacervatio*' for the heaping – the iteration – of syllogisms. As such, he treats it as a valid form of argument. He gives the example: 'Whatever I wish, my mother wishes; whatever my mother wishes, Themistocles wishes; whatever Themistocles wishes, the Athenians wish; therefore whatever I wish, the Athenians wish'. This is a clearly valid sorites syllogism; *if* its premises are true, so is its conclusion. Valla's main concern is to bring out the difference between this kind of reasoning and invalid kinds easily mistaken for it. One of his examples of the latter is: 'Rome is the most beautiful of all cities; this district is the most beautiful of all districts in Rome; this house is the most beautiful of all houses in this district; therefore this house is the most beautiful of all houses in all cities'. This argument is clearly fallacious; the beauty of the whole is not determined only by the beauty of its most beautiful part. The individual links in the chain are non-syllogistic. The problem lies not in the heaping but in what is heaped. Valla suggests that such arguments may be invalid when they concern wealth, nobility, strength, beauty and knowledge, but will be valid when they concern warmth, cold, height and distance; perhaps the intended contrast is between many-dimensional and one-dimensional phenomena. Valla's suggestion is at best a fallible rule of thumb; the warmest house may not be found in the warmest city. Such a rule is of little use to the theory of reasoning, but may be a useful contribution to its practice. Valla also considers standard sorites puzzles, mainly in question and answer form. He insists that each grain and each hair makes *some* relevant difference, however small, and ends by quoting Cicero's remark that in such cases we can never know exactly where to draw the line. This is quite consistent with the claim that we can learn to know unsound slippery slope reasoning when we see it.[49]

After Valla, the history of the sorites syllogism diverged from the history of the sorites sophism. For textbooks of logic, a sorites was a multiple syllogism in the orthodox sense of 'syllogism'.[50] A favourite example came from Romans 8: 29,30: 'For whom he did foreknow, he also did predestinate to be conformed to the image of his Son. . . . Moreover whom he did predestinate, them he also called: and whom he called, them he also justified: and whom he justified, them he also glorified.' This sense of the term 'sorites' survived as long as syllogistic logic, and is still in occasional use. The liveliness of this tradition may be judged from its major innovation. Rudolf Goclenius (1547–1628), Professor of Logic and Metaphysics at Marburg, listed the premises of sorites syllogisms in reverse order in his *Isagoge in Organum Aristotelis.* Sorites syllogisms were thenceforth known as Goclenian when so presented, and as Aristotelian when in the original order.[51] The distinction between Aristotelian and Goclenian forms became the main item of information in many a textbook's chapter on 'Sorites'. The etymology of the word was sometimes explained as a heaping up of premises, without reference to the paradox.

Sorites paradoxes appeared far less often in logic textbooks, and then usually under alternative terminology such as '*sophisma polyzeteseos*', the fallacy of continual questioning. They were preserved by their appearance in sceptical works by Cicero and Sextus Empiricus which became well known in the seventeenth century.[52] One philosopher of sceptical cast gave them attention – Pierre Gassendi, in his history of logic *De Origine et Varietate Logicae* – but little use was made of the sorites for sceptical purposes in the period. The sceptics of the Academy had exploited it against Stoic logic; only with the rise of modern logic at the end of the nineteenth century would it again be perceived as a serious threat.

Leibniz mentions both the Heap and the Bald Man in *New Essays on Human Understanding*, his reply to Locke's *Essay Concerning Human Understanding*. Locke had argued that the boundaries of sorts and species are fixed by the mind, not by the nature of the things classified, supporting his claim by appeal to the existence of monsters on the boundaries between species. Leibniz replies that such borderline cases could not show that all classification is purely stipulative. The line between man and beast, or between stabbing and slashing, is fixed by the nature of what it divides. However, he allows that in cases of insensible variation some stipulation is required, just as we must fix our units of measurement (nature does not require us to measure in inches), and associates such cases with sorites paradoxes. Bald men do not form a natural kind; nor do heaps.[53] A borderline case is then a matter of opinion; different opinions may be equally good. Leibniz gives the outer limits of colours as an example. He also mentions Locke's remark that the same horse will be regarded as large by a Welshman and as small by a Fleming, because they have different comparison classes in mind – although here one might wish to distinguish between vagueness and context dependence. Leibniz concludes that, although different species could be linked by sorites series of insensible transitions, in practice they usually are not, so that few arbitrary stipulations are required in the classification of what actually exists.[54]

Unlike the Stoics, Leibniz holds that some of the questions in a sorites interrogation do not have right and wrong answers. The indeterminacy would remain even if we were perfectly acquainted with the inner natures of the creatures in question. Our inability to answer is not simply a product of ignorance. A similar moral would later be drawn by Alexander Bain in his *Logic*. Under the canons of definition he discussed borderline cases, connecting them with the Heap, and concluded 'A certain *margin* must be allowed as *indetermined*, and as open to difference of opinion; and such a margin of ambiguity is not to be held as invalidating the radical contrast of qualities on either side'.[55]

Sorites paradoxes play a quite different role in Hegel's logic. He uses them to illustrate not indeterminacy but the way in which a quantitative difference can issue in a qualitative one. The temperature of water rises or falls, until it turns to steam or ice. The peasant adds ounce after ounce to his ass's load, until its spine snaps. Grain after grain is added and a heap starts to exist. The assimilation of these cases is surprising, for it seems to require a sharp cut-off point between non-heap and heap, so that quality can be a discontinuous function of quantity. Hegel had something broader in mind, however, as his next examples show. A small increase or decrease in expenditure may not matter, but more will amount to prodigality or avarice. The constitution of a state need not be changed every time a citizen is born or dies, or an acre is added to or subtracted from its territory, but nevertheless the constitution appropriate to a small city state is inappropriate to a vast empire. Differences of degree, however slight, cannot be considered without implication for differences of kind. In unHegelian terms, sorites paradoxes refute generalizations of the form: if x and y differ by less than quantity q, and x has quality Q, then y has quality Q.[56]

According to the principle of excluded middle, everything is either a heap or not a heap. Given Hegel's low opinion of this principle, 'the maxim of the definite understanding, which would fain avoid contradiction, but in doing so falls into it', one may find it surprising that he did not use the sorites as a stick with which to beat it.[57] Only when philosophers accorded more respect to the principle did the paradox emerge as one of the most significant obstacles to its application.

Chapter 2

The ideal of precision

2.1 THE EMERGENCE OF VAGUENESS

Only two traditions in the history of philosophy have found a serious problem in sorites paradoxes. One culminated in Stoicism; the other is modern analytic philosophy. The reason has to do with the centrality of formal logic in both traditions. The conclusiveness of formal proof makes a sorites paradox hard to fudge. Its premises once granted, its disastrous conclusion seems inescapable. This chapter describes the way in which the rise of modern logic at the end of the nineteenth century caused sorites paradoxes to become again a topic of philosophical discussion after two thousand years. There are, of course, obvious differences in the treatment of sorites paradoxes between the Stoics and early analytic philosophers. They can to some extent be explained by differences in their conceptions of logic.

This chapter is also the history of the emergence of a concept. Contemporary analytic philosophy treats sorites paradoxes under the heading of 'vagueness', as in this book. The word has been appropriated as a term of art for the phenomenon of blurred boundaries, which results in sorites susceptibility. Its use in ordinary English is much less restricted. Someone who asks 'How tall is he?' may reasonably criticize the answer 'Between two foot and eleven foot' as 'vague' simply because it is uninformative.[1] The criticism is that its boundaries include too much, not that they are blurred. In current philosophical usage, that would count as a misuse of the word. The emergence of the problem of vagueness required the emergence of the concept of vagueness.

Three philosophers at the beginning of the analytic tradition will be discussed: Gottlob Frege, Charles Sanders Peirce and Bertrand Russell. For Frege, a logical calculus is a powerful but delicate engine which cannot tolerate grit. Blurred boundaries make one form of grit. Logic is to be formulated in an artificial language of perfect precision. Vagueness, like madness, must be mentioned in order to be excluded. Unlike Frege, Peirce and Russell produced theories of vagueness. Although Peirce often seems to be talking about something quite different, it would be a mistake to ignore what he says, for his usage reflects the inchoate ordinary sense of the term, and enables the subsequent distillation of the technical concept to be seen for the philosophical achievement it was. Russell helped to make the technical sense canonical, although even his theory fails to demarcate vagueness according to his intentions. With him, the problem of vagueness is systematically presented for the first time in something close to its current form.

2.2 FREGE

1879 is more than conventionally taken to mark the beginning of modern logic. That year Frege published his *Begriffsschrift*, subtitled ‘a formula language, modelled upon that of arithmetic, for pure thought’. He constructs a language in which the central concepts of logic – e.g. ‘not’, ‘if’, ‘all’, ‘identical’ – can be expressed and combined: in modern terms, a predicate calculus.[2] The syntax of the language is rigorously defined by rules specifying exactly which patterns of symbols count as well-formed formulas; its semantics is defined by rules specifying the meanings of all well-formed formulas. Logical axioms and rules of inference for the language are also given. One can mechanically check whether a proof is permitted in the system, by seeing whether the formula at each step is either an axiom or follows by a rule of inference from formulas at previous steps. Frege describes his system as a partial realization of Leibniz’s project for a *calculus ratiocinator*, in which all valid reasoning is reduced to mechanical steps. A vastly wider range of inferences can be derived in it than in any previous system of logic.

The *Begriffsschrift* language is ‘for pure thought’. Its lexicon contains only logical symbols and variables. It has no words for specific empirical objects, properties or relations. Such words play no role in Frege’s long-term aim of defining mathematical concepts in terms of purely logical concepts, and deriving the axioms of mathematics from axioms of pure logic. This logicist project was a partial failure, as was Russell and Whitehead's later attempt in *Principia Mathematica.* But even if the axioms of mathematics cannot be derived from axioms of logic, all ordinary mathematical reasoning from axioms to theorems can be carried out in a system like Frege's: by itself an extraordinary advance.

To call Frege's formula language artificial is not merely to point out that it was deliberately constructed and is no one's first language. It is to assign his syntactic and semantic rule-making a different status from that possessed by the activity of a grammarian for a natural language. Frege's definitions are to be treated as stipulative; unlike dictionary definitions, they do not describe an independently existing practice. What it is for a symbol in an artificial language to mean something is for it to be stipulated to have that meaning. If no meaning is stipulated, then the symbol has no meaning. This is the status usually assigned to definitions in mathematics.[3] If no sense has been given to division by zero, then '0/0' is senseless; it cannot, unknown to us, really mean something. Any defect in a stipulative definition automatically makes the defined term defective in meaning. Frege held that many definitions in the foundations of mathematics were indeed defective, often because they made no stipulation for certain cases. By making his definitions complete and consistent, he aimed to ensure that no semantically defective expressions entered his formula language. But the possibilities for defects of definition gave him a way of thinking about semantically defective expressions in natural languages too. Thus a vague term such as 'heap' may be thought of as incompletely defined, because in borderline cases it is neither stipulated to apply nor stipulated not to apply. Of course, expressions of natural language do not literally have their meanings by stipulation, but our use of them may be thought of as having the effect of a stipulation. Frege did indeed tend to lump vagueness with incompleteness of definition.[4]

The Stoics' epistemic solution of sorites paradoxes would have been of no use to Frege. If a definition is stipulative, it cannot draw a line at a case for which no stipulation has been made. In such cases there is no question of unknowable cut-off points. They might still be postulated in natural languages, where definitions are not stipulative. Frege had no objection to unknowable cut-off points as such; they are quite congenial to his uncompromising realism. But he was not interested in natural language for its own sake, and his concern with artificial language suggested an alternative model of sorites susceptibility, inconsistent with the Stoic model.

An epistemic solution to sorites paradoxes has been proposed within the analytic tradition only recently. It is tempting to attribute its previous unpopularity to the residue of a definitional picture of meaning. On this picture, our use of a term implicitly defines its meaning. Since explicit definitions are liable to various defects, so equally are implicit definitions. Where we do not draw a line, no line is drawn. If the Stoic view is correct, the implicit definition account is radically misleading: but that is for another chapter.

Frege's specific remarks on vagueness must be seen in the light of his conception of a logically perfect language.[5] What must a language be like to permit the rigorous conduct of mathematical reasoning? Mathematical symbols can be combined in arbitrarily complex formulas, so the language will contain infinitely many possible sentences. Thus one cannot demarcate the well-formed formulas, and assign them meaning, by listing them. Rather, one will stipulate *recursive rules*. The syntactic rules will list the simple well-formed expressions of various types, and then the ways in which complex expressions may be well formed by combination of simpler well-formed expressions. The semantic rules will assign the well-formed expressions meaning in a parallel way. If a rigorous system of logic is to be formulated, syntax and semantics must somehow correspond. Rigour demands that a rule of inference be specified in purely syntactic terms, so that adherence to it can be mechanically checked. But a rule of inference must also be valid: it must preserve truth from premises to conclusion. Thus the visible syntactic structure of the premises and conclusion must somehow reflect their semantic structure, for the latter determines whether they are true. Indeed, for logical purposes the *only* relevant aspects of meaning are those which contribute to making sentences true or false. The contribution of an expression to making arguments valid or invalid by making sentences in which it occurs true or false is its *referent*. The truth-value of a sentence is worked out recursively as a function of the referents of its constituent expressions. In a logically perfect language, every well-formed expression has one and only one referent, invariant from context to context, for any variation would constitute a semantic difference not manifested in any syntactic difference, and might therefore invalidate rules of inference formulated in purely syntactic terms.

In the simplest case, a sentence is formed from a name and a predicate: for example, '7 is prime'. The sentence is true because the object named by '7' falls under the concept predicated by 'is prime'. Thus the referent of the name '7' is simply the number 7. More generally, Frege takes the referent of a name to be its bearer. The role of the predicate 'is prime' is to divide objects into two classes, the primes and the rest; '7 is prime' is true because 7 falls in the former class, '8 is prime' false because 8 falls in the latter. Thus Frege takes the referent of a predicate to be a function associating each object with a truth-value, truth or falsity. The truth-value of the sentence is the result of applying the referent of the predicate, a function, to the referent of the name, an object. The referent of 'is prime' is a function which maps each prime to truth and each non-prime to falsity. Any function mapping objects to truth-values is a *concept* in Frege's technical sense.[6] An object *falls under* a concept if and only if the concept maps the object to truth.

In his later work, Frege contrasted the referent of an expression with its *sense*. The latter is the way in which the expression presents its referent to those who understand it. For example, the complex names '7 + 5' and '10 + 2' have the same referent, because they name the same number, but different senses, because they present it in different ways. Since the referent is what the sense presents, sense determines reference. Since the referent of a complex expression is presented as a function of the presented referents of the component expressions, the sense of the former is composed of the senses of the latter.

How does vagueness fit into this framework? Recall Frege's assimilation of it to failures of definition. To define a concept, one must lay down a rule mapping each object to a truth-value. In attempting to do so, one might forget to stipulate for certain cases. For example, suppose that one stipulates that all positive integers be mapped to truth, and all negative ones to falsity, but makes no stipulation for zero. One has failed to specify a concept. Some conditions have been laid down, but they are met both by a concept mapping zero to truth and by one mapping it to falsity. No one concept has been singled out. It would be quite wrong to say that one has specified a partial concept, a concept undefined for some objects. To do that, one would need to stipulate that the concept be undefined for zero, but one has not got even that far. To fail to stipulate a value is not to stipulate that there be no value.[7] Frege would insist that *some* stipulation be made even for non-numerical cases, such as the moon. Unless a stipulation has been made for every case, nothing has been defined.

Mathematicians sometimes define a function to be 'undefined' in certain cases; this is quite different from not making a stipulation for those cases.[8] This practice is coherent by Fregean lights only if one thinks of 'undefined' as another value that a function can take. After all, there is nothing *indeterminate* about such cases. Frege denied the existence of partial functions which leave some objects unmapped to anything. Thus there are no partial concepts. In more familiar terms, all concepts have sharp boundaries. There are no vague concepts. If concepts are likened to areas on a plane, a vague concept would be an area without sharp boundary lines, gradually shading off into the background, and that, according to Frege, would be no area at all.[9] But the crucial point is that if our attempt to define a predicate is incomplete, it has not been assigned any one concept. It lacks a referent. Since Frege treated vagueness as incompleteness of definition, he held that a vague predicate has no referent. It is like a would-be name that does not name anything.[10]

What would happen if one introduced a vague predicate into Frege's formula language? In this language, the truth-value of a sentence is a function of the referents of its constituents. But a function cannot deliver a value until it is applied to something. Since a vague predicate lacks a referent, it provides nothing to which the relevant function can be applied. Thus the function does not deliver a value. No sentence containing a vague predicate has a truth-value, so no sentence containing it is true. For example, the formalization of 'Either Jack is bald or Jack is not bald' is not true; in other words, even instances of the law of excluded middle are invalid. Frege often identifies this law with the requirement that all concepts have sharp boundaries.[11] On his view, the formalization of 'Either Jack is bald or Jack is not bald' will lack a truth-value whether or not Jack is a borderline case of baldness, because 'is bald' simply lacks a referent. But even if 'is bald' referred to a partial concept, so that 'Jack is bald' lacked a referent only when Jack was a borderline case, that would be enough to generate counterexamples to the law of excluded middle, given Frege's semantic principle that the truth-value of a sentence is a function of the referents of its constituents. There is nothing special about excluded middle in this respect. Both 'If Jack is bald then Jack is bald' and 'Jack is not both bald and not bald' will lack a truth-value when Jack is a borderline case of baldness.[12] It is just that 'Either Jack is bald or Jack is not bald' reflects more vividly the semantic principle that the result of applying a concept to an object is either truth or falsity.

If predicate-involving formulas are to be recognizable as logically valid by their syntactic structure, all predicates in the language must be everywhere defined.[13] On Frege's view, this entails a ban on vague predicates. When such precautions have been taken, the result is a logically perfect language in which every sentence is either true or false.

Frege treats sorites paradoxes as another symptom of what goes wrong when vague predicates are admitted into the language. The Heap is mentioned in *Begriffsschrift*, in connection with a theorem which generalizes the principle of mathematical induction, that if 0 has a property P, and whenever a number n has P so does $n + 1$, then all natural numbers (0, 1, 2, . . .) have P.[14] Thus if removing one bean from a heap of beans always leaves a heap of beans, removing any finite number of beans from a heap of beans leaves a heap of beans. Some later writers have even held mathematical induction responsible for sorites paradoxes.[15] That is a mistake; if one specifies the number of beans in the original heap, one can reach the paradoxical conclusion without appeal to mathematical induction, as the Stoics did. Frege does not make that mistake; he uses the apparent failure of mathematical induction only to illustrate the disastrous results of applying rigorous logic to vague predicates. He could not concentrate the blame on mathematical induction alone, for he has derived it from other principles of logic, at least one of which must also fail for vague predicates.

Frege sometimes speaks of sorites paradoxes as the result of treating a vague predicate as though it stood for a sharp concept.[16] His description is not quite right, for someone who does think that 'heap' stands for a concept with a sharp cut-off point should not agree that one bean never makes the difference between a heap and a non-heap. The paradox arises just because we do try to acknowledge the vagueness of 'heap', by saying that one bean could never make the difference, but then find that our concession leads us into absurdities. Frege's point should be that, if we make this concession to vagueness, we cannot also expect to get away with using rigorous logic, which is designed only for referring and therefore precise expressions.

The account of vagueness is so far wholly negative. Yet Frege can hardly deny that vague language serves us perfectly well for many ordinary purposes. We use it to communicate information to each other. Some explanation is needed of our success, given that vague expressions fail to refer. Frege makes some gestures in this direction. However, Michael Dummett has argued that they should not be taken very seriously, since Frege holds that vagueness and other failures of definition are 'defects because no fully coherent account of a language exhibiting such features is possible . . . if a coherent theory of vague expressions could be constructed, vagueness would not be a defect'.[17] On this reading, it would be bad news for Frege if his explanation turned out to be coherent.

Dummett has overstated his case. No correct account can be given of the referents of vague expressions, because they have none. Vagueness has no logic, for logical laws are to be formulated in a language in which there is no reference failure. So much is Fregean. But Frege lumps vagueness together with ambiguity and partiality, of which coherent accounts surely can be given. If a mathematician uses the name '7' ambiguously for two different objects, one can coherently say what they are. Again, Frege compares a defective definition to an equation that is not guaranteed to have one and only one solution; a coherent account can be given of the equation, and therefore of the class of its solutions, even if that class is empty or multi-membered.[18] A bad definition imposes too many or too few constraints on the reference of the expression to be defined, but the constraints themselves can be stated. If vagueness is harder to describe in this way, Frege does not say so. His position does not rule out the possibility of a coherent account of vague expressions. His gestures towards one are therefore worth examining. What cannot be expected is a coherent Fregean theory that matches a vague expression of the language being theorized about with a synonymous expression of the language in which the theorizing is being done, for the latter expression would then have to be vague too, and therefore fail to refer; but reference failure in the theory itself would make it incoherent. In principle, a coherent Fregean account of a vague language must be formulated in a perfectly precise language. The difficulty of this task is evident; Frege did not claim to have come close to meeting it.[19]

Frege once suggested that ordinary language succeeds as a medium of communication because the senses of its sentences are not determined by the senses of their constituent words (contrast a logically perfect language). Thus I can use a sentence made up of vague words to get across to you a thought with a truth-value, even though my words individually fail to refer.[20]He may have had something similar in mind in *Begriffsschrift*, saying that, for some objects *a*, '*a* is a heap' cannot express a judgement. The idea is that it expresses a judgement only if it has a truth-value. Frege need not claim that 'heap' refers to a concept mysteriously applicable to some objects and not others to give a truth-value. Rather, he can deny that 'heap' expresses a concept at all, while allowing that some judgements of the form '*a* is a heap' express possible judgements and others do not. Logic requires expressions to have the same referents wherever they occur; vague natural languages violate this constraint.[21] On this view, there is no question of locating a specific principle of logic that fails in vague languages, for just about any principle fails when words change their meaning from occurrence to occurrence.

Unfortunately, Frege's account is quite implausible. It treats vague sentences like idioms. The verb phrase 'kick the bucket' is an idiom because we understand it to mean die not by understanding its component words but by remembering what we specifically learned, that it does mean that. In other cases when we do not understand the individual words, the context of a remark allows us to guess its meaning. However, most utterances are not understood like that, as Frege himself points out in other connections. If I say to you 'A heap of beans will be placed on your doorstep at noon tomorrow', you understand what I have said even though you have never encountered the sentence before and the context gives you no special clue as to its meaning. Frege would doubtless have been unperturbed by the failure of his explanation, not because he thought it had to fail, but because it is not a cartographer's job to explain why travellers with bad maps or none at all sometimes reach their destinations.

Frege had another line of explanation, by appeal to his distinction between sense and referent. He toyed with the idea that vague expressions have senses but not referents, comparing them to names in fiction.[22] One can understand 'heap', know its sense, and use it as though it referred to something, when really it does not. If a difference of sense does not entail a difference of referent, why should the existence of a sense entail the existence of a referent? However, it is controversial whether Frege's theoretical apparatus can sustain a sense without a referent. If an expression has no referent, there is no way in which it presents its referent; is that not just to say that it has no sense? Or can the expression have a sense without a referent by employing a way of presenting a referent while, in the event, presenting no referent? An assassin may employ a way of killing a man while, in the event, killing no man. The issue cannot and need not be decided here.[23]

Frege never developed the idea that vague expressions have senses without referents. His theoretical apparatus gives it little scope for development, for an expression is permitted to contribute to determining a truth-value only through its referent. Even if vague expressions have senses, their lack of referents entails that they contribute nothing to determining truth-values for sentences in which they occur; semantically they are idle. Suppose, for example, that 'There is a heap of sand on most building sites' is true. If the sense of 'heap' were a component of the sense of the whole sentence, the truth-value of the sentence would be determined by a determination of the referent of 'heap', amongst other things, for the sense of a sentence is the way in which its truth-value (its referent, according to Frege) is determined. Since the vagueness of 'heap' deprives it of a referent, the sentence would lack a truth-value. But the sentence is true, so its sense does not have the sense of 'heap' as a component. 'Heap' is semantically inert in the sentence. This is to return to Frege's other line of explanation, of vague sentences as idioms, which has already been seen to fail. It does not do justice to the manner in which our understanding of 'heap', our grasp of its sense, enters into our understanding of the whole sentence. If Frege allowed the sense of a vague word to be a component of the senses of sentences in which it occurs, he would have to deny that any such sentence has a truth-value: an equally extreme view. He was later inclined to the most ruthless view of all, denying both sense and truth-value to sentences involving vague expressions.[24]

Frege never gave an adequate account of vagueness, whether or not his overall theory entails that no such account can be given. Vague expressions do not simply fail to refer. Perhaps one could give a better account of vagueness in Fregean terms by assimilating it to referential ambiguity rather than to referential vacuity. But Frege had no interest in giving such an account. He discussed vagueness only in his efforts to escape it.[25]

2.3 PEIRCE

That successful inquiry involves a movement from vagueness towards precision is a commonplace. Both Frege and Peirce subscribed to it. They differed in this: for Frege, vagueness is to be eliminated at the beginning of inquiry; for Peirce, it is not to be eliminated before its end. On Frege's view, we cannot reason reliably until we have a precise language. On Peirce's, our language will always be vague. Vagueness is harmful only when it leaves the question at hand too unclear to be answered. We can then hope to clarify the question in relevant respects until it can be answered. What we cannot hope is ever to have achieved perfect clarity in all respects. Indeed, unnecessary precision does positive harm, cluttering up our theories with irrelevant complexity and rendering them too rigid to adapt to new evidence. If rational inquiry continues long enough, perhaps any specific question will eventually be made clear enough to be answered; but there will never come a time by which *every* question has been made clear enough to be answered.

Peirce gave a sustained discussion of vagueness in a paper of 1905, 'Issues of pragmaticism'. It figures as a characteristic of the indubitable tenets of common sense with which inquiry begins, a precondition of their certainty. For example, we cannot seriously doubt that there is order in nature, but the claim is a very vague one; if it were made more precise, the result would be a sophisticated hypothesis, open to serious doubt. Again, that fire burns is certain only because it is vague. The belief is not falsified when fire fails to burn a stone, for it does not specify precisely what fire burns in what circumstances.[26]

The certainty of many vague beliefs lies in their unspecificity. Now there are different ways of lacking specificity. If I believe that you are between 11 mm and 9,437 mm in height, my belief is unspecific because a wide range of heights would make it clearly true. If I believe that you are of average height, my belief is unspecific because a wide range of heights would make it neither clearly true nor clearly false, even though only a narrow range would make it clearly true. The former belief draws a sharp line around a wide area; the latter draws a blurred line around a small one. What the beliefs have in common is that a wide range of heights make them not clearly false. Of the two, most contemporary philosophers would classify only the belief that you are of average height as particularly vague. For them, vagueness is a matter of blurred boundaries, of cases neither clearly included nor clearly excluded. By this standard, the belief that you are between 11 mm and 9,437 mm in height is hardly vague at all. Peirce, in contrast, would have counted both beliefs as vague, for he regarded all unspecificity as a kind of vagueness. On this view, to be vague is to leave a wide range of cases not clearly excluded.

Peirce's wide use of 'vague' distances him from later work on vagueness. He had in fact shown himself master of the use now standard in philosophy in his entry for 'Vague (in logic)' for a dictionary of philosophy and psychology:

> Indeterminate in intention.
>
> A proposition is vague when there are possible states of things concerning which it is intrinsically uncertain whether, had they been contemplated by the speaker, he would have regarded them as excluded or allowed by the proposition. By intrinsically uncertain we mean not uncertain in consequence of any ignorance of the interpreter, but because the speaker's habits of language were indeterminate; so that one day he would regard the proposition as excluding, another as admitting, those states of things. Yet this must be understood to have reference to what might be *deduced* from a perfect knowledge of his state of mind; for it is

> precisely because those questions never did, or did not frequently, present themselves that his habit remained indeterminate.[27]

This is obviously aimed at the kind of unspecificity in the belief that you are of average height, not at the kind in the belief that you are between 11 mm and 9,437 mm in height; the 'possible states of things' inducing uncertainty are borderline cases for the belief. Peirce is attempting to explain indeterminacy of meaning by indeterminacy of use.[28]

When Peirce expounded his own views, he appropriated the word 'vague' for a wider concept, one that he took to be of more theoretical importance. Vagueness is contrasted with generality and determinacy. 'Man is mortal' is general; it requires that *every* man be mortal. 'This month a great event is to happen' is vague; it requires only that *some* great event happen this month, without specifying what.[29] Both vague and general sentences are indeterminate in the sense that their grammatical subjects are not perfectly specific: there are many possible men, and many possible great events. In contrast to the general 'Every man is brave' and the vague 'Some man is brave', the sentence 'Philip is brave', with a proper name as subject, may be treated as determinate in that respect.[30] In the sense of Peirce's dictionary definition, the negation of a vague sentence is equally vague, for they share the same blurred boundary. In the present sense, the negation of a vague sentence is general rather than vague; the negation of the vague 'Some man is immortal', for instance, is equivalent to the general 'Every man is mortal'.

Peirce does not simply define vagueness and generality in terms of the distinction between 'some' and 'every'. Rather, a sign is said to be vague when its further determination depends on the utterer, general when it depends on the interpreter.[31] Peirce gives as an example the vague remark 'A man whom I could mention seems to be a little conceited'. Although it may be very likely that the utterer had the person addressed in mind, she has not committed herself to the claim that he is a little conceited; she is free to specify someone else. The utterer, not the interpreter, has the right to determine the sign further. This is supposed to be the normal case, obtaining in default of any stipulation to the contrary. Sometimes, however, a convention transfers the right to determine the sign further to the interpreter. 'Man is mortal' can be applied to any man the interpreter chooses. Determining a sign in this context is a way of testing its truth.[32]

The contrast between determination by the utterer and determination by the interpreter can be connected with the contrast between 'some' and 'every' thus.

Suppose that you make an assertion of the form 'Some *F* is *G*', which I dispute. You point at an *F* and the dispute continues over your more specific assertion 'This *F* is *G*'. If I come to agree with you that this *F* is *G*, I must concede that your original assertion was correct. If you come to agree with me that this F is not *G*, it does not follow that your original assertion was incorrect, but your attempt to prove it correct has failed. Moreover, if your original assertion was correct, you should be able to prove it correct by pointing at a well-chosen *F*. The truth of your assertion 'Some *F* is *G*' corresponds to the existence of a winning strategy for you in the game in which you choose an *F* and we dispute over the more specific assertion.

Now suppose, in contrast, that you make an assertion of the form 'Every *F* is *G*', which I dispute. *I* point at an *F* and the dispute continues over your more specific assertion 'This *F* is *G*'. If you come to agree with me that this *F* is not *G*, you must concede that your original assertion was incorrect. If I come to agree with you that this *F* is *G*, it does not follow that your original assertion was correct, but my attempt to prove it incorrect has failed. Moreover, if your original assertion was incorrect, I should be able to prove it incorrect by pointing at a well-chosen *F*. The truth of your assertion 'Every *F* is *G*' corresponds to the existence of a winning strategy for me in the game in which I choose what to point at and we dispute over the more specific assertion.

If the truth of an assertion is to be tested by an example, the example should be chosen by the utterer of 'Some *F* is *G*' but by the interpreter of 'Every *F* is *G*'. Such an account has been formalized in Jaakko Hintikka's concept of a semantic game.[33]

One should not suppose that all general propositions are universally quantified, like 'Every *F* is *G*', and all vague ones existentially quantified, like 'Some *F* is *G*'.[34] On Peirce's account, conjunctions too are general, and disjunctions too are vague. If you assert 'John has been to Paris and London', I can choose which conjunct to attack. If you assert 'John has been to Paris or London', you can choose which disjunct to defend. Semantic games can be defined for conjunctions and disjunctions too. Indeed, the contrast between vagueness and generality can be drawn even for syntactically simple sentences. 'To affirm of anything that it is a horse is to yield to it *every* essential character of a horse; to deny of anything that it is a horse is vaguely to refuse to it *some* one or more of those essential characters of the horse.'[35] Here the explanation uses the quantifiers 'some' and 'every', but it does not imply that a quantifier occurs in the semantic structure of 'Dobbin is a horse'. Even if the concept 'horse' includes a list of essential properties, that is to conjoin them, not to quantify universally over them.

A vague utterance is true if some way of determining it results in a truth (hence the certainty of vague common sense). A general utterance is true only if every way of determining it results in a truth. The process of determination is not just a matter of progressively simplifying semantic structure, for if it were, it could be completed in practice, which is not Peirce's view. 'No communication of one person to another can be entirely definite, i.e. non-vague . . . wherever degree or any other possibility of continuous variation subsists, absolute precision is impossible. Much else must be vague, because no man's interpretation of words is based on exactly the same experience as any other man's.'[36] If the sharpening of blurred boundaries is treated as another form of determination, vagueness as Peirce defined it for the dictionary can be subsumed under his more general account. 'In another sense, honest people, when not joking, intend to make the meaning of their words determinate, so that there shall be no latitude of interpretation at all'; they 'intend to fix what is implied and what is not implied. They believe that they succeed in doing so, and if their chat is about the theory of numbers, perhaps they may.'[37] The aspiration of the honest people is evidently to eliminate not quantifiers but blurred boundaries. To move from 'Someone is sad' to 'He is sad' is not to fix the implications of 'Someone is sad', and to fix the implications of 'London is prosperous' by saying 'I don't mean that everyone in it is prosperous' is not to replace 'London is prosperous' by a semantically simpler sentence. Peirce's qualification 'In another sense' indicates an uneasy awareness of ambiguity in the use of 'vague', but he does not articulate the distinction.[38]

The distinction between vagueness and generality applies to the determining of unfixed implications too. Normally, utterers fix the implications of their assertions; this is vagueness. Occasionally, the fixing is left to the interpreter; this is generality. Peirce's example is 'That creature is filthy, in every sense of the term'.[39] It is not quite apt, for although 'every' hands the next move in the game to the interpreter, it does not permit the interpreter to decide which sense of 'filthy' is implied, for the utterer has already stipulated that every sense is. Some recent literary theorists claim that it is up to the interpreter rather than the author of a text to fix its implications, which is what one might expect interpreters to say. In the sharpening of blurred boundaries, the relative priority of utterer and interpreter may often depend on the context of the utterance rather than its structure. 'That is a heap' may be vague in some contexts, general in others.

Peirce's broad notion of indeterminacy, covering all kinds of vagueness and generality, gets its unity and point from his model of determinacy. A determinate proposition is absolutely specific. Whatever falls short of the ideal is to that extent indeterminate; having blurred boundaries is just one way of falling short. The ideal is unattainable, but it can be approximated better and better. The sharpening of blurred boundaries is one part of this cognitive task.

Peirce gives a second pair of definitions of vagueness and generality, said to be perhaps 'more scientific' than his first pair. At first sight, they look like a radical challenge to classical logic: 'anything is *general* in so far as the principle of excluded middle does not apply to it and is *vague* in so far as the principle of contradiction does not apply to it'.[40] In Peirce's terminology, a proposition satisfies the principle of excluded middle if it is either true or false, and satisfies the principle of contradiction if it is not both true and false. He allows that any given proposition satisfies both principles, but says that they may fail when it is indeterminate what proposition is in question. The new definitions might be rationalized thus.

An utterance is generally true only if every determination of it results in a truth, and generally false only if every determination of it results in a falsehood. Since some determinations of 'The number of bald men is even' result in a true proposition and others in a false one, it is neither generally true nor generally false. In this sense the principle of excluded middle does not apply.

An utterance is vaguely true if some determination of it results in a truth, and vaguely false if some determination of it results in a falsehood. Thus 'The number of bald men is even' is both vaguely true and vaguely false. In this sense the principle of contradiction does not apply.

Peirce's definitions are also intended for the case of ordinary quantified sentences. In his sense, a sentence such as 'A woman wrote *Middlemarch*' is vague, and the principle of contradiction does not apply to it. Yet the sentence is straightforwardly true. It is hard to see that his assimilation of the two cases casts much light on either. The kinds of determination at issue are too disparate. If rational inquiry falsifies one determination of 'A woman wrote *Middlemarch*', e.g. 'Jane Austen wrote *Middle-march*', the next step is to test a different determination. But if rational inquiry falsifies a determination of 'The number of bald men is even' with one stipulated cut-off point for 'bald', there is no point in testing it again with another; one stipulation is enough. Inquiry could not progress until vagueness was distinguished from unspecificity.

2.4 RUSSELL

Russell made the distinction between vagueness and unspecificity; his theorizing unmade it again. Vagueness is a recurrent theme in his work from 1913 to 1948. It appears as a natural phenomenon, one of the primitive ways in which representations – from memory images to words – diverge from what they represent. It also makes a logically perfect language physically impossible. His main ideas on the topic are formulated in the 1923 paper 'Vagueness', which set the agenda for most subsequent work.[41] The ensuing discussion is organized around ten of its theses.

(1) *Only representations are vague.* Russell insists that vagueness and precision are properties only of symbols, such as words, in relation to what they represent. Were there no representation, there would be no vagueness: 'Nothing is more or less what it is, or to a certain extent possessed of the properties which it possesses'. But the attribution to the world at large of properties of language is the occupational hazard of metaphysics, and Russell suggests that philosophers (such as Bergson?)

who emphasize 'the flux and the continuum and the unanalysability of the Universe' have projected the vagueness of language onto the world. Idealism receives a characteristic snub: 'It is thought that there must be some kind of identity between the knower and the known, and hence the knower infers that the known also is muddle-headed'.[42]

In confining vagueness to representations, Russell does not confine it to language. He allows thoughts and images as private representations. Not even all public representations are linguistic. Russell regards a thing's appearance as a public representation of it, recordable in a photograph or tape, and therefore potentially vague. Indeed, he holds that 'all vagueness in language and thought is essentially analogous to this vagueness which may exist in a photograph'.[43]

(2) *All language is vague.* Russell's next step is to argue that vagueness infects all words. His argument does not rely on a complete taxonomy of words; rather, he chooses examples for which it is plausible that, if any word were precise, some of them would be, and argues that none of them is precise.

The word 'red' is vague for 'there are shades of colour concerning which we shall be in doubt whether to call them red or not, not because we are ignorant of the meaning of the word "red", but because it is a word the extent of whose application is essentially doubtful'.[44] 'Doubtful' here means 'undefined' rather than 'unknown'; Russell is not suggesting that we might understand the word 'red' while merely being unable to know, of something we can see, whether it is red. His position is that it would not be true to say 'It must be either red or not red': mere ignorance of which disjunct was true would not undermine the truth of the disjunction.

'Second' (for a unit of time) is the kind of word that science has tried its hardest to make precise, but without perfect success. When Russell wrote, the second was defined by reference to the rotation of the earth; he points out that since the earth is not a rigid body and different parts of its surface rotate in different periods, the definition does not single out an exact length of time. He adds that since all observations have a margin of error, we can never know that an occurrence took exactly a second, but this seems to confuse the issue, which is whether the question ‘Did the occurrence last exactly a second?’ has a right answer, not whether we can know what the right answer is.

Even proper names are vague. The name ‘Ebenezer Wilkes Smith’ may be correctly applied to only one person, but it is vague in the temporal boundaries of its application, for birth and death are not absolutely instantaneous. Since Russell has said that only representations are vague, his claim is not that there is an individual thing, Ebenezer Wilkes Smith, the beginning and end of whose existence is only vaguely defined. Rather, he must hold that the name does not single out a unique thing; it is agreed that it should name a person, but our concept ‘person’ is itself vague, being compatible with different standards for defining the beginning and end of existence; they determine very slightly different sorts of ‘person’, some of which live fractionally longer than others.

The words with the best chance of avoiding vagueness are logical connectives, such as ‘or’ and ‘not’. In arguing that even they are not perfectly precise, Russell begins with the words ‘true’ and ‘false’. He argues that the vagueness of the judgements to which they are applied induces a corresponding vagueness in ‘true’ and ‘false’ themselves. If ‘This is a man’ is vague, so is ‘“This is a man” is true’. Now the meaning of a logical connective is given in terms of the concepts of truth and falsity; for example, ‘*P* or *Q*’ is true if and only if either ‘*P*’ is true or ‘*Q*’ is true. Thus the vagueness of ‘true’ and ‘false’ infects the logical connectives with vagueness too. All words are vague. Russell concludes that all language is vague.

Russell’s argument neglects the relation between the vagueness of words and the vagueness of complex expressions.[45] For example, although the noun phrase ‘enormous and tiny heap’ is made up of vague words, it is not itself vague in the same way, for it has no borderline cases; it is just a contradiction in terms. Even if all words can have borderline cases, not all complex expressions can.

Perhaps Russell would have been willing to define a complex expression as vague if and only if at least one of its simple constituents was vague, whether or not the expression as a whole was capable of borderline cases. Since all complex expressions are made up of words, in order to show that all expressions are vague it would be sufficient to show that all words are vague. One difficulty for Russell with this definition is that, if a complex expression is incapable of borderline cases, then we can use it to define a precise word. For example, we could define an 'eth' as an enormous and tiny heap. There cannot be a borderline eth. Would Russell therefore have to allow 'eth' as an exception to his claim that all words are vague? In order not to do so, he would need to work at a level of semantic structure or logical form rather than of surface syntax, so that he could classify 'eth' as semantically or logically complex, although it is a single word. The notion of a word, being defined at the level of surface syntax, is not quite the one he needs.

There is a deeper problem. On Russell's account, a word is vague just in case it can have a borderline case, in which its application is 'essentially doubtful'. Now what is essentially doubtful is a judgement, whose proper linguistic expression is in a sentence rather than a word. It is essentially doubtful whether *this is red.* A demonstrative and copula must be added to 'red' to form a suitable sentence. Words of other categories require more complex additions. For example, 'many' seems to be vague; to use it to express an essentially doubtful judgement we need a sentence like 'Many numbers under 20 are prime'. The primary application of 'vague' is to sentences, not to words. But the vagueness of a sentence does not imply the vagueness of every constituent word.[46] One vague word is enough. It may be essentially doubtful whether this is a red shape, because it is essentially doubtful whether this is red, although beyond doubt that it is a shape. The vagueness of 'This is a red shape' does not imply the vagueness of 'This is a shape'. But how then can the vagueness of 'Many numbers under 20 are prime' show 'many' to be vague, unless the other words in the sentence are assumed *not* to be vague? For if they were vague, they might be alone responsible for the vagueness of the sentence. How can even the vagueness of 'This is red' show 'red' to be vague, unless 'this' is assumed to be precise? The trouble is that Russell's claim, that all words (including 'prime' and 'this') are vague, contradicts the assumption which he seems to need in arguing for that claim. If all the basic constituents of a sentence are precise, then the sentence will be precise too.[47] Thus if all the basic constituents but one in a vague sentence are precise, the remaining one must be vague. Unfortunately, Russell's position forbids him to use this principle in its defence. How then can he show that all words are vague?

Russell might try to argue as follows. The vagueness of ‘this’ and ‘is’ does not prevent them from having clear application in a favourable context. Thus if one can find such a context in which the sentence ‘This is red’ expresses an essentially doubtful judgement, the blame must lie with ‘red’, so ‘red’ will have been shown to be vague. However, Russell’s theory gives little scope even to this form of argument. If ‘this’ singles out a unique object in a particular context, we should be able to use it to give that object a proper name; but then the name will be precise, contrary to Russell’s claim. The best Russell can do may be to find a context in which the application of ‘this’ and ‘is’ is not *relevantly* unclear, but ‘This is red’ expresses an essentially doubtful judgement. For example, different candidates for the referent of ‘this’ may all be the same in colour. The vagueness of ‘red’ can then be blamed. Although arguments of this kind are unlikely to be very rigorous, they may still be very plausible.

How do these considerations affect Russell’s argument for the vagueness of logical words? Suppose that ‘“This is a man” is true’ expresses an essentially doubtful judgement in a given context; can the vagueness of ‘true’ be blamed? Russell may seem to have undermined his positive answer by deriving the vagueness of ‘“This is a man” is true’ from that of ‘This is a man’. Does this not show that the subject of ‘“This is a man” is true’ rather than its predicate is to blame for its vagueness? However, it must be recalled that ‘This is a man’ occurs in ‘“This is a man” is true’ only as part of a quotation. What matters is whether the quotation ‘“This is a man”’ is vague, not whether the sentence ‘This is a man’ is. This is a delicate issue. On the one hand, we can think of the quotation as like a name of an English sentence. As such, it is relevantly clear *which* English sentence it is a name of. A name of a vague sentence need not itself be vague. On this view, the vagueness of ‘“This is a man” is true’ is to be blamed on the vagueness of ‘true’. On the other hand, we can think of the quotation as a name for a precise proposition. Its reference will then be indeterminate between various precise propositions, so the quotation will be relevantly vague. On this view, the vagueness of ‘“This is a man” is true’ is not to be blamed on the vagueness of ‘true’. If ‘true’ is applied to vague sentences, it is vague; if it is applied to precise propositions, it is precise. Russell claimed that 'true' and 'false' can have precise meanings only when they are applied to precise representations; the analysis confirms his claim.

What is much more dubious is Russell's claim that the vagueness of 'true' and 'false' infects all logical words. For example, if it is clear in a given context whether '*P*' and '*Q*' separately are true, it will also be clear whether '*P* or *Q*' is true. The disjunction expresses an essentially doubtful judgement only when at least one of its disjuncts does. Since 'or' itself is not to blame, it has not been shown to be vague.[48] Thus Russell has not succeeded in showing that all words are vague, still less that all language is vague. What he has done is to make it plausible that most words have blurred boundaries.

(3) *There is higher-order vagueness.* Some failures of definition can be described in relevantly precise terms. For example, if we are talking about integers, we can say, precisely enough, that 'half of' is well-defined for even numbers but not odd ones. But if the vagueness of a term can only be properly described in terms that are themselves vague, then it is said to have second-order vagueness. If the terms used to describe the vagueness themselves have second-order vagueness, the original term has third-order vagueness, and so on. According to Russell, a vague word has a region of definite application, a region of definite non-application and a penumbra in which it neither definitely applies nor definitely fails to apply. He goes further, saying that the penumbra itself shades off, lacking a sharp boundary.[49]Thus he acknowledges vagueness in the notion of definite application with which he characterizes vagueness. This is a form of second-order vagueness. There are borderline cases of 'red', but there are also shades on the borderline between the definite cases of 'red' and its borderline cases, and others on the border-line between the borderline cases and the definite cases of 'non-red'.

Russell points out that it is second-order vagueness which blocks the attempt to introduce precision by stipulating that a word should not apply in its penumbra. If the penumbra had sharp edges, the result would be a new precise sense. For example, we can introduce a new word 'dred' for definite cases of 'red'. If 'red' had first-order but not second-order vagueness, 'dred' would be precise. On Russell's view, 'dred' can have borderline cases too, for 'red' has second-order vagueness. The argument iterates. We can introduce a new word ‘ddred’ for definite cases of ‘dred’. If ‘red’ had second-order but not third-order vagueness, ‘ddred’ would be precise. But ‘ddred’ can have borderline cases too, for ‘red’ has third-order vagueness. It would be in the spirit of Russell’s position to attribute all orders of vagueness to all words.

Higher-order vagueness is also needed for Russell’s argument that ‘true’ and ‘false’ are vague, although he does not make the connection explicit. Consider, for example, the vague sentence ‘Ebenezer is bald’. Russell seems to assume that if Ebenezer is in the penumbra of ‘bald’, then the sentence is neither true nor false, for he asserts ‘There are men of whom it is not true to say that they must either be bald or not bald’.[50] Now if ‘Ebenezer is bald’ had only first-order vagueness, we could divide all cases into three sharply distinguished groups: those to which ‘bald’ definitely applies, those to which it definitely does not apply, and the penumbral cases. ‘Ebenezer is bald’ would be definitely true if Ebenezer belonged to the first group, definitely false if he belonged to the second, and definitely neither true nor false if he belonged to the third. In each case, the sentence would be either definitely true or definitely not true, and either definitely false or definitely not false. It would not be a borderline case of ‘true’, or of ‘false’. When Russell says that the application of ‘true’ and ‘false’ to vague representations is itself vague, he is assuming that the representations have second-order vagueness. If ‘“Ebenezer is bald” is true’ has nth-order vagueness, then ‘Ebenezer is bald’ has $(n + 1)$th-order vagueness.[51]

(4) *Vagueness invalidates classical logic.* Russell denies the validity of the law of excluded middle for vague languages. As already noted, he holds that if Ebenezer is in the penumbra of ‘bald’, then ‘Ebenezer is bald or Ebenezer is not bald’ is not true. His view seems to be that if the disjunction were true, then one of its disjuncts would be true, and that neither ‘Ebenezer is bald’ nor its negation is true in a borderline case.

Russell uses the supposed failure of excluded middle as a diagnosis of the sorites paradox that, if a man goes bald, ‘there must have been one hair the loss of which converted him into a bald man’. Unfortunately, he does not explain how sorites reasoning relies on the law of excluded middle. It is not used by the Stoic versions considered in Section 1.3, for they do not employ disjunctions. Perhaps Russell meant only that sorites paradoxes and the failure of excluded middle have a common cause in the occurrence of borderline cases. He does not say that excluded middle is the only law of traditional logic invalidated by vagueness. His position quite generally is that 'All traditional logic habitually assumes that precise symbols are being employed'.[52] His response is not to seek some non-traditional system of logic better adapted to vagueness; he would have taken that to miss the point that no system of logic worth the name is reliable when applied to vague symbols. Rather, he tries to show how we can use logic in spite of its inapplicability to our actual language.

Although our language is vague, we can conceive a precise one; we have the concept of imprecision only because we have the concept of precision. If our non-logical words were precise, our logical words would thereby become precise too, and 'We can, in fact, see precisely what they would mean if our symbolism were precise'.[53] A logically perfect language is imaginable, and could in principle exist. Thus if we can infer a conclusion about the nature of the world from the hypothesis of a logically perfect language, that conclusion is actually true, for the nature of the world does not depend on whether there is in fact a logically perfect language. Unfortunately, Russell does not say what standards of validity apply to our inference. Although its premise is *about* a logically perfect language, the inference is made *in* our logically imperfect language, to which classical logic does not apply. How can we see 'precisely' what logical words would mean if our symbolism were precise, unless we form precise representations of those hypothetical meanings? Yet, by hypothesis, we cannot form precise representations.

The problem is general. We need and want to reason, but the only propositions available to us are vague, for Russell assumes that our concepts as well as our words are vague. We cannot rely on classical logic, if Russell is right, because it will lead us into sorites paradoxes. What should we do instead? Like Frege, Russell does not tell us.[54]

(5) *Vagueness is not generality.* Russell initially makes a clear distinction between vagueness and generality. He uses the latter term more widely than Peirce, for unspecificity: 'A proposition involving a general concept – e.g. "This is a man" – will be verified by a number of facts, such as "This" being Brown or Jones or Robinson. But if "man" were a precise idea, the set of possible facts that would verify "this is a

man" would be quite definite.'[55] Russell's point is that, although "This is a man" is in fact vague, it could in principle have been both general and precise, for there is no contradiction in supposing that the set of possible facts that would verify it is sharply defined but contains more than one member. Thus generality does not entail vagueness.

When Russell speaks of a fact verifying a proposition, the relation he means is not an epistemological one. The fact is what makes the proposition true; it need not enable an observer to know that the proposition is true. A possible fact that would verify 'This is a man' is a way in which 'This is a man' could be true. Of course, the notion of the number of ways in which a proposition could be true is rather obscure, and Russell's theory of facts is far from dispelling all its obscurity. However, we can probably make enough sense of the notion for present purposes. For example, the proposition 'John is in Australia or Canada' might be true because John is in Australia, and it might be true because he is in Canada, so there are at least two possible facts that would verify it.

Unfortunately, Russell soon backslides from his distinction between vagueness and generality. A few pages later, in his analysis of vagueness, he explains 'A belief is *precise* when only one fact would verify it'.[56] By the previous definition of generality, a belief is general when not only one fact would verify it. Thus generality is the contradictory of precision. But Russell also treats vagueness as the contradictory of precision. He therefore confuses generality with vagueness.

Some part of Russell's account should be dropped. It is his later definition of precision that seems most at fault. It classifies the proposition 'John is in Australia or Canada' as not precise, and therefore as vague, merely because more than one fact would verify it. It does so irrespective of whether the disjuncts 'John is in Australia' and 'John is in Canada' are vague or precise. More generally, imagine that 'P_1' and 'P_2' are precise propositions such that 'P_1' would be verified by the possible fact f_1 and 'P_2' would be verified by the distinct possible fact f_2. In a logically perfect language one should be able to form the disjunction 'P_1 or P_2', for it has no penumbra and satisfies the law of excluded middle.

Yet Russell's definition counts 'P_1 or P_2' as not precise, for there are two possible facts, f_1 and f_2, that would verify it. That is absurd. The earlier distinction between vagueness and generality is much closer to the mark.[57]

(6) *Accuracy is isomorphism.* Russell's account of precision is easier to understand if one begins with his account of accuracy. They are not the same thing, as he points out. A belief 'is *accurate* when it is both precise and true', so a false precise belief is inaccurate but not vague.[58] A large-scale map drawn by a cartographer with an over-vivid imagination may be both inaccurate and precise. Russell's aim is a definition of vagueness applicable not just to language but to representation of all kinds. For him, maps and photographs display the fundamental nature of representation more clearly than words do. This emphasis helps to explain his definitions of accuracy and precision.

Russell defines accurate representation as isomorphism: the representing system should have exactly the same structure as the represented system. Accurate representations mirror form, not content. The concept of isomorphism is taken from mathematics. The representing system is conceived as a set X of objects, with various relations $R_1, \ldots, R_m$ which its members can have to each other. For example, X might be the set of points on a map of Africa, R_1 the relation of being to the left of, and R_2 the relation of being coloured darker than. The represented system likewise is a set Y of objects, with various relations $S_1, \ldots, S_\mathrm{n}$ which members of Y can have to each other. Thus Y might be the set of points on the surface of Africa itself, S_1 the relation of being to the west of, and S_2 the relation of being higher than. On Russell's account, the system of X and $R_1, \ldots, R_m$ accurately represents the system of Y and $S_1, \ldots, S_n$ if and only if there is a relation, correspondence, with the following features:

(i) Each member of X corresponds only to a member of Y, each R_i corresponds only to an S_j, and nothing else corresponds to anything.
(ii) Each member of X corresponds to at least one object and each R_i corresponds to at least one relation.
(iii) To each member of Y corresponds at least one object and to each S_j corresponds at least one relation.

(iv) Each member of X corresponds to at most one object and each R_i corresponds to at most one relation.
(v) To each member of Y corresponds at most one object and to each *Sj* corresponds at most one relation.
(vi) If *Ri* corresponds to *Sj*, and the members $x_1, \ldots, x_k$ of X correspond to the members $y_1, \ldots, y_k$ of Y respectively then R_i relates $x_1, \ldots, x_k$ (in that order) if and only if Sj relates $y_1, \ldots, y_k$ (in that order).

The word 'correspond' is used only for convenience; what matters is only whether some relation or other satisfies (i)–(vi). The first five clauses say that correspondence is a one–one correlation of the members of the systems; the last clause says that it also preserves relationships. Together, they say that correspondence is an isomorphism between the systems.[59]

Consider, for example, a relation that projects each point on the map of Africa to a point on the surface of Africa itself, and associates being to the left of and being coloured darker than (as relations between points on the map) with being to the west of and being higher than (as relations between points on the surface of Africa) respectively. Suppose that each point on the map is projected onto exactly one point on the surface of Africa ((ii) and (iv)), onto each point on the surface of Africa is projected exactly one point on the map ((iii) and (v)), one point on the map is to the left of another if and only if the former is projected onto a point to the west of the point onto which the latter is projected, and one point on the map is coloured darker than another if and only if the former is projected onto a point higher than the point onto which the latter is projected ((vi)). Then the system consisting of the points on the map with the relations of being to the left of and being coloured darker than accurately represents the system consisting of the points on the surface of Africa with the relations of being to the west of and being higher than. Speaking loosely, we may say that the map accurately represents Africa. However, although Russell calls a map a representing system, he does not specify what its members are, or what relations between them are to be constituents of the system. Yet many different systems can be abstracted from the same map; there is no such thing as *the* structure of a map.[60]For example, we could consider lines rather than points on the map, and the relation of crossing. The question of accurate representation makes sense only with respect to a choice of representing and represented systems.

Russell holds that no actual map is perfectly accurate. A dot on the map is projected onto a large area on the ground. The projection is not one–one but one–many, violating clause (iv). The smaller the scale of the map, the worse the violation of (iv). Russell's conflation of generality and vagueness is again at work. What he calls the inaccuracy of small-scale maps would usually be regarded as their uninformativeness. A shortage of information is not the same thing as misinformation.

Russell's definition of accurate representation does not require that someone intend the representing system to represent the represented system; it does not even require a causal connection between the two systems. He thereby leaves open the possibility that what we should regard as a very inaccurate map of Africa is accurate in his sense, because some complex unintended and accidental relation between points on the map and points in Africa, quite other than the usual projection, happens to be an isomorphism. The map might also turn out to be an accurate representation of someone's digestive system. It seems wrong to say that one system represents another in virtue of such an accidental relation, and equally wrong to say that it is accurate in virtue of a non-representing relation.

The official definition has other strange consequences. For example, since any system is isomorphic with itself, it accurately represents itself. Similarly, since being isomorphic to is a symmetric relation, an accurate map of Africa is accurately represented by Africa. Again, for any system of a set X with relations $R_1, \ldots, R_m$, one could choose any set Y one liked with the same number of elements as X, and use a one–one correspondence between X and Y to *define* relations $S_1, \ldots, S_m$ corresponding to $R_1, \ldots, R_m$; since this correspondence will be an isomorphism from the original system to the artificial one, the former accurately represents the latter. For example, any square kilometre of ground will accurately represent any other square kilometre of ground for a suitable choice of relations on the latter.

Could Russell keep his official definition by maintaining that in a broad sense any isomorphism is *ipso facto* a representing relation? His concept of 'cognitive or mechanical' representation was certainly broad, and he hoped to explain it in naturalistic terms. However, the paper does not define representation in general, which can be inaccurate. It defines only accurate representation. In order to make a theoretical case for classifying every isomorphism as a representing relation, Russell would need to provide an independent account of representation and (without discrediting it) show it to have the strange consequence that every isomorphism is a representing relation. Such an attempt seems unlikely to succeed.

We can finesse the issue by using the notion of a representing relation without attempting to define it. We can then rationally reconstruct Russell's views thus:

(a) If a relation is representing and an isomorphism then it is accurate.[61]
(b) If a relation is representing and not an isomorphism then it is inaccurate.
(c) If a relation is not representing then it is neither accurate nor inaccurate.
(d) One system accurately represents another if and only if the former has an accurate relation to the latter.
(e) One system inaccurately represents another if and only if the former has an inaccurate relation to the latter.[62]

These theses are neutral with respect to the claim that every isomorphism is an accurate representing relation. If Russell were willing to deny this claim, he could escape all the counterintuitive results mentioned above, simply by denying that the relations in question are representing.

(7) *Precision is one–one correlation.* According to Russell, 'accurate' means something like 'precise and true'. Thus if we can find a clause in his full definition of 'accurate' that generalizes 'true', its deletion should give his definition of 'precise'. Russell does this elliptically, saying 'a representation is *vague* when the relation of the representing system to the represented system is not one–one, but one–many'; for example, one point on the map represents many points on the ground.[63] Clauses (i)–(v) above suffice to make correspondence a one–one relation. Russell has deleted clause (vi), whose addition made correspondence an isomorphism by having the representing system preserve the relationships in the represented system.

In effect, Russell treats (vi) as a generalized truth requirement. Perhaps we should not use the word 'true' when it comes to maps, but the accuracy of a map does seem to involve a component very like the truth of a proposition. If the relation R_i represents the relation Sj and the objects $x_1, . . . , x_k$ represent the objects $y_1,. . . , y_k$ respectively, one could say that it is true for R_i to relate $x_1, . . . , x_k$ (in that order) if and only if Sj relates $y_1, . . . , y_k$ (in that order). This is a simple version of the correspondence theory of truth. If correspondence is a representing relation, clause (vi) then says that the relations of the representing system relate its elements in all and only the true ways. Within the expressive limits of the system, (vi) demands the whole truth as well as nothing but the truth.

Russell's account of precision and vagueness can be rationally reconstructed in parallel with his account of accuracy and inaccuracy:

(a′) If a relation is representing and a one–one correlation then it is precise.[64]
(b′) If a relation is representing and not a one–one correlation then it is vague.
(c′) If a relation is not representing then it is neither precise nor vague.
(d′) One system precisely represents another if and only if the former has a precise relation to the latter.
(e′) One system vaguely represents another if and only if the former has a vague relation to the latter.[65]

Thus one–one correlation is not sufficient for precision; by (c′), the correlation must also be a representing relation. The account escapes the implication that any system precisely represents any other system with the same number of elements.

Unfortunately, Russell does not abide by a neat distinction between accuracy and precision. He calls languages accurate, although they cannot be true or false. He does not apply his definitions straightforwardly to propositions in words, although they are the subject of the original distinction between accuracy and precision. Nevertheless, (a′)–(e′) do develop a strand in his thinking.

Russell's definitions are not equivalent to our ordinary concepts of precision and vagueness. The abstractness and sketchiness of his account makes the gap hard to measure, but a sense of its size can be gained from the linguistic case.

(8) *Meaning is a special case of representing.* Russell calls words, sentences and languages vague. However, he does not treat words and sentences as vaguely representing systems whose elements are not correlated one–one with elements of the represented system. The constituents of a word (such as syllables) do not usually represent at all, and although constituents of a sentence (such as words) do represent, Russell locates its vagueness in the representation of several items (possible facts) by the sentence as a whole. Thus words and sentences are treated as elements of a representing system which is the whole language, rather than as representing systems in their own right. This involves an extension of terminology, since 'vague' and 'precise' have so far been defined only for representing systems, not for their elements. When a system is vague because one of its elements represents more or less than one item, it is natural enough to call that element vague too, and these are the cases Russell has in mind. A violation of clause (ii) or (iv), unlike a violation of clause (iii) or (v), can be blamed on a single representing item.

Russell takes meaning as the representing relation for a language. He regards the meaning of a word as the object, property or relation it means. His discussion presupposes something like the theory of meaning expounded in his *Lectures on Logical Atomism*, where proper names mean particular objects, verbs mean properties and relations, and so on. In those lectures he described a logically perfect language, free of all vagueness, in terms close to those later applied to an accurate language: 'In a logically perfect language, there will be one word and no more for every simple object'.[66] The same idea appears in the later paper: 'In an accurate language, meaning would be a one–one relation; no word would have two meanings, and no two words would have the same meaning'. Both ambiguity (violating clause (iv) and synonymy (violating clause (v)) are treated as obstacles to accuracy, and indeed to precision. Synonymy is 'easily avoided', but ambiguity endemic; 'The fact that meaning is a one–many relation is the precise statement of the fact that all language is more or less vague'.[67]

Ambiguity and vagueness seem to be different phenomena. A word may have one vague sense, or two precise ones. The various meanings of a vague word are presumably supposed to be the different objects, properties or relations picked out by different stipulations of sharp boundaries, but Russell does not elaborate.[68]

Linguistic isomorphism may fail in a different way. All words are individual objects, and some mean individual objects, but some do not: 'the word 'precedes', though it means a relation, is not a relation'.[69] This violates clause (i) (an S_j is represented by something other than an R_i). Order is restored in the sentence 'This precedes that', for it represents the possible fact that this, that and the relation of precedence stand in a certain relation by the fact that 'this', 'that' and 'precedes' stand in a certain syntactic relation in the sentence. At this point Russell seems to treat the sentence rather than the language as the representing system, contrary to his practice elsewhere.

Russell's view is at its strangest when he applies it to sentences. He treats a proposition as meaning the possible facts that would verify it. Thus, as already noted, it is vague if it could be made true by more than one possible fact. According to Russell, any sentence whose predicate expresses a general concept could be made true by more than one fact. 'This is a man' could be verified by this being Brown or Jones or Robinson. So any such sentence is vague. Although a more careful application of his theory of facts might have enabled him to avoid some consequences of this kind, his official theory of vagueness is worse than an over-simplification; it radically misconstrues the phenomenon.[70]

(9) *Precision diminishes probability.* The precise assertions of science are more informative than the vague assertions of common sense. They are therefore more likely to be false, but more useful if true. Russell contrasts the assertion that a man is tall with the assertion that his height is between 6 ft 2 in. and 6 ft 3 in. Here too specificity has been conflated with sharpness of definition. The latter belief is no more sharply defined than the assertion that the man's height is more than 2 in., but this assertion is much less informative than the assertion that he is tall. The best one can say is that no vague assertion is *maximally* informative. Vague assertions may also be harder to prove false than precise ones, other things being equal, but since an assertion is vague if and only if its negation is, it is equally hard to prove vague assertions true. This is no doubt part of the value of precision.

(10) *Vagueness is a natural phenomenon.* Although Russell confines vagueness to representations, he regards it as a natural phenomenon, because he regards representation as a natural phenomenon. He traces vagueness to what he calls a law of physics: 'the appearances of a thing at different places are less and less differentiated as we get further away from the thing'.[71] The appearances of a thing are its publicly observable physical effects. As they spread outwards, information is lost; different close-up appearances give rise to the same distant appearance, so the latter is vague as a representation of the former. In the case of perception, Russell treats our sensations as appearances of their stimuli. Sensations caused by different stimuli are identical, or at least indiscriminable – Russell is not sure which, but hopes that quantum physics will eventually settle the matter. He holds this feature of our physiology responsible for the ineliminable vagueness of our knowledge, including its higher-order vagueness.

Russell's physical explanation of vagueness again confuses it with unspecificity. Different scenes may give rise to the same photograph, and in that sense the information in the photograph about the original scene is unspecific. But if there is a sharp line between the scenes that would give rise to exactly similar photographs and those that would give rise to different ones, then the information in the photograph may still be sharply defined. It 'says' that the original scene belonged to a particular sharply defined set. The information is unspecific only because the set contains more than one member. This is a point at which quantum physics might have helped Russell. If the relation between scene and photograph is indeterministic, so that an individual scene gives rise to different photographs with various probabilities, an individual photograph does not 'say' that the original scene belonged to a particular sharply defined set. There would be a penumbra of scenes which could have resulted in that photograph, but which would more probably have resulted in a different one. They could be regarded as neither definitely permitted nor definitely excluded by the information it contains.

A similar account might be given in the case of perception, with a probabilistic connection between stimulus and sensation. The mere loss of information from stimulus to sensation does not explain first-order vagueness any more than it does for photographs, let alone the higher-order vagueness which Russell offers to explain. What he needs is something like a threshold of discrimination. Two stimuli whose difference is below the threshold cannot be discriminated. Since many indiscriminable differences can add up to a discriminable difference, one can have a series of stimuli each indiscriminable from its successor, of which the first member is discriminable from the last. Indiscriminability is a non-transitive relation. We are unable to make sharp distinctions in the continuum of stimuli because any line through it separates stimuli whose difference is below the threshold of discrimination. The non-transitivity of indiscriminability might in turn be explained by the probabilistic model. The greater the difference between two stimuli, the more likely they are to evoke different reactions; they are discriminable only if they evoke different reactions with a sufficiently high probability, which defines the threshold of discrimination. Russell did later connect the non-transitivity of indiscriminability with our inability to make sharp perceptual distinctions.[72] Unfortunately, he never firmly separated the phenomenon of blurred boundaries from that of unspecificity.

Chapter 3

The rehabilitation of vagueness

3.1 VAGUENESS AND ORDINARY LANGUAGE

Russell's paper of 1923 on vagueness raised many of the issues central to subsequent debate. Its effects, however, were delayed. Nothing of significance, and little of insignificance, was published on vagueness between Russell's work and that of Max Black in 1937.[1]The latter provoked several responses, and the topic has remained alive, if not always kicking, ever since. Although chance doubtless plays as great a role in the history of philosophy as in the history of everything else, it is not fanciful to suppose that the time was riper for vagueness in 1937 than in 1923. When Russell wrote, philosophy was concerned with language primarily as the medium of science. That could include a wide range of uses, from reporting observed data to formulating mathematical theories about unobservable entities. But it encouraged the use of 'vague' as a dustbin category, into which one dumped any failure to meet the ideal of precision, without prying too closely. To vary the metaphor, vagueness was the noise through which a signal had to be discerned. By 1937, philosophers had begun to find ordinary language interesting in its own right. The irreducible variety of its uses was taken to be one of its chief characteristics, and the indiscriminate assimilation of one kind of use to another diagnosed as the origin of many philosophical puzzles. Scientific uses of language were losing their privileged status; for most other uses, the ideal of precision looked irrelevant or damaging. Vagueness is a precondition of the flexibility of ordinary language. This thought converged with an older pragmatist idea, that too much precision is a bad thing even in scientific language, restricting its adaptability to new evidence.

The ground might seem to have been cleared for the investigation of vagueness as a positive feature of ordinary language, and indeed something of the kind appeared. However, ordinary language provided a new method as well as a new topic. The philosophical investigation of an area was to consist in the description of the ordinary uses of the vocabulary appropriate to that area. Thus vagueness would be studied not by constructing theories about it but by describing the everyday use of 'vague' and related words. Wittgenstein, having rejected Tractarian exactness, discussed the use of 'exact'. The Oxford philosopher J.L. Austin contrasted the uses of 'exactly', 'precisely' and 'accurately': 'If I measure a banana with a ruler, I may find it to be precisely 5 5/8 inches long. If I measure my ruler with bananas, I may find it to be exactly six bananas long, though I couldn't claim any great precision for my method of measurement'.[2] Austin's account brought out the variety of features covered by 'vague' in different contexts: roughness, ambiguity, imprecision, lack of detail, generality, inaccuracy, incompleteness. As he says, 'vague' is vague.[3] Its everyday meaning is indeed so diffuse that it can be the object only of the most desultory investigation.

To describe the uses of 'vague' is not to describe the vagueness of vague words. It is not enough to know that 'heap' is vague; in what respects is it vague, and to what degree? An answer to the question is part of a description of the ordinary uses of the word 'heap'. One might expect answers to be given within a standard framework, so that the vagueness of one word can be compared with that of another. The framework would be articulated in a fixed vocabulary, whose elements would tend to evolve senses more precise and specific than their ordinary ones, so that different dimensions of vagueness could be kept separate. Philosophy, like every other discipline, needs and is entitled to its own technical terms. That is what 'vague' has since become: it demarcates a definite problem only because it has been artificially but legitimately restricted, within the philosophical profession, to the case of blurred boundaries.

A standard framework for description is an incipient theory; it embodies a view of the important dimensions of the phenomena to be described. Since Wittgenstein and Austin were notoriously suspicious of philosophical theory, they inhibited theory-making even of this mild kind.

Of course, many philosophers of the period escaped their influence. Austin himself permitted philosophical theories, if they were not premature; it was just that he put the age of maturity so late. Wittgenstein held that philosophical theories were symptoms of philosophical puzzlement, not answers to it, but that was itself one of his philosophical theories. His work was always driven by theoretical concerns. This applies in particular to his account of family resemblance terms, his specific contribution to the study of vagueness, as it does to Friedrich Waismann's similar notion of open texture, developed under Wittgenstein's influence. However, theory does not flourish when it must be done on the quiet. It needs to be kept in the open, where it can be properly criticized.

Theory needs more than the right to speak its name; it also needs problems. In this way, too, exclusive preoccupation with the actual use of ordinary language was almost as bad for the study of vagueness as exclusive preoccupation with the possible use of a logically perfect language, by starving it of problems. The matter of vagueness gets its urgency from sorites paradoxes. They cannot arise in a logically perfect language, but when they do arise in ordinary language, they can usually be shrugged off. For practical purposes, the remark 'You have to draw the line somewhere; it doesn't matter exactly where' deals quite adequately with most sorites paradoxes. More is required, of course, when a slippery slope is of moral concern, as in the case of abortion or euthanasia, but then the importance of the example eclipses the need for a general account (hard cases make bad law). Only the standards of formal logic, applied to ordinary language, demand something more: a precise and general answer to the question 'What forms of inference are valid in a vague language?', underwritten by an appropriate theory of meaning for such a language. Most work on vagueness in the period 1925–65 suffers from a certain blandness because it does not take sorites paradoxes seriously. The felt need to make sense of them gives more recent work its flavour, however unpleasant. That need is felt by philosophers who take an ordinary language as a model of what is to be understood but a logically perfect one as a model of what it is to understand. Such a combination became more common in the 1960s, with the linguistics of Noam Chomsky as one precedent. The work to be discussed in this chapter has other preoccupations.[4]

3.2 THE BLACK–HEMPEL DEBATE

The earliest attempt to construct a framework for the systematic description of vagueness in natural language is probably Black's 1937 paper 'Vagueness: an exercise in logical analysis'. Carl Hempel published a reply to Black in 1939. The debate between them raised fundamental questions about the nature of vagueness.

Russell's work on vagueness provides the intellectual context of Black's paper. However, he rejects both Russell's analysis of vagueness as one–many representation and his blithe eviction of all vague language, and therefore all actual language, from the domain of logic. Black distinguishes vagueness from both generality and ambiguity. Vagueness consists in 'the existence of objects concerning which it is intrinsically impossible to say either that the symbol in question does, or does not, apply'.[5]Black means that neither thing can be said because neither is true, not just because neither can be known. If vagueness is blurring of boundaries, generality is breadth in the area enclosed; Russell's analysis is criticized for confusing the two. Ambiguity is another phenomenon again; one can resolve it, as one cannot resolve vagueness, by supplying an alternative word.[6] As for Russell's claim that logic does not apply to vague expressions, Black takes it just to evade the responsibility of systematically describing natural languages. His aim is to analyse vagueness as a positive phenomenon, an adaptation to our need not to clutter up our medium of communication with irrelevant precision, not a mere defect.

Black sharpens the challenge by arguing elaborately that incoherence results if classical negation is applied to vague statements. The underlying idea is simple. To make a statement is just to exclude certain cases.[7] Since a statement '*P*' does not exclude its borderline cases, '*P*' is correctly assertible and true in such cases. Equally, they are borderline cases for 'Not *P*', which is therefore correctly assertible and true for the same reason. Thus both '*P*' and 'Not *P*' are true in borderline cases. Moreover, since 'Not *P*' is classically defined as true if and only if '*P*' is false, '*P*' should be false and therefore excluded in borderline cases, contrary to what was said before.

One may question Black's assumption that to assert '*P*' is just to exclude certain cases, and not positively to include others, when borderline cases are in question. He is not merely attributing the assumption to classical semantics, for the latter has already been ruled out by his initial description of borderline cases; no elaborate argument would be needed. The target of his attack is rather the view that, even if '*P*' is neither true nor false, its negation can still have the classical property of being true if and only if '*P*' is false: but proponents of such a view should not grant Black's assumption that to assert '*P*' is merely to exclude its falsity condition. Nevertheless, vague statements do violate classical semantics, if Black is right that they are neither true nor false in borderline cases.

According to Black, vague language runs into trouble by trying to express matters of degree without being explicit about degrees. 'He is tall' treats tallness as though it were an all-or-nothing matter. The proposed remedy is a notation in which degrees are explicitly registered.

Black begins his construction by arguing that vagueness is 'objective' rather than 'subjective': the willingness of a speaker to assert '*x* is *L*' depends on what the object *x* is, not just on who is making the assertion, even when *L* is vague. Whether speakers of English are willing to call a man 'tall' depends on his height, not just on their mood. In a borderline case, there will be variation between different speakers at the same time and in the same speaker at different times, but there will also be statistical regularities in the variation. It is on these that Black bases his construction.

The crucial notion is *the consistency of application of L to x*, where *L* is a predicate and *x* an object. It is defined as the ratio of the number of cases in which *L* is judged to apply to *x* to the number of cases in which *L* is judged not to apply to *x*. More precisely, Black considers situations in which a user of the language is forced to decide between the two judgements, and defines the consistency of application as the limiting value of the ratio as the number of situations increases indefinitely and all users of the language are taken into account. Black expects the objectivity of vagueness to ensure that the ratio does converge to a limiting value. This value should be very large if *L* clearly applies to *x*, near 1 if *x* is a borderline case of *L* and near 0 if *L* clearly does not apply to *x*. For example, we may suppose that a given height determines the consistency of application of 'tall' to anyone of that height, and that it is virtually 0 for heights under 5 ft, rises smoothly thereafter and is huge for heights over 7 ft.[8]Consistency of application can be plotted against height on a graph. Black calls the resulting curve the *consistency profile* for 'tall'. In contrast, the consistency profile for a more precise term such as 'at least 6 ft tall' will rise more steeply; we may suppose that its consistency of application is virtually 0 for heights under 5 ft 9 in. and huge for heights over 6 ft 3 in. The more precise a predicate, the more its consistency profile resembles a right angle. Black makes the dubious assumption that objects are ranged along a relevant quantifiable dimension (like height) for any predicate *L*, and uses the shape of the consistency profile as a criterion for the vagueness or precision of *L*.[9] The mathematical detail should not disguise the fundamental point that the vagueness of a symbol is being equated with disagreement in its application.

Black says that the vague symbol *L* can be replaced by a new symbol L' which differs from *L* in being explicit about degrees of application. In the old language one said '*x* is *L*'; in the new language one says 'L' is present in x with degree *c*', which is to be equivalent to 'The consistency of application of *L* to *x* is *c*'. Black calls L' the analysis of *L*.[10] However, sentences of the old language cannot be translated into sentences of the new one, for '*x* is *L*' implies nothing about consistencies of application. It is not equivalent to 'All the users of the language apply *L* to *x*' or 'More than half the users of the language apply *L* to *x*', for example. Although Black equates correctness in the use of a symbol with 'statistical conformity with the behaviour of a certain group of users', even he denies that the ability to use it correctly involves extensive statistical knowledge of the behaviour of other users, on pain of vicious circularity.[11]Moreover, the intended precision of the new sentences does not match the vagueness of the old. L' is not a genuine analysis of *L*. Rather, *L* and L' are supposed to reflect the presence or absence of the same feature, the former treating it vaguely as an all-or-nothing matter, the latter precisely as a matter of degree. The statement that a certain proportion of speakers call *x* 'tall' is made to double as the statement that *x* is tall to that degree. A poll of speakers supposedly defines the degree to which a non-linguistic feature is present. Black's meta-language for a vague object-language is at the same time to be a precise replacement for it.

Black writes as though the logic of the *new* language is non-classical, classical logic emerging as a special case when the consistency of application of the terms is either infinite or zero (effective unanimity). However, it could be precisely true (or false) that the consistency of application of L to x is c even when c is neither infinite nor zero. The notion of consistency of application may satisfy some special principles, just as the notion of comparative tallness satisfies the special principle that if x is taller than y and y is taller than z then x is taller than z. However, principles of this kind in no way violate classical logic; they merely add to it. Perhaps the old vague language violates classical logic, but the principles Black gives are for reasoning in the new language, not the old one. At most they are principles for reasoning *about* the non-classical semantics of the old language *in* the new one. Although Black's account is confused in both letter and spirit, several points can be discerned.

If L is precise, then either L applies to x or its negation applies to x; this is one form of excluded middle. If L is vague, then the consistency of application of its negation to x is still the reciprocal of the consistency of application of L to x, so either the former is at least 1 or the latter is; this is a kind of replacement principle.[12]

Black also defines notions of implication between predicates. If 'scarlet' and 'red' were precise, one could say that 'scarlet' implies 'red' in the sense that, for every object x, if 'scarlet' applies to x then 'red' applies to x. One can adapt this to the vagueness of the words by saying that, for every object x, the consistency of application of 'scarlet' to x is less than or equal to the consistency of application of 'red' to x. In such a case, where the consistency profile of M (e.g. 'scarlet') lies wholly under the consistency profile of L (e.g. 'red'), Black says that the field of L includes the field of M. Inclusion is transitive: if the field of L includes the field of M and the field of M includes the field of N, then the field of L includes the field of N. If the field of 'not L' includes the field of K, then Black says that the field of L excludes the field of K. For example, the field of 'red' excludes the field of 'green' if the consistency of application of 'not red' is never lower than the consistency of application of 'green', which will be the case if everyone willing to call an object 'green' is also willing to call it 'not red', or if the anomalous occasions on which it is called both 'green' and 'red' are balanced by others on which it is called both 'not green' and 'not red'.[13]

In many cases, the concept of inclusion must be loosened up to be of use. For example, suppose that *x* is called 'intelligent' on almost but not quite all occasions on which x is called 'wise', the exceptions not sufficing to define a subgroup of speakers who use the words in a non-standard sense. The consistency profile of 'wise' may then lie mainly under that of 'intelligent', rising above it at just a few points. Black then wants to say that the field of 'intelligent' includes that of 'wise' to some degree. He suggests that one could measure the degree to which the field of *L* fails to include that of *M* by a number, zero for perfect inclusion and positive otherwise. Unlike perfect inclusion, loose inclusion is not transitive, for small failures of inclusion can add up to large ones. However, one might be able to replace transitivity by the principle that the degree to which the field of *L* fails to include the field of *N* is not greater than the sum of the degree to which the field of *L* fails to include the field of *M* and the degree to which the field of *M* fails to include the field of *N*.[14] Thus in the inference from 'The field of *L* includes the field of *M*' and 'The field of *M* includes the field of *N*', the degree of error in the premises limits the degree of error in the conclusion. The inference thereby satisfies, so far as can be expected, the insight of the physicist Duhem, quoted by Black, that one needs inferences for which not only does the exact truth of the premises guarantee the exact truth of the conclusion (the usual definition of deductive validity), but also the *approximate* truth of the premises guarantees the *approximate* truth of the conclusion.

Black's account requires an appropriate concept of a user of a language, for an over-wide or over-narrow concept would yield mistaken attributions of vagueness or precision. If the consistency profile of 'red' took into account the judgements of the colour-blind, 'red' would come out as vaguer than it actually is. The problem is accentuated by the definition of degrees of application in terms of the dispositions of users of the language: it would then be circular to define a user as one disposed to apply predicates to objects in rough proportion to the degrees to which the former do in fact apply to the latter. However, Black argues that the circularity can be eliminated. Given a person and a predicate, one can construct a consistency profile by considering the person's dispositions to apply the predicate on different occasions. In the simplest case, a language community can then be defined as a group of people each with the same consistency profile, not contained in any larger such group. In practice, languages do not have clear boundaries; no two people have exactly the same consistency profiles, and languages fade into each other through series of intermediate dialects. One might say that a language community is a group of people each with *approximately* the same consistency profile, not contained in any larger such group. This depends on a standard for 'approximately the same', whose fixing is to some extent arbitrary. It loosens the concept of a language community, but need not reintroduce circularity.

Although Black speaks of defining 'user of a language', the concept he really needs is 'user of an expression'. In respect of the word 'red', the relevant community includes the hard-of-hearing with good eyesight but not the colour-blind with good hearing; in respect of 'melodious', the position is reversed.[15]

One of Black's most striking claims is that vagueness can be measured experimentally. He even reports an experiment supposed to illustrate the construction of a consistency profile. Eighty-three subjects were presented with a set of rectangles of varying lengths and breadths, and were asked to divide it into two subsets 'at what seems the most NATURAL place'.[16]Unsurprisingly, different subjects made the division at different places, but some places were more favoured than others. The subjects' task was not even of the kind involved in the formation of a consistency profile, for they were not given a predicate and asked to which of the rectangles they would apply it. They were told to classify, but not by what principle.[17] Black's experiment served only to discourage further experiments.

Black reads the failure of classical laws of logic in a vague language directly off the behaviour of its users. The heart of Hempel's case against Black is that this move confuses two levels of description. At one level, an observer can describe the users' behaviour without even understanding their language. At a more abstract level, the language can be described as a system governed by rules of syntax and semantics. The relation between the two levels is far from direct. Many aspects of linguistic behaviour are irrelevant to language as a system, and the rules of the system are not empirical generalizations about the behaviour, which often violates those rules. Hempel compares abstracting a language from behaviour to learning the rules of chess from watching it played. The observer will notice that 'before moving a chessman, a player will often frown thoughtfully', and that 'a player, when pronouncing the words "checkmate", displays in general more signs of pleasure than his partner', but these regularities are not rules of chess.[18] Now a consistency profile is a behavioural regularity, whereas the logic of a language depends on its syntax and semantics. Since the phenomena in which Black locates vagueness are just of the kind from which syntax and semantics abstract, he has not shown that vagueness makes any difference to logic. If one examines the actual use of language, one will observe standard principles of logic being violated in all sorts of ways: but the principles are not thereby refuted, any more than the rules of chess are refuted when incompetent or dishonest players violate them. Once one has abstracted enough away to talk about logic, one may already have abstracted away vagueness.

Hempel frames his argument within C.W. Morris's general theory of signs. For Morris, the use of a sign involves at least three factors: the sign itself, that to which it refers, and its users. A *strictly semiotical* term is one whose definition must advert to all three factors; in contrast, a *semantical* term can be defined by adverting simply to the first two factors, without mention of those who use the sign. Vagueness, as defined in terms of consistency profiles, is a strictly semiotical term, not a semantical one. The concept of logical validity is semantical. Black's vagueness is strictly irrelevant to logic.

Hempel considers the idea of introducing a semantical concept of vagueness, which would be relevant to logic. There are precedents for a strictly semiotical term having a semantical analogue: the strictly semiotical three-place relation 'the group *z* of speakers designates by the French term *x* the property *y*' is analogous to the semantical two-place relation 'the French term *x* designates the property *y*', in which reference to the third factor has been eliminated (henceforth, 'designate' will always be used in the semantical sense). Can the same be done for vagueness? Hempel does not deny that the strictly semiotical three-place relation 'the group *z* of speakers applies the term *x* to the object *y*' is analogous to the semantical two-place relation 'the term *x* applies to the object *y*'. Nor does he deny Black's contention that the strictly semiotical relation is *gradable*: it comes in degrees. The question is whether the semantical relation is gradable too.

According to Hempel, a predicate applies to an object if and only if the former designates a property possessed by the latter. 'Chaud' applies to the sun because it designates the property of being hot and the sun possesses that property. He takes it for granted that property possession is not itself gradable.[19] Thus, for predicates, if the semantical relation of application is gradable, it must be because designation is gradable. As for names and other singular terms, Hempel equates application with designation; if the former is gradable then so automatically is the latter. The question is now whether a language could employ a gradable designation relation. Hempel imagines a putative example. The noun 'sol' designates the sun to degree 0.7; the adjective 'cal' designates the property of being hot to degree 0.9. If the noun '*a*' designates the object *a* to degree α, and the adjective '*B*' designates the property *B* to degree β, then the sentence '*a* esti *B*' designates the state of affairs that a has *B* to degree $\alpha\beta$. For example, the sentence 'sol esti cal' designates the state of affairs that the sun is hot to degree 0.63. However, Hempel argues that such a language could not be translated into English. In particular, 'The sun is hot' is not an exact translation of 'sol esti cal'. Presumably, 'The sun is hot' designates the state of affairs that the sun is hot to no degree less than 1, and is therefore not an exact translation of any sentence which does designate that state of affairs to a degree less than 1. Since the language cannot be translated into English, 'it is not an interpreted language in the usual sense'.[20] Hempel therefore denies that designation is gradable. There is no semantical concept of vagueness, so vagueness does not invalidate classical logic.

In reply to Hempel, Black agreed that the concept of vagueness in the original paper was strictly semiotical, but disputed the argument against the possibility of an analogous semantical concept.[21] In particular, he denies that a language with a gradable designation relation cannot be translated into English. Since 'sol esti cal' designates to a high degree the state of affairs that the sun is hot, 'The sun is hot' is an adequate, albeit approximate, translation. Some parts of the language may resist even this kind of translation into English, but why should English not be extended by the terminology needed to accommodate them? A semantical concept of vagueness is both desirable and possible. Thus vagueness may yet demand a change of logic. Although we might try to retain classical logic by making our language more precise, that is not a promising strategy. While admitting that logic always abstracts from linguistic practice, Black suggests that classical logic involves an excessive degree of abstraction.

A curious feature of the debate is the willingness of both parties to assume that English does not itself have a gradable designation relation. If it has, then there is, quite trivially, no problem about the exact translatability of at least one language with a gradable designation relation into English. Is not the vagueness of English and other natural languages just what was supposed to grade the designation relation in the first place?

The source of the confusion may lie in the use of English to discuss its own semantics. How can 'hot' designate the property of being hot to any degree less than one? Yet 'hot' *is* vague. One may say that it designates the property of being hot to degree 1, but if properties are sharp, why should Black accept that 'hot' singles out a unique property? If he had to state its designation in a precise language, he might need to say that it designated different properties to various degrees. The precise language would contain no synonym of 'hot', just because 'hot' is vague, but it might still provide a clearer view of the relation between the word and the world. Thus Hempel's argument against the possibility of a gradable designation relation does not seem compelling. One may also endorse Black's objection that if we can learn a language with a gradable designation relation by the direct method then we shall not need to translate it into the vocabulary we now have.[22]

Hempel is more sympathetic to Black's account of vagueness when read as a strictly semiotical description. Like Peirce and Black, he equates vagueness with a kind of variation in use both between and within speakers. He argues that both kinds of vagueness are already present in the message '60 kg' as printed out by a weighing machine, and that all words – even logical ones – are ineradicably vague, not least because they are learned from particular instances of their use.[23]

Genuine vagueness must be distinguished from other kinds of variation in use. Hempel understates the problem. Even when liars and other deviants have been expelled from the speech community, its remaining members are not infallible. In defining 'consistency of application', Black gives as an exemplary test situation 'an engine driver on a foggy night [. . .] trying to decide whether the light in the signal box is really a red or a green light'.[24] If x is the green light, the occasions on which fog tricks the driver into judging x to be red will make the consistency of application of 'red' to x non-zero. Black's analysis counts this as vagueness in 'red', and x as to some degree a borderline case of 'red', even though x may be a paradigm of 'green'. The problem is comparatively limited for 'red', since red things usually look red and red-looking things usually are red, but Black's analysis is intended for all predicates, not just observational ones. For example, widespread disagreement on whether y is a spy makes y a borderline case of 'spy': yet y may be a guileless bystander or a master-spy. Black's statistical survey must be confined to situations in which the subject knows 'the full facts' – and is able to integrate this knowledge (a deluge of information can lead to mistakes). But knowledge of the full facts is not a behaviourally defined condition; indeed, it may not be well defined at all. If one tries to imagine in any detail what it would be like to carry through Black's programme, such problems multiply. Vagueness has certainly not been captured in behavioural terms.

Consistency of application has turned out to be a very vague notion: it is hard to agree with Black's claim that by speaking of it we can make our language more precise.[25] Hempel unfavourably contrasts Black's method for reducing vagueness with the scientist's. The latter replaces the vague 'x is hot' with 'the temperature of x in centigrades is t', which can be tested by the dilatation of a mercury thread. Black replaces 'x is hot' with 'the consistency of application of "hot" to x is c', which must be tested by a statistical survey of language use. If replacement rather than analysis is at issue, the scientist's candidate looks the more fruitful.[26]

Black modified his position in a paper of 1963. There is no mention of consistency profiles and the other statistical apparatus. Revision of logic is no longer proposed, even though the application of classical logic to sorites series is said to violate the intended looseness of our concepts, and borderline cases are described as those to which the law of excluded middle does not apply. Yet Black claims that such cases do not constitute exceptions to the law; they are 'simply irrelevant' to it.[27] Logic is not intended (by whom?) to apply to borderline cases. The applicability of logical rules is not itself specified by a rule; one must use one's judgement to decide.[28] There is no exact point at which they become inapplicable. Laws of logic are treated as rules of thumb, to be followed in the absence of good reason to the contrary. Black's attitude may be compared with Frege's. Both hold that the uncompromising application of logic in a vague language yields incoherence and that this is not a defect in logic. Black differs from Frege in holding that it is not a defect in language either: we live happily with compromises. He has lost interest in the possibility of a precise language in which the uncompromising application of logic would not lead to incoherence. The only convincing reply to such pessimism would be a positive and logically rigorous solution to the sorites paradoxes, showing them to involve untrue premises or fallacious inferences.

Black's later position tacitly concedes an important methodological point to Hempel. When logical puzzles exploit vagueness, they are to be solved not by the choice of some particular non-classical system of logic appropriate to vagueness, but by a better general understanding of the very indirect relation between logical systems and linguistic practice. If that is the specific moral of the Black–Hempel debate, it has not been widely learned.

The debate raises wider issues about the relation between meaning and use. Black attempted to reduce semantic descriptions to descriptions of use. Semantic descriptions are normative; they classify uses as correct or incorrect. Black's descriptions of use were non-normative. In effect, he was trying to reduce the normative to the non-normative. His attempt failed. Arguably, any such attempt must fail. The attempt to connect meaning and use is hard to sustain unless use is already described in normative terms, something later philosophers of ordinary language were willing to do. Their notion of 'what we say' was a notion of correct usage in our community, not a matter of statistics. To invoke non-normative descriptions of use in analysing vagueness, as in statistical accounts of degree of truth, is to misconceive the relation between meaning and use.

3.3 FAMILY RESEMBLANCES

The first part of Wittgenstein's *Philosophical Investigations* may be read as a sustained critique of his *Tractatus Logico-Philosophicus*, in particular of its claim that analysis must reveal everyday language to be, in spite of all appearances, logically perfect by Tractarian standards. One may say that Wittgenstein repudiates formal logic as a model of natural language, subject to two qualifications.

(a) No alternative model is proposed: what Wittgenstein attacks is the idea that anything less than a natural language is a philosophically adequate model of natural language. For nothing less has the diversity to permeate our lives in the way our language does, and this permeation is philosophically central. Wittgenstein does not deny that specific areas of natural language can usefully be modelled; indeed, he offers such models himself. But such a model will break down if applied too widely. It points not to a hidden uniformity, a way in which language must work, but to a way in which it *can* work; the model brings out differences between that area of language and others.

(b) Although the *Tractatus* did model natural language on formal logic, the model idealized the latter as well as the former. 'For the crystalline purity of logic was, of course, not *a result of investigation*: it was a requirement' (PI §107). The *Tractatus* imposed the requirement of crystalline purity on both formal logic and natural language. In logic, it is sometimes an appropriate ideal, doubtless never fully realized. Even in the best of actual machines, parts can bend, break off or melt (PI §193). The *Tractatus* made the double error of generalizing the normative requirement to all language and mistaking it for a description of what is necessarily the case.[29]

Our everyday language is in working order as it is, even though it does not meet Tractarian requirements. In particular, it does not meet the ideal of precision. But it would be inexact to say that what we say is always inexact, for 'exact' is itself a word of our everyday language, and answerable to its standards. What counts as 'exact' depends on the purpose in hand; people sometimes come to dinner exactly on the hour, whatever laboratory results may show. 'No *single* ideal of exactness has been laid down' (PI §88). The same goes for precision. It would be wrong to say that all language is vague, for something is vague only if it falls short of a standard set by the purpose in hand.[30] Nevertheless, there is a recognizable direction to the demands of the Tractarian ideal. In *Philosophical Investigations*, the model of family resemblances shows how a word can function successfully without meeting those demands.

Wittgenstein is arguing that the various phenomena properly called 'language' have no one thing in common, and that this does not make the word 'language' senseless or ambiguous (PI §65). He thereby undermines the *Tractatus* model at two levels. If linguistic phenomena have no one thing in common, then in particular they do not have what the *Tractatus* says they have in common. At the same time, the word 'language' itself functions incompatibly with the *Tractatus* model. Wittgenstein avoids the confusions which could easily result from the interaction of the two levels by developing his argument with the words 'game' and 'number' in place of the word 'language'.

Games have no one feature in common. Amusement, winning and losing, competition between players, skill, luck: each is missing from some games. The concept of a game cannot be analysed by means of a necessary and sufficient condition for x to be a game. Rather, games resemble and differ from each other as members of a family do. 'We see a complicated network of similarities overlapping and criss-crossing: sometimes overall similarities, sometimes similarities of detail' (PI §66). So too for the concept of number: numbers of different kinds (integral and fractional, rational and irrational, real and complex, finite and infinite) have no one feature in common. 'Why do we call something a "number"? Well, perhaps because it has a – direct – relationship with several things that have hitherto been called number; and this can be said to give it an indirect relationship to other things we call the same name' (PI §67).

The sense of a family resemblance term has a dynamic quality, for the extent of its legitimate application can grow over time. At a given time, the word can legitimately be applied to something to which it has not yet been applied, but which sufficiently resembles things to which the word has already been generally applied. That further application once made, the scope for legitimate applications in the future is correspondingly enlarged. To put the point schematically, suppose that at time t activities $x_0, \ldots, x_i$ are the only things to have been called 'games', that each of the activities $y_0, \ldots, y_j$ sufficiently resembles $x_0, \ldots, x_i$ to be legitimately called a 'game', but has not yet been so called, and that activity z does not sufficiently resemble $x_0, \ldots, x_i$ for that purpose. Now suppose that, by a later time t^*, $y_0, \ldots, y_j$ have been called 'games' and – as is quite possible – that z sufficiently resembles $y_0, \ldots, y_j$ to be legitimately called a 'game' if they have been. Then it is legitimate after t^* to call z a game, but it was not legitimate before t. Such a change in the legitimate applicability of a term can occur in virtue of a series of legitimate applications. The legitimacy of the process makes it natural to think of a single sense developing over time, not of a single word having first one sense, then another. On this view, the concept of a game has persisted through change.[31]

The dynamic quality of family resemblance concepts prevents them from being analysed as disjunctions of precise concepts. If *F* and *G* are precise concepts, so too is their disjunction, even though no one feature need be common to everything that is *F* or *G*. If one defined 'game' by a list ('cricket or chess or ring-a-ring-a-roses or . . .', the dots being actually filled in), then it would not apply to any newly encountered activity not on the list, no matter how much it resembled those that were. That is not how we use the word; one would merely have stipulated a new sense for it, more precise but less flexible. The extension of our concept is not 'closed by a frontier' (PI §68). To explain our concept, one must leave the dots after the list of examples.

Family resemblances do not themselves constitute vagueness.[32] Nevertheless, family resemblance concepts are obviously susceptible to borderline cases. How much resemblance to previous cases is sufficient for something to be a game? Speakers of the same language are bound to vary somewhat in their judgements of resemblance and their inheritance of precedents. The concept 'game' has blurred edges (PI §71). This blurring in no way prevents the concept from functioning successfully. Wittgenstein turns Frege's comparison of a vague concept to an area with blurred boundaries against him, for it is not always useless to indicate such an area. The request 'Stand roughly there' can be clearly carried out. We are not faced with an enemy who will use any gap in our concepts to escape.[33]

Family resemblance concepts appear susceptible to sorites paradoxes, indeed positively to invite them. What is to stop us from gradually extending the concept of a game to any activity we choose, such as nuclear warfare? At each stage, the new applications of the concept would resemble the old ones enough to be legitimate, even clearly so (the present sorites is quite different from one exploiting the vagueness of 'sufficient resemblance'). Wittgenstein does not discuss the question. Perhaps he would have said that the concept is in trouble only if we *do* make such extensions, not whenever we *could.* But is there any conceptual block to such extensions?

The paradox would be blocked by a requirement that any two games should sufficiently resemble each other (not in the same respect for every pair), sufficiency being determined by speakers' judgements. However, such a requirement violates the spirit of Wittgenstein's account of family resemblance. So too does the requirement that any game should sufficiently resemble a certain paradigm game (again, not in the same respect in every case): no game has that privileged status. If there are paradigm games, there are paradigm games of many kinds, and 'paradigm game' is itself a family resemblance term, threatened by its own sorites paradox. One hardly wants to block a paradox about paradigm games by invoking paradigm paradigm games.

A more hopeful suggestion is that the negation of a family resemblance concept is itself a family resemblance concept. If it is legitimate to deny that *x* is a game when it sufficiently resembles things that in the past have been legitimately denied to be games – just as it is legitimate to assert that *x* is a game when it sufficiently resembles things that in the past have been legitimately asserted to be games – then the expansionist tendencies of assertion and denial should hold each other in check. Equilibrium will be maintained by tension, since the conditions for legitimate assertion and legitimate denial may be met simultaneously, but a disputed no man's land of borderline cases was only to be expected. On a more complex account, the tension might be between 'game' and various specific contrary concepts, such as 'warfare', rather than between 'game' and its explicit contradictory 'non-game'. The underlying point remains. A sorites paradox is stopped when it collides with a sorites paradox going in the opposite direction.

That account will not strike a logician as solving the sorites paradox. It does not explain which activities *are* games. Nevertheless, it suggests a device in the functioning of family resemblance concepts that might prevent the paradoxes from arising in practice. From the standpoint of the *Philosophical Investigations*, may one ask for more?[34] One is not entitled to assume that there is a general cure for sorites paradoxes. However, hard questions arise even in particular cases. Faced with a particular sorites series for 'game', one can ask 'How many of these are games?'. The family resemblance model may not have, and may not be intended to have, the resources to explain why that is a bad question.

Wittgenstein does not claim that all vague concepts are family resemblance concepts. He points out that we do acquire some concepts by seeing what is common to different things. I might learn the meaning of 'yellow ochre' by seeing several patches of just that shade. In not quite the same way, I might see blue as what is common to light and dark shades of blue (PI §72). 'Blue' is vague by some standards, for it has borderline cases, but that does not make it a family resemblance term, for all shades of blue resemble each other in the same respect.

Wittgenstein's treatment of family resemblances is inconsistent with a kind of nominalism often attributed to him. On that view, a feature common to different things is nothing over and above the legitimate application of the same name to all of them; the only standard of resemblance is a linguistic one. If the legitimate application of a family resemblance word to different things does not entail the presence of a common feature, the nominalist view is false.

The argument can be transposed from words to concepts. Wittgenstein's treatment of family resemblances is inconsistent with a kind of conceptualism on which a feature common to different things is nothing over and above the legitimate application of the same concept to all of them, the only standard of resemblance being a conceptual one. If the legitimate application of a family resemblance concept to different things does not entail the presence of a common feature, the conceptualist view is false.

Wittgenstein compares a family resemblance concept to a thread made of overlapping fibres, no one of which runs through it all (PI §67). The individual fibres represent genuine common features; the thread does not. However, it does not follow that common features are wholly independent of human cognitive capacities, for Wittgenstein is speaking of common features that we can recognize as such. His point is that the recognition of a common feature is a far more specific and limited phenomenon, or group of phenomena, than philosophers have usually been prepared to acknowledge. It occurs against a pre-existing conceptual background, and cannot serve as a general explanation of the acquisition of concepts.

The point of the family resemblance model is negative. It is intended to undermine certain conceptions of language. It does not amount to a positive conception of vague language. Equally, it does not show the desire for such an account to be illicit. The account may be desired, not because it would give our language a foundation of a kind it cannot have, but because without it we lack a clear view of sorites paradoxes.

3.4 OPEN TEXTURE

The possibility of vagueness is central to Friedrich Waismann's conception of language, developed under Wittgenstein's influence. Like Wittgenstein, he emphasized the heterogeneity of language. Unlike him, he theorized explicitly about vagueness. The two concerns interacted; Waismann discussed vagueness with reference to contrasts between, for example, the language of physical objects and the language of sense impressions. Perhaps as a result, he never achieved a sharp focus on the phenomenon of vagueness itself.

Both Wittgenstein and Waismann had been struck in the 1930s by the logical positivists' identification of the meaning of a statement with its method of verification. For example, the meaning of an empirical statement such as 'That is a cat' might be identified with the kind of experience that would conclusively verify it. For a mathematical statement, the method of verification would be the requisite kind of proof. Metaphysical statements such as 'God is love' were dismissed as meaningless on the grounds that they had no method of verification. Since a verification is something carried out by a speaker of the language, the positivist theory promised to anchor the semantics of a language in the practice of using it. However, grave difficulties for the programme soon emerged. If the meaning of 'That is a cat' is exhausted by the kind of experience that would conclusively verify it, then one might expect it to be translatable into a statement about those experiences: but no one could provide a remotely plausible translation.

Waismann tried to describe a subtler connection between meaning and verification. On his view, empirical statements are never conclusively verified, for two reasons. First, although a finite number of experiences provide good enough evidence to warrant the assertion of 'That is a cat', they are never logically sufficient for its truth; in principle one might always turn out to have been the victim of an elaborate hoax or hallucinogenic drug. Second, suppose that all the relevant experiences, future as well as past, are somehow given. They may still be logically insufficient to decide the truth or falsity of 'That is a cat'. Waismann does not have in mind the sceptical possibility that all my experiences might be the work of an evil demon who gives me the best evidence experience can provide for the truth of 'That is a cat', while making the statement in fact false; he does not assume that the truth-value of the statement can transcend all possible experiential evidence. Rather, he points to cases in which experience takes an unexpected turn, not provided for in the meaning of the statement. The cat-like object disappears into thin air, then something just like it appears again; it grows to an enormous size, or is revived from death. Our understanding of 'That is a cat' gives us no basis on which to choose between the statement and its negation in such a case. It is neither verified nor falsified; equally, Waismann assumes, it is neither true nor false. When in this way a concept (such as 'cat') does not provide for every case, Waismann says that it has *open texture*.[35]

Most of our empirical concepts have open texture. In contrast, mathematical concepts have closed texture; they provide in advance for all possible cases.[36] Waismann does not identify vagueness with open texture. He regards a concept as vague only if unlegislated cases actually occur; it has open texture if they could occur. Open texture is 'something like *possibility of vagueness*'.[37] All vague concepts have open texture, but a concept like 'gold' may be open textured without vagueness if we can imagine unlegislated cases for it but they never in fact occur. Vagueness can be removed by stipulations, for they need only cover actual cases. Open texture can be reduced but not wholly removed by stipulations, for we can never foresee all the kinds of hitherto unlegislated case that could possibly arise. Even if we could, it might be added, the terms in which we made the stipulations would themselves have open texture.

Waismann tries to use the open texture of statements about physical objects to explain why they cannot be translated into statements about experiences. Similarly, he tries to use the open texture of statements attributing psychological properties, such as ‘She is intelligent’, to explain why they cannot be translated into statements about behaviour. However, the explanations are incomplete. A statement with open texture can be translated into another statement with open texture. Both statements about experiences and statements about behaviour have open texture too. What Waismann needs to show is that the open texture in a statement about experiences cannot match the open texture in a statement about physical objects, and that the open texture in a statement about behaviour cannot match the open texture in a psychological statement. He does not show that.

The idea of open texture was also intended to soften a number of philosophical dichotomies, most notably that between the analytic and the synthetic. Analytic truths such as ‘All brothers are siblings’ are supposedly true in virtue of their meanings; synthetic truths such as ‘All brothers are mortal’ require a contribution from the world as well as from language. One might expect the distinction not to be sharp in the presence of open texture. However, two claims of this sort need to be separated. A banal claim is that ‘analytic’ and ‘synthetic’, like most other words, have open texture; a statement can be on the borderline between them. A more contentious claim is that open texture in the object-language *entails* open texture in the metalinguistic terms ‘analytic’ and ‘synthetic’; it would require a more elaborate argument than Waismann supplies. But both claims imply that semantic relations are not an all-or-nothing matter.[38]

Does the law of excluded middle apply in the presence of vagueness? Waismann considers the question with reference to the description of sense impressions, which often requires vague terms.[39] I look at the night sky; I can say that I saw a fair number of stars, but I cannot say how many. The insistence that either I saw 735 stars or I didn’t is empty, Waismann suggests: we can cling to the law of excluded middle if we like, but it has become pointless to do so.[40] A disjunction is worth asserting only if one knows how to get oneself into a position to assert one or other disjunct; I do not know how to do that for the disjunction 'Either I saw 735 stars or I didn't'. Waismann's view of the link between meaning and verification seems to leave no room for one or other disjunct to be unverifiably true.

As Waismann points out, the case challenges our usual assumptions about vagueness. We expect excluded middle to fail for vague statements, not precise ones; here it fails for the apparently precise 'I saw 735 stars', not the vague 'I saw many stars'. We expect vagueness to be in the description, not the thing described; here the reverse seems to obtain. What is going on?

Waismann seems to have conflated several separate doubts about the law of excluded middle.

1 The phrase 'star seen by me' may have borderline cases: for example, stars on the threshold of visibility. This is just a particular case of the general problem of vagueness; it shows nothing about the nature of sense impressions. 'I saw 735 stars' is after all not a perfectly precise statement.

2 Waismann emphasizes the past tense of 'I saw'. It is too late to count the stars. If the meaning of statements about the past is explained in terms of our methods for verifying them, the likely outcome is an anti-realist view of the past as existing only in present traces. On this view, the unspecificity and vagueness of our memories correspond to unspecificity and vagueness in the past itself. The law of excluded middle may fail for past tense statements, whether or not they concern sense impressions. I may know that I saw all and only those stars that cast light on my spectacles; if the law fails for 'I saw 735 stars', then it fails for '735 stars cast light on my spectacles'. Yet Waismann does not propose a generalized anti-realism about the past, and nothing he says would make it plausible. It might rather be treated as a *reductio ad absurdum* of a verificationist theory of meaning.[41]

3 'I saw 735 stars' might be taken to mean 'I saw that there were 735 stars [in that part of the sky]'. There may well be no number *n* such that I saw that there were *n* stars. I may equally not have seen that there were not 735 stars. My visually derived information was incomplete. However, this is no threat to the law of excluded middle. For on this reading, the negation of 'I saw 735 stars' is not 'I saw that

there were not 735 stars' but 'I did not see that there were 735 stars', which is obviously true in the envisaged case (we are no longer concerned with problem (1)).

Sense impressions may pose a special threat to the law of excluded middle, but Waismann has not shown that they do. However, his discussion suggests a better point. I saw that there were many stars in the sky. If the content of my visual impression is what I saw in that sense, visual impressions themselves – not just our descriptions of them – can be vague. This does not contradict Russell's thesis that only representations are vague, if visual impressions are representations. My visual impression represents the sky as containing many stars. In effect, Waismann argues that since our visual impressions have vague content, vague language is needed to report them. This is close to the argument that since people use vague language, vague language is needed to report what they have said. The point is not trivial; it is a serious difficulty for the idea that all truths can be stated in precise language. Beliefs, desires, intentions and other psychological states also have vague contents; how can they be fully described in a precise language?

Vague words are sometimes better than precise ones for reporting sense impressions; often neither are adequate. Although 'That there were many stars in the sky' may correctly answer both 'What did he see?' and 'What did he say?', it can completely answer the latter question in a way in which it cannot completely answer the former. Indeed, a well-chosen 'that'-clause *reproduces* what was said; it cannot reproduce what was seen. The content of sense impressions may not even be capable of full conceptualization. That would not undermine the preceding argument. In particular, a sense impression with a wholly or partly unconceptualized content may still be vague, if its content can be treated as perceptual information. The information will of course be incomplete, being clearly compatible with more than one possible state of affairs. For example, my impression of the night sky fails to determine an exact number of stars, but that is a matter of unspecificity rather than vagueness (compare (3) above). However, the information may also be genuinely vague, being neither clearly compatible nor clearly incompatible with some possible state of affairs. For example, there are numbers n such that it is not clear that I see that there are at least n stars and not clear that I do not see that there are at least n stars (compare (1) above).

One might expect a picture to be worth a thousand words in reproducing the content of a visual impression. That is not Waismann's view. He treats pictures as precise, and in that respect less able than vague words to convey a vague impression. If one must use pictures, several are needed, between which the impression is indeterminate. One shows pictures with different numbers of stars, saying each time 'It looked roughly like that'. Waismann's argument may underestimate the capacities of pictures. An impressionist painting of the night sky can leave the number of stars vague; even the colour of the sky is not precisely specified by flecks of paint of many colours. If there are deeper reasons why the vagueness and unspecificity in a picture cannot match the vagueness and unspecificity in a visual impression, Waismann does not explain them. Since he does not distinguish between representational and non-representational properties of a picture, he may have been misled by comparing specific physical non-representational properties of paintings with the vague and unspecific representational properties of visual impressions, for it is not obvious what the non-representational properties of visual impressions are.

On Waismann's view, a visual impression can be conveyed with equal roughness by different pictures, and so by different, even incompatible, precise descriptions. Only the class of those descriptions as a whole adequately describes the impression. He therefore tries to construct a logic for talk about sense impressions as a logic of classes of propositions. If any proposition in such a class is true, the class itself counts as true; if every proposition in the class is false, the proposition itself counts as false. A class is implied by any of its subclasses. So far, one might think of the class whose members are the propositions '*P*', '*Q*', . . . , '*Z*' as simply the disjunctive proposition '*P* or *Q* or . . . or *Z*', but Waismann rejects this view. He reads the disjunction classically, so that it is true only if at least one of its disjuncts is true; in contrast, he counts some classes as true when all their members are indeterminate in truth-value (or some are indeterminate and some false). For example, the truth that I saw many stars is identified with a class of propositions of the form 'I saw n stars', none being true or false.

The last point indicates a limitation of Waismann's project. If a logic of classes is constructed out of a logic of propositions, one might hope that puzzling properties of the former would be explained in terms of unpuzzling properties of the latter. But Waismann requires the base propositions to have puzzling properties. Some of them are neither true nor false, for otherwise every class is equivalent to the disjunction of its members after all, and so is nothing new.[42] Thus Waismann's first need is for a logic of propositions that can be neither true nor false, but this he does not supply.

A more promising remark is that the membership of the class of propositions will often be indeterminate. For a class with determinate and multiple membership corresponds merely to an impression whose content is precise but unspecific. On this view, what corresponds to an impression whose content is vague is a class whose membership is indeterminate. Even if all the base propositions are true or false, and a class is equivalent to the disjunction of its members, a class of indeterminate membership might count as indeterminate in truth-value. For if only false propositions are determinately members of the class, but the membership of some true proposition is indeterminate, then the class will be indeterminate between a false disjunction and a true one. However, Waismann's formal remarks, like Black's, were too inchoate to inspire significant development.

Chapter 4

Many-valued logic and degrees of truth

4.1 OVERVIEW

According to the principle of bivalence, an utterance that says something is either true or false. The sentence uttered may itself be classified as true or false, relative to the context of utterance (this context relativity will be left tacit). Bivalence is integral to all standard explanations of the formal systems at the core of modern logic. For Frege, Russell and the younger Wittgenstein, a logically perfect language is two-valued. Every well-formed formula in it is true or false.

Bivalence is problematic for vague languages, as the Greeks knew. Borderline sentences seem impossible to classify as true or false, yet something is said to be the case by utterances of them. This may prompt a search for a scheme of classification with more than two categories. Perhaps a third category is needed, 'neither true nor false', or a whole range of new categories of the form 'true to such-and-such a degree'. These new categories may be called 'truth-values' in an extended sense. Since most or all of our reasoning is conducted in vague language, it would then be natural to seek logical systems that stand to the new many-valued classificatory scheme as the classical systems stand to the bivalent scheme. However, the analogy can be developed in more than one way. The systems considered in this chapter select a specific feature of two-valued logic for generalization: truth-functionality.

The chapter is organized as follows. Section 4.2 explains the notion of truth-functionality and the principles by which it has been generalized from two-valued to many-valued logic. Sections 4.3–4.6 describe early attempts to apply three-valued logic to the problem of vagueness, and the difficulties they faced. The objections to two-valued logic, if sound, generalize to three-valued logic, and the generalization of truth-functionality to three values is vulnerable to further objections of its own. Section 4.7 gives a familiar rationale for preferring logics with infinitely many values to those with three or some larger finite number when vagueness is in question. Section 4.8 specifies the detailed workings of generalized truth-functionality in a logic with infinitely many values. Section 4.9 sketches early attempts to develop such logics. Section 4.10 applies the framework of infinitely many values to sorites paradoxes. Section 4.11 applies it to some constructions in natural languages; however, the relevant notion of degree of truth turns out to have no clear connection with the problem of vagueness. Section 4.12 discusses higher-order vagueness, arguing that it makes most work on many-valued logic irrelevant to the problem of vagueness. Section 4.13 shows that attempts to replace numerical values in many-valued logic by non-numerical degrees of truth face unexpected difficulties, particularly in the case of negation. Section 4.14 argues that the attempt to generalize truth-functionality to degrees of truth is fundamentally mistaken. As an appendix, Section 4.15 sketches some technical issues in the axiomatization of logics with infinitely many values.

Conclusion: the nature of vagueness is not captured by any approach that generalizes truth-functionality.

4.2 TRUTH-FUNCTIONALITY

The basis of both modern and Stoic logic is propositional logic. It studies forms of inference characterized in terms of sentence functors such as 'it is not the case that . . .' (~), 'either . . . or —' (∨), '. . . and —' (&), 'if . . . then —' (⊃) and ' . . . if and only if —' (≡). Any sentence functor constructs complex declarative sentences out of simpler ones. But not all sentence functors are studied. Those chosen have, on their preferred readings, a special feature, truth-functionality. If a truth-functional sentence functor * is used to construct a complex sentence $*(p_1, \ldots, p_n)$ out of the simpler sentences $p_1, \ldots, p_n$, then the truth-values of $p_1, \ldots, p_n$ determine the truth-value of $*(p_1, \ldots, p_n)$. For example, if one knows whether the simple sentences 'It is hot' and 'It is wet' are true or false, then one can work out whether the complex sentence 'Either it is not the case that it is hot or it is wet' is true or false. Such dependences of truth-value are displayed in the truth-tables for these functors:

p	$\sim p$
T	F
F	T

$p\ q$	$p \vee q$	$p \& q$	$p \supset q$	$p \equiv q$
T T	T	T	T	T
T F	T	F	F	F
F T	T	F	T	F
F F	F	F	T	T

Granted bivalence, each possible combination of truth-values for the constituent sentences is represented by a line of the truth-table, from which one can read off the truth-value of the complex sentence in the appropriate column. But if there is a possibility that the sentence p is neither true nor false, the truth-tables above say nothing about it. They might still be correct in what they say about the cases they do cover, but they would not cover all the cases.[1]

A sentence functor is truth-functional just in case it has a complete truth-table. If one follows Frege in treating the truth-value of a sentence as its referent, then to say that the functor * is truth-functional is just to say that the referent of the sentence $*(p_1, \ldots, p_n)$ is a function of the referents of the sentences $p_1, \ldots, p_n$, just as the complex sentence itself is a function of the simpler ones. This function from the truth-values of $p_1, \ldots, p_n$ to the truth-value of $*(p_1, \ldots, p_n)$ will then be treated as the referent of the functor *; thus the truth-tables above display the referents of $\sim$, $\vee$, &, $\supset$ and $\equiv$. On this view, the truth-functionality of sentence functors follows from the condition that the referent of a complex expression is determined by the referents of its constituents. Since the logically perfect languages of Frege, Russell and the younger Wittgenstein meet this condition, all their sentence functors are truth-functional.

In natural languages, not all sentence functors are truth-functional. The sentence functor 'It is a contingent fact that . . .' is not truth-functional, for although 'A war began in 1914' and '2 + 2 = 4' have the same truth-value, 'It is a contingent fact that a war began in 1914' and 'It is a contingent fact that 2 + 2 = 4' have different truth-values. The line of the truth-table for 'It is a contingent fact that p' on which p is true cannot be completed. In such a case, either the referent of the complex sentence is not determined by the referents of its constituents or the truth-value of a sentence is not its referent. This may or may not be an insuperable difficulty for the traditional conception of a logically perfect language. In any case, faced with a sentence functor in a natural language, one has no automatic right to assume that it is truth-functional.

The truth-tables provide a mechanical test of validity for finite truth-functional forms of inference. An example is *disjunctive syllogism*, with the formulas $p \vee q$ and $\sim p$ as premises and q as conclusion. An instance of this form is the argument from 'Either he went left or he went right' and 'It is not the case that he went left' to 'He went right'. Disjunctive syllogism is valid because no assignment of truth or falsity to the sentence variables p and q results in the assignment of truth to both premises and falsity to the conclusion, as one can quickly check from the truth-tables. Thus any particular argument of that form, the result of substituting particular sentences for the sentence variables, will have a true conclusion if it has true premises. Similarly, the law of excluded middle, the unpremised conclusion $p \vee \sim p$, is valid because it is assigned truth by both assignments to its sentence variable. The form of inference with $p \vee q$ and p as premises and $\sim q$ as conclusion is invalid, since the assignment of truth to both p and q results in the assignment of truth to both premises and falsity to the conclusion. Thus if particular true sentences are substituted for the sentence variables, the result is a particular argument of that form with true premises and a false conclusion.

The use of the two-valued truth-tables to establish validity depends on the assumption of bivalence, whereas their use to establish invalidity does not. For validity, the method must take account of *every* argument of the relevant form. For invalidity, it suffices that *some* argument of that form has true premises and a false conclusion.

To reject bivalence is to lose both the account of the meaning of basic logical operators and the mechanical test of validity for inference forms involving them provided by the two-valued truth-tables. But if the classification of sentences into truths and falsehoods has been replaced by a classification into more than two categories, it is natural to seek to replace the two-valued truth-tables by tables involving more than two values. Such tables would classify complex sentences into the relevant categories according to the classification of their components. They would thereby provide both a new account of the meaning of basic logical functors and a new test of validity for inference forms involving them.[2] This is the central idea of many-valued logic: generalized truth-functionality.

If 'true' and 'false' survive as categories within the new classification, the two-valued truth-tables will presumably be embedded within the new tables. For example, the conjunction of a true sentence with a false one will still be classified as false. The new tables will merely add lines to cover those combinations in which at least one component is other than true or false. In this sense the two-valued tables are treated as incomplete rather than incorrect. Even if 'true' and 'false' are each subdivided into several new categories, the new tables will presumably imply the correctness of the old ones for the cases covered by the latter.

A generalization of the concept of a truth-value makes possible a generalization of the concept of truth-functionality. However, the classification of sentences into more than two categories by itself carries no commitment to the claim that logical connectives in natural language are truth-functional in the generalized sense, that the classification of complex sentences formed by their means depends only on the classification of their components. Suppose, for example, that 'neither true nor false' is treated as an alternative to 'true' and 'false'. If the corresponding generalization of truth-functionality holds for the conditional $p \supset q$, then the status of p and of q on this threefold classsification determines the status of $p \supset q$. It follows in particular that if any such conditional whose antecedent and consequent are neither true nor false is true, then every such conditional is true. Thus if both 'n grains make a heap' and '$n + 1$ grains make a heap' are neither true nor false, and 'If n grains make a heap then $n + 1$ grains make a heap' is true, then 'If $n + 1$ grains make a heap then n grains make a heap' is also true. This rules out the view that in such a case the former conditional is true (because another grain can only help) while the latter is neither true nor false (because sorites paradoxes are to be avoided). Generalized truth-functionality is therefore a contentious assumption.

Those who replace the true/false dichotomy by a many-fold classification have reason to hope that generalized truthfunctionality holds, for then they gain the advantages noted above: an account of the meaning of logical operators and a test of validity for arguments involving them. But reasons to hope that generalized truth-functionality holds are not always reasons to believe that it does. Even the reasons to hope for generalized truth-functionality are defeasible. Some who use a threefold classification of sentences reject the corresponding generalization of truth-functionality and make up the losses thereby incurred in other ways (examples will be discussed in the next chapter). Unfortunately, few proponents of many-valued logic have recognized, still less discharged, the need to provide reasons for believing basic logical operators to be truth-functional in the generalized sense.

The notion of validity must also be generalized in many-valued logic. In two-valued logic, a form of inference is valid just in case it preserves truth from premises to conclusion. Equally, it is valid just in case it preserves non-falsity, for non-falsity is equivalent to truth. In many-valued logics these definitions of validity are not equivalent; the new classification may not even contain the categories of truth and falsity. Thus a many-valued logic has not been fully given until a notion of validity is specified. One technique is to *designate* certain values, and define validity as the preservation of designated values: if all the premises take designated values, so must the conclusion. If the three values are truth, falsity and neutrality, one might designate just truth, or both truth and neutrality. A form of inference may count as valid on one policy and not on the other.

The use of designated values is sometimes held to reintroduce bivalence by the back door, since every sentence takes either a designated or an undesignated value. However, this is not to reintroduce two-valued truth-functionality. For example, suppose that truth is the only designated value and that negation has the following table:

p	~p
T	F
N	N
F	T

Thus if p is false and q is neutral, they both take undesignated values; but $\sim p$ is assigned truth, a designated value, while $\sim q$ is assigned neutrality, an undesignated value. The status of a sentence as designated or undesignated does not determine the status of its negation as designated or undesignated.

Moreover, the use of designated values is not the only way of defining validity in many-valued logic. An ordering of values may be specified ('truer than') and an argument defined to be valid just in case the value of the conclusion is at least as high as the lowest value of a premise. Alternatively, validity itself may be made a matter of degree. Many-valued logic does not betray itself merely by its conception of validity.

With these explanations in place, we may turn to the history of applications of many-valued logic to the problem of vagueness.

4.3 THREE-VALUED LOGIC: BEGINNINGS

One of the first to conceive of many-valued logic was also a pioneer of modern two-valued logic: C.S. Peirce. He introduced the standard two-valued truth-tables; he later wrote down three-valued tables but never published them. His idea was that two-valued logic, although not wholly incorrect, is valid only within a limited domain, three-valued logic being needed for full generality. In a manuscript dated 23 February 1909 he states 'Triadic Logic is that logic which, though not rejecting entirely the Principle of Excluded Middle, nevertheless recognizes that every proposition, S is P, is either true, or false, or else S has a lower mode of being such that it can neither be determinately P, nor determinately not P, but is at the limit between P and not P'. It is tempting in retrospect to connect Peirce's rationale for three-valued logic with his work on the problem of vagueness, but we cannot be sure; his surviving notes are too fragmentary. Nor is it clear how fixed was Peirce's belief in triadic logic; he wrote 'All this is mighty close to nonsense' on one page of the manuscript.[3]

Many-valued logic appeared in print in 1920, with a paper by the Polish logician Jan Lukasiewicz. However, his concern was not the problem of vagueness but that of free will. He thought that fatalism can be avoided only if some statements about the future, such as 'There will be a sea-fight tomorrow', are not yet true or false. The same problem had worried the Greeks, amongst whom it was a more common reason than sorites paradoxes for rejection of bivalence. Lukasiewicz assigned future contingents a third value and developed a three-valued logic to match.[4] Mathematical research on many-valued logic developed rapidly. The connection with vagueness now seems obvious, for two ideas became familiar in the 1920s: that the law of excluded middle does not generally hold in many-valued logic, and that it does not generally apply to vague languages. Nevertheless, the connection was slow to be made. Although Black's consistencies of application might be regarded as a continuum of alternative values, he gave nothing like consistency of application tables in any case except that of negation (the consistency of application of ~*p* is the reciprocal of the consistency of application of *p*). He did not postulate generalized truth-functionality.

4.4 THREE-VALUED LOGIC: HALLDÉN

Perhaps the first serious attempt to treat vagueness with many-valued logic was published by the Swedish logician Sören Halldén in 1949. The title of his monograph, *The Logic of Nonsense*, sounds like a contradiction in terms. Pure nonsense is sheer gibberish; how can there be a logic of sheer gibberish? However, Halldén specifies at once that by calling a proposition 'nonsensical' or 'meaningless' he means only that it is neither true nor false.[5] A proposition is 'meaningful' just in case it is either true or false. As an example of a 'meaningless' proposition, Halldén gives the Liar: 'This proposition is false'. Unlike most writers of the period, he also takes the sorites paradoxes seriously. He suggests that the borderline case question 'Is the man with a hundred hairs bald?' should be rejected as 'meaningless', since its only possible answers are 'meaningless' propositions. Neither 'The man with a hundred hairs is bald' nor 'The man with a hundred hairs is not bald' is true or false.[6] These borderline propositions are clearly not meaningless in the usual sense. If Jack is the man with a hundred hairs, 'Jack is bald' is 'meaningless' in Halldén's sense, but one understands it quite well; what it actually says would have been clearly true in some alternative circumstances and clearly false in others. The point is brought out by Halldén's remark that the 'meaningfulness' or 'meaninglessness' of such a proposition is often a contingent matter.[7] Whether 'Jack is bald' is 'meaningless' depends on the state of Jack's scalp; whether it is sheer gibberish does not. The scare quotes round 'meaningful' and 'meaningless' are therefore best left in place.

Halldén adapts logic to ‘meaningless’ language by adopting a three-valued logic whose values are truth, falsity and ‘meaninglessness’. In modifying the two-valued truth-tables, he follows a simple policy. If every component is true or false, the complex proposition has the same truth-value as in the two-valued table. If any component is ‘meaningless’ (N), the complex proposition is ‘meaningless’ too:[8]

p	$\sim p$
T	F
F	T

p q	$p \vee q$	$p \& q$	$p \supset q$	$p \equiv q$
T T	T	T	T	T
T N	N	N	N	N
T F	T	F	F	F
N T	N	N	N	N
N N	N	N	N	N
N F	N	N	N	N
F T	T	F	T	F
F N	N	N	N	N
F F	F	F	T	T

Halldén also introduces a new one-place operator + on a different principle. $+p$ is to mean that p is ‘meaningful’. Thus if p is ‘meaningless’, $+p$ is false rather than ‘meaningless’. Otherwise, it is true:

p	$+p$
T	T
N	F
F	T

Thus $\sim +p$ will mean that p is ‘meaningless’.

Halldén’s policy with the standard functors corresponds to the Fregean idea that the referent of a complex expression is a function of the referents of its components, the referent of a declarative sentence being truth or falsity. The function has nothing to work on unless the components deliver their referents, so the complex expression refers only if all its components do. However, the table for + treats the failure to refer (to be true or false) as itself a way to refer in an extended sense (to be true, false or ‘meaningless’), a very unFregean step. Thus it would be inconsistent to give the other tables the Fregean rationale. It is unclear what alternative rationale they have.

Surprisingly, Halldén develops a logic whose valid formulas coincide with those of classical logic (for formulas not involving +). One might have expected the law of excluded middle, $p \vee \sim p$, to be invalid, since it is 'meaningless' when p is, according to Halldén's tables. However, all he demands of a valid formula is that it cannot be false, not that it must be true, for he treats 'meaninglessness' as well as truth as a designated value. Although $p \vee \sim p$ is not always true, it is never false. More generally, if A is classically valid, then Halldén's tables make it true whenever all its sentence variables are 'meaningful', and 'meaningless' otherwise; thus it is valid in Halldén's sense. Conversely, if A is valid in his sense, then it is classically valid, for when all its sentence variables are true or false, it will be 'meaningful' and not false, so true. The rationale for Halldén's designation policy is clear. If truth were the only designated value, so that a formula would be valid only if always true, then no formula not involving + would be valid, since any such formula is 'meaningless' when some of its sentence variables are. Not even $p \supset p$ would be valid. But surely some formulas not involving + deserve to count as valid. Thus 'meaninglessness' must be a designated value too.

The upshot of Halldén's discussion is that no sentence built up from vague component sentences by means of the standard sentence functors is logically guaranteed to be true, but any classical tautology is guaranteed not to be false. However, this vindication of classical logic is more limited than it looks. Logic deals with *reasoning*. We expect it to pick out good inferences, not just good propositions; we are less interested in believing tautologies than in safely advancing from old beliefs to new ones. Many classical forms of inference are invalid on Halldén's account.

One example is the rule of *modus ponens*. An application of it is the inference from $(p \vee \sim p) \supset +p$ and $p \vee \sim p$ *to* $+p$. Its premises are Halldén-valid, since neither is ever false, but its conclusion is not Halldén-valid, since it is false when p is 'meaningless'. Halldén is therefore forced to restrict *modus ponens*. The example depends on the operator +; when the premises of *modus ponens* are Halldén-valid formulas not involving +, so is the conclusion. Halldén provides rules for inferring valid formulas from valid formulas. However, we also want to infer new contingently correct beliefs (such as 'The builders have started work') from old ones (such as 'If there is a heap of sand in the garden then the builders have started work' and 'There is a heap of sand in the garden'). If being correct is equated with taking a designated value, then we are interested in forms of inference whose conclusions take a designated value whenever their premises do, i.e. in valid forms. *Modus ponens* lacks this property, even when it is restricted to formulas not involving +. If p is 'meaningless' and q false, then the premises $p \supset q$ and p are not false while the conclusion q is false. Thus even the restricted form of *modus ponens* counts as invalid. Similarly, the rule of &-elimination – from p & q to infer p (alternatively, q) – counts as invalid. If Jack is a non-philosopher on the borderline of baldness, the premise 'Jack is a bald philosopher' ('Jack is bald and Jack is a philosopher') is 'meaningless' while the conclusion 'Jack is a philosopher' is false. More generally, one can show that, on Halldén's tables, a form of inference not involving + preserves non-falsity if and only if its conclusion follows classically from those premises involving only sentence variables which also occur in the conclusion.[9] Any premise involving a sentence variable not involved in the conclusion is irrelevant to the preservation of non-falsity.

On Halldén's tables, the preservation of non-falsity is not a very useful notion of validity. The vagueness of 'bald' should not invalidate the inference from 'Jack is a bald philosopher' to 'Jack is a philosopher'. Even on Halldén's tables, the truth of p & q guarantees the truth of q, as does the truth of p and $p \supset q$. One might therefore decide to make truth the only designated value after all, acknowledging the consequence that no individual formula not involving + will be valid, but taking comfort in the thought that more inferences will be valid. We want our beliefs and assertions to be true, not just non-false; 'meaninglessness' is not enough. However, the new criterion of truth preservation does not restore all classically valid inferences, and even invalidates some new ones. For example, the $\vee$-introduction rule – from p (alternatively, q) to infer $p \vee q$ – does not preserve truth on Halldén's tables, since it has a true premise and a 'meaningless' conclusion when p is true and q 'meaningless'. More generally, one can show that on Halldén's tables a form of inference not involving + is truth-preserving if and only if it is classically valid and either every sentence variable in the conclusion is also in some premise or the premises are classically inconsistent.[10] The argument from 'Jack is not a philosopher' to 'Jack is not a bald philosopher' has a true premise and a 'meaningless' conclusion if Jack is a non-philosopher on the borderline of baldness. This result is hard to accept.[11]

The trouble lies in Halldén's tables, not in the choice of designated values. No rationale they may have in the case of paradoxes such as the Liar extends to mere borderline cases. If Jack is a lifelong non-philosopher, 'Jack is a bald philosopher' should remain false throughout his life, rather than being false when he has a full head of hair, becoming 'meaningless' as he loses his hair and then returning to falsity when he is clearly bald. The falsity of 'Jack is a philosopher' suffices for the falsity of 'Jack is a bald philosopher'. Similarly, 'Jack is either a non-philosopher or a bald philosopher' should remain true throughout his life, rather than being true when he has a full head of hair, becoming 'meaningless' as he loses his hair and then returning to truth when he is clearly bald. The truth of 'Jack is a non-philosopher' suffices for the truth of 'Jack is either a non-philosopher or a bald philosopher'. At least in the case of vagueness, a conjunction with a false conjunct is false, and a disjunction with a true disjunct is true. These principles guided later attempts to construct many-valued logics of vagueness.

Some of Halldén's discussion of vagueness can be separated from his three-valued logic. He introduced the idea that if a vague property can be attributed at all, it can be attributed on the basis of precise properties. If we know that someone has a head covered with hair, then we can deduce that he is not bald; if we know that he has no hair on his head, then we can deduce that he is bald. Here the concepts 'has a head covered with hair' and 'has no hair on his head' do not involve the concept 'bald'. The set of independently specifiable properties either entailing baldness or inconsistent with it is the *decision-range* of baldness. On Halldén's view, it is true to say that x has a vague property F if and only if x has some property in the decision-range of F that decides in favour of F. It is false to say that x has F if and only if x has some property in the decision-range of F that decides against F. Thus it is 'meaningful' to say that x has F if and only if x has some property in the decision-range of F. Otherwise, x is a borderline case of F.[12] Given that all properties in decision-ranges are precise, the vague facts are determined by the precise ones. If two possible situations are alike in respect of all precise properties, then they are alike in respect of all vague properties too. In modern terminology, vague properties *supervene* on precise ones. A tempting conclusion to draw, but not a necessary one, is that the world can be completely described without resort to vague language.[13]

4.5 THREE-VALUED LOGIC: KÖRNER

Vagueness was treated with a different version of three-valued logic in a series of works by Stephan Körner from 1955 onward.[14]He sketched a 'logic of inexact concepts' and applied it to the philosophy of science. The inexactness stems from borderline cases for concepts defined by examples. Such a concept F divides objects into positive candidates, which must be 'elected' as positive instances of F, negative candidates, which must he elected as negative instances, and neutral candidates, which may be elected as positive or as negative instances, depending on a free choice. In effect, there is a three-valued pre-election logic and a two-valued post-election logic. In classifying relations between concepts, Körner takes both stages into account. For example, all the positive candidates for F may also be positive candidates for G, but this inclusion is not preserved by an election in which some neutral candidates for both concepts are elected as positive instances of F but as negative instances of G.

To the three kinds of candidate for a concept correspond three kinds of proposition: true, false and neutral. Truth and falsity are conceived as stable states, neutrality as a provisional one. We can elect a neutral proposition as true or false by a free choice. This suggests a specific policy for adapting the two-valued tables to the possibility of neutral propositions. Consider a line of the three-valued table for $*(p_1, \ldots, p_n)$, and all possible elections of any neutral p_is as true or as false. If $*(p_1, \ldots, p_n)$ is true on the two-valued table for all possible elections, it is assigned truth on the given line of the three-valued table. If it is false on the two-valued table for all possible elections, it is assigned falsity. If it is true on the two-valued table for some possible elections and false for others, it is assigned neutrality. The result is the following set of tables:[15]

p	~p
T	F
N	N
F	T

p q	p∨q	p&*q*	p⊃*q*	p ≡ q
T T	T	T	T	T
T N	T	N	N	N
T F	T	F	F	F
N T	T	N	T	N
N N	N	N	N	N
N F	N	F	N	N
F T	T	F	T	F
F N	N	F	T	N
F F	F	F	T	T

The tables for ~ and ≡ are just like Halldén's. For &, there is a difference only when one conjunct is false and the other intermediate. Halldén makes the conjunction intermediate, because one conjunct is; Körner makes it false, because the conjunction will be false whether the neutral conjunct is elected as true or as false. Given the earlier criticism of Halldén's table for &, this difference tells in Körner's favour. For ∨, there is a difference only when one conjunct is true and the other intermediate. Halldén makes the disjunction intermediate, because one disjunct is; Körner makes it true, because the disjunction will be true whether the neutral conjunct is elected as true or as false. This difference too tells in Körner's favour. There is a similar difference for ⊃: when the antecedent is intermediate and the consequent true, and when the antecedent is false and the consequent intermediate, Halldén makes the conditional intermediate while Körner makes it true.

Körner does not introduce an operator corresponding to Halldén's 'meaningfulness' operator +. This has a rationale in his conception of neutrality as a provisional status. To make $+p$ false when p is neutral would be to give permanent recognition to something temporary.

What is the appropriate notion of validity for Körner's tables? If truth were the only designated value, no formula would be valid, for a formula is neutral when all its sentence variables are. Even formulas such as $\sim(p \,\&\, \sim p)$, $p \supset p$ and $p \equiv p$ would be invalid. If both truth and neutrality were designated values, the valid formulas would coincide with those of classical logic, as in Halldén's system. However, *modus ponens* would then be invalid, since $p \supset q$ and p would take designated values and q an undesignated one when p is neutral and q false. *Modus ponens* is valid when truth is the only designated value, but other classical inferences are invalid, such as that from p to $q \supset (p \& q)$.[16]

A further problem faces Körner's approach. If p is neutral, his tables make $\sim(p \& \sim p)$ neutral too. The law of non-contradiction fails. Yet neutrality is a provisional status. Whether p is elected as true or as false, $\sim(p \& \sim p)$ will be false after the election. It should therefore count as false before the election too, by the rationale for Körner's three-valued tables, but it does not. The problem lies in the line of the three-valued table on which a conjunction is neutral when its conjuncts are, for it takes no account of possible interdependence between the conjuncts, as when one is the negation of the other. It would not help to substitute another value at that point in the table, for the conjunction of a neutral conjunct with itself *should* be neutral. Rather, the supposed rationale for the truth-tables is inconsistent with generalized truth-functionality. If the value of a complex proposition depends on the results of all possible elections of its neutral constituents as true or as false, it is not determined by the values of the constituents alone, for it also depends on connections between the effect of an election on one constituent and its effect on another.[17]

The effect of Körner's three-valued tables is to treat different occurrences of the same constituent in a complex proposition as though they were independent. Thus $p \& \sim p$ is treated like $p \& \sim q$ when both p and q are neutral, a candidate for truth as well as falsity. Remarkably, Körner sometimes endorses this effect. He explicitly permits an object to be elected as a positive instance of a concept even if it has already been elected as a negative instance of that concept.[18] Thus if p is 'Jack is bald', $p \& \sim p$ might indeed come out true if Jack were elected as a positive instance of baldness with respect to the first occurrence of p and as a negative instance with respect to the second. Jack would be both bald and not bald, and a contradiction would be true. Indeed, logic would be abolished, for no form of inference is valid in Körner's system once different occurrences of the same expression are treated independently.[19] The rot could be stopped by a general convention mooted by Körner, according to which nothing may be elected as a positive instance of a concept once elected as a negative instance, and vice versa. Such a convention seems needed if elections are to have more than momentary significance, but in reintroducing the interdependence of different occurrences of the same expression it would undermine Körner's three-valued tables. A formal technique better suited to Körner's philosophical approach might be that of supervaluations, discussed in the next chapter.

4.6 THREE-VALUED LOGIC: SECOND-ORDER VAGUENESS

Neither Halldén nor Körner managed to give a plausible three-valued logic for vagueness. Indeed, a further and more general objection faces the project. The objection to two-valued logic was the supposed impossibility of classifying all vague propositions as true or false: but the phenomenon of second-order vagueness makes it equally hard to classify all vague propositions as true, false or neither. As grain is piled on grain, we cannot identify a precise point at which 'That is a heap' switches from false to true. We are equally unable to identify two precise points, one for a switch from false to neutral, the other for a switch from neutral to true. If two values are not enough, three are not enough.

Both Halldén and Körner briefly address the problem. For Halldén, it takes a sharp form. If a proposition is 'meaningful', it is true to say that it is 'meaningful', and therefore 'meaningful' to say that it is 'meaningful'. If the proposition is 'meaningless', it is false to say that it is 'meaningful', and therefore again 'meaningful' to say that it is 'meaningful'. In terms of the three-valued table for the 'meaningfulness' operator +, any of the three values for *p* makes +*p* true or false and therefore ++*p* true. Thus ++*p* counts as valid. Now Halldén supposes that propositions on the boundary between truth and falsity are 'meaningless'; if *p* were on the boundary between falsity (or truth) and 'meaninglessness', +*p* would be on the boundary between truth and falsity, and would therefore be 'meaningless': but then ++*p* would be false. Thus ++*p* is valid if and only if there is no second-order vagueness. Halldén describes a syntactic modification of his system in which ++*p* is not derivable, but hopes to avoid its extra complexity. He suggests that any second-order vagueness might be removed by a convention.[20] This does not seem very promising, for the terms in which the convention was stated would themselves be vague. Moreover, if it is appropriate to remove second-order vagueness by a convention in order to retain the simplicity of the three-valued system, why is it not appropriate to remove first-order vagueness by another convention in order to retain the still greater simplicity of the two-valued system?

Körner mentions higher-order vagueness only to admit that it exists. The tripartite classification of objects into positive, negative and neutral candidates for an inexact concept may itself be inexact. Something can be a neutral candidate for the concept of a neutral candidate for the concept *F*.[21] This shows how far from literally the notion of a candidate is to be taken. If nothing could be elected as a positive instance of the concept *F* without previously having been elected as a positive instance of the concept of a positive or neutral candidate for the concept *F*, a vicious regress would obviously threaten. Something can be a positive instance of a concept without having been elected as such in any interesting sense.

Second-order vagueness poses a general threat to three-valued logic. If three-valued tables define the standard functors then (a) the three values should exhaust all possibilities (otherwise the definition would not cover all cases) and (b) any three-valued truth-table should define some operator (otherwise some constraint on meaning has been left unaccounted for). By (b), it is legitimate to introduce Halldén's functor +. By (a), $+p \vee {\sim}{+p}$ is always true (since $+p$ is true or false on every line of its truth-table and the three-valued tables extend the two-valued ones). Now the usual rationale for three-valued logic is the belief that vagueness invalidates the law of excluded middle. But if there is reason to believe that vagueness makes $p \vee {\sim}p$ not universally true, there is equal reason to believe that second-order vagueness makes $+p \vee {\sim}{+p}$ not universally true.

Similar arguments can be given without appeal to the special operator +. Consider, for example, the following disjunction:

(*) $(p{\equiv}q) \vee (p{\equiv}r) \vee (p{\equiv}s) \vee (q{\equiv}r) \vee (q{\equiv}s) \vee (r{\equiv}s).$

Now the trivial equivalences $p \equiv p$, $q \equiv q$, $r \equiv r$ and $s \equiv s$ should be true. Indeed, by (b), the three-valued logician is committed to the possibility of defining $\equiv$ in a way which has that effect. By (a) above, p, q, r and s take values from a set of three, so at least two of them will take the same value. Hence, by truth-functionality, at least one disjunct in (*) is true. But now let p be clearly true, q clearly false, r clearly a borderline case and s neither clearly true nor clearly false nor clearly a borderline case (an example of second-order vagueness). Which disjunct of (*) is true?

Such cases tell against (*) as much as borderline cases tell against the law of excluded middle.

The answer is not to replace three-valued logic by, for example, seventeen-valued logic. The foregoing arguments against three-valued logic can be generalized to arguments against *n*-valued logic for any finite number *n* greater than two. Moreover, any such choice of *n* looks arbitrary. More recent applications of many-valued logic to vague languages have therefore tended to use infinitely many values. Whether this move gets to the root of the problem will be seen below.

4.7 CONTINUUM-VALUED LOGIC: A RATIONALE

Imagine a patch darkening continuously from white to black. At each moment during the process the patch is darker than it was at any earlier moment. Darkness comes in degrees. The patch is dark to a greater degree than it was a second before, even if the difference is too small to be discriminable by the naked eye. Given that there are as many moments in the interval of time as there are real numbers between 0 and 1, there are at least as many degrees of darkness as there are real numbers between 0 and 1, an uncountable infinity of them. Such numbers can be used to measure degrees of darkness. Now at the beginning of the process, the sentence 'The patch is dark' is perfectly false, for the patch is white. At the end, the sentence is perfectly true, for the patch is black.[22] In the middle, the sentence is true to just the degree to which the patch is dark. Truth comes in degrees. For 'The patch is dark' to be true just is for the patch to be dark; for 'The patch is dark' to be true to a certain degree just is for the patch to be dark to that degree. Even if we cannot discriminate between all these degrees in practice, we have made the truth of our sentence depend on a property which does in fact come in such degrees. Thus there are at least as many degrees of truth as there are degrees of darkness, and so at least as many as there are real numbers between 0 and 1, an uncountable infinity of them. Such numbers can be used to measure degrees of truth. So the thought goes.

Any treatment of the case within a semantics of finitely many values must divide the continuous process of darkening into a finite sequence of discrete segments, corresponding to the different values through which the sentence ‘The patch is dark’ is supposed to pass. Any particular choice of segments seems arbitrary. A continuum of degrees of truth is attractive because it promises to avoid such arbitrary choices. The appearance of continuity can be taken at face value.

If the dichotomy of true and false gives way to a continuum of degrees of truth, should two-valued logic give way to infinitely many-valued logic, with degrees of truth as truth-values? The obvious candidates for those values are the real numbers between 0 and 1 inclusive, the closed interval [0, 1]. Of course, examples cannot show the proposal to be correct, but they do make it a natural one to explore. It has at least the virtues of simplicity and boldness. If it is even approximately correct, the best policy will be to develop it as a clear working model whose over-simplifications and over-precision can be removed later. If it is not even approximately correct, that can best be shown by an attempt to make it work. For the time being we may therefore work with numerical degrees of truth as though they were unproblematic; the problems will emerge soon enough.

4.8 CONTINUUM-VALUED LOGIC: TRUTH-TABLES

Continuum-valued logic assumes more than a continuum of degrees of truth. It makes the further assumption that the main sentence functors satisfy generalizations of truth-functionality to those degrees. When the application of such functors builds a complex sentence out of simpler ones, the degree of truth of the former is held to be determined as a function of the degrees of truth of the latter. What degree-tables are appropriate?

Some notation will help. The degree of truth of the sentence p is $[p]$, which is assumed to be a real number between 0 and 1. $[p] = 1$ when p is perfectly true; $[p] = 0$ when p is perfectly false; $[p] \leqslant [q]$ when q is at least as true as p.

Conjunction may be considered first. How does the degree of truth of ‘The patch is dark and my head hurts’ depend on the degree of truth of ‘The patch is dark’ and the degree of truth of ‘My head hurts’? The task is to express $[p \,\&\, q]$ as a function of $[p]$ and $[q]$.

Three assumptions about the degrees of truth of conjunctions seem plausible. First, repetition of a conjunct does not lower degree of truth; ‘The patch is dark and the patch is dark’ is as true as ‘The patch is dark’:

($\&_1$) $[p] \leqslant [p \,\&\, p]$.

Second, any conjunct is at least as true as the conjunction; if ‘The patch is dark’ fails to be true to a certain degree, ‘The patch is dark and my head hurts’ also fails to be true to that degree, for the latter claim merely adds something to the former:

($\&_2$) $[p \,\&\, q] \leqslant [P]$ and $[p \,\&\, q] \leqslant [q]$.

Third, if each conjunct is replaced by one at least as true as it, the new conjunction is at least as true as the old one; if the room is at least as dark as the patch and your head hurts at least as much as mine, then ‘The room is dark and your head hurts’ is at least as true as ‘The patch is dark and my head hurts’:

($\&_3$) If $[p'] \leqslant [p]$ and $[q'] \leqslant [q]$ then $[p' \,\&\, q'] \leqslant [p \,\&\, q]$.

Degree-functionality for &, the assumption that the degree of truth of a conjunction is determined by the degrees of truth of its conjuncts, is a consequence of ($\&_3$).

From ($\&_1$), ($\&_2$) and ($\&_3$) one can show that the degree of truth of a conjunction is simply the minimum of the degrees of truth of its conjuncts:

(&) $[p \,\&\, q] = \min \{[p], [q]\}$.

The proof is as follows. Suppose that $[p] \leqslant [q]$. By ($\&_2$), $[p \,\&\, q] \leqslant [p]$. Put $p' = q' = p$ in ($\&_3$); then $[p \,\&\, p] \leqslant [p \,\&\, q]$, so $[p] \leqslant [p \,\&\, q]$ by ($\&_1$). Thus $[p \,\&\, q] = [p]$. Similarly, if $[q] \leqslant [p]$ then $[p \,\&\, q] = [q]$. Either way, (&) holds. If degrees of truth are restricted to 0 and 1, then ($\&_1$), ($\&_2$) and ($\&_3$) all hold for classical conjunction, and (&) summarizes its two-valued truth-table. In this sense the continuum-valued treatment generalizes the classical treatment.

The treatment of conjunction can be generalized to the universal quantifier. ‘Everything is F’ is treated as the conjunction of its instances ‘x is F’ for each member of the domain. The degree of truth of the universally quantified sentence is then the minimum, or more accurately the greatest lower bound, of the degrees of truth of its instances. In an infinite domain, a universally quantified sentence is perfectly false if for each positive real number $\in$ it has an instance true to a degree less than $\in$, even if it has no perfectly false instance.

Disjunction can be given a similar treatment. To ($\&_1$), ($\&_2$) and ($\&_3$) correspond the following three plausible assumptions:

($\vee_1$) $[p \vee p] \leqslant [p]$.

($\vee_2$) $[p] \leqslant [p \vee q]$ and $[q] \leqslant [p \vee q]$.

($\vee_3$) If $[p'] \leqslant [p]$ and $[q'] \leqslant [q]$ then $[p' \vee q'] \leqslant [p \vee q]$.

($\vee_1$) and ($\vee_2$) reverse the ordering in ($\&_1$) and ($\&_2$) respectively because the addition of a disjunct subtracts from the original claim rather than adding to it. ($\vee_3$) replaces & by $\vee$ in ($\&_2$), and has the same rationale. From ($\vee_1$), ($\vee_2$) and ($\vee_3$) one can show that the degree of truth of a disjunction is simply the maximum of the degrees of truth of its disjuncts:

($\vee$) $[p \vee q] = \max\{[p], [q]\}$.

The proof is like that of (&).

The treatment of disjunction can be generalized to the existential quantifier. ‘Something is *F*’ is treated as the disjunction of its instances. The degree of truth of the existentially quantified sentence is then the maximum, or more accurately the least upper bound, of the degrees of truth of its instances. In an infinite domain, the existentially quantified sentence is perfectly true if for each positive real number ε it has an instance true to a degree greater than $1 - \varepsilon$, even if it has no perfectly true instance.

Negation, the conditional and the biconditional are less straightforward. Their degree tables, unlike those for conjunction and disjunction, cannot be derived from considerations about the comparative ordering of degrees of truth alone; other mathematical relations must be taken into account. Consider the biconditional. If p and q have exactly the same degree of truth, then $p \equiv q$ should be perfectly true. If p is perfectly true and q perfectly false, or vice versa, then $p \equiv q$ should be perfectly false. When the truer component decreases in degree of truth and the less true component increases, the biconditional should increase in degree of truth, but at what rate? The simplest assumption is that the degree of truth of the biconditional is perfect truth minus the difference between the degrees of truth of its components:

(≡) $[p \equiv q] = 1 + \min\{[p], [q]\} - \max\{[p], [q]\}$.

As usual, the restriction of degrees of truth to 0 and 1 gives back the standard two-valued truth-table for ≡ from (≡).

The conditional $p \supset q$ can be defined in terms of the biconditional and conjunction, by $p \equiv (p \,\&\, q)$. This enables us to calculate $[p \supset q]$ as a function of $[p]$ and $[q]$, using (≡) and (&). For

$$\begin{aligned}[p \supset q] &= [p \equiv (p \,\&\, q)] \\ &= 1 + \min\{[p], [p \,\&\, q]\} - \max\{[p], [p \,\&\, q]\} \\ &= 1 + \min\{[p], \min\{[p], [q]\}\} - \max\{[p], \min\{[p], [q]\}\}.\end{aligned}$$

Since $\min\{[p], \min\{[p], [q]\}\}$ simplifies to $\min\{[p], [q]\}$ and $\max\{[p], \min\{[p], [q]\}\}$ to $[p]$, we have

(⊃) $[p \supset q] = 1 + \min\{[p], [q]\} - [p]$.

Thus

if $[p] \leqslant [q]$ then $[p \supset q] = 1$;

if $[q] \leqslant [p]$ then $[p \supset q] = 1 + [q] - [p]$.

In other words, if the consequent is at least as true as the antecedent, then the conditional is perfectly true; if the consequent drops below the antecedent in degree of truth, the conditional is less than perfectly true to the extent of that drop. We could also have defined $p \supset q$ in terms of the biconditional and disjunction, by $(p \vee q) \equiv q$. From (≡) and (∨) we could then have calculated that $[p \supset q] = 1 + [q] - \max\{[p], [q]\}$, which is mathematically equivalent to (⊃). Alternatively, we could have begun with (⊃) and defined $p \equiv q$ as $(p \supset q) \,\&\, (q \supset p)$; we could then have recovered (≡) from (⊃) and (&).

For negation, it helps to suppose that the language contains an absurd sentence ⊥ (e.g. '2 = 3' or 'Pigs can fly') such that $[\bot] = 0$. We can thendefine $\sim p$ as $p \equiv \bot$ or as $p \supset \bot$ ('If he apologized then pigs can fly' amounts to 'He did not apologize'). From (≡) or (⊃) we then have

(~) $[\sim p] = 1 - [p]$.

This is in any case the most natural table for ~; it ensures that, the truer a sentence is, the less true its negation, and vice versa. Once again, the restriction of degrees of truth to 0 and 1 gives back the standard two-valued truth-tables for $\supset$ and ~ from ($\supset$) and (~).

The foregoing tables were devised by Lukasiewicz for other purposes. He made 1 the only designated value; thus a formula is valid just in case it comes out perfectly true on any assignment of degrees of truth to its atomic constituents, and a form of inference is valid just in case its conclusion is perfectly true on any assignment on which all its premises are perfectly true.[23] Unlike Halldén and Körner, Lukasiewicz allows a formula to be perfectly true when all its atomic components have intermediate degrees of truth, so that some formulas are valid. For example, $[p]$ and $[\sim\sim p]$ are always equal, so $[p \equiv \sim\sim p]$ is always 1, so $p \equiv \sim\sim p$ is valid. Much that is classically valid is also valid on Lukasiewicz's tables, and the converse always holds. However, the law of excluded middle fails. If p is neither perfectly true nor perfectly false, then $p \vee \sim p$ is not perfectly true. If p is half-true ($[p] = 1/2$) then so is $p \vee \sim p$ ($[p \vee \sim p] = \min\{[p], [\sim p]\} = \min\{[p], 1 - [p]\} = 1/2$). The best that can be said for $p \vee \sim p$ is that it is never less than half-true.

The failure of excluded middle may seem natural enough in borderline cases. More disturbing is that the law of non-contradiction fails in the same way. $\sim(p \ \& \sim p)$ always has the same degree of truth as $p \vee \sim p$, and thus is perfectly true only when p is either perfectly true or perfectly false. When p is half-true, so are both $p \ \& \sim p$ and $\sim (p \ \& \sim p)$.

Not all two-valued tautologies are as much as half-true on the infinite tables. For example, when p is half-true, the two-valued contradiction $p \equiv \sim p$ is perfectly true (because $\sim p$ is true to exactly the same degree as p), so the two-valued tautology $\sim(p \equiv \sim p)$ is perfectly false. That a formula is a two-valued tautology tells one by itself nothing whatsoever about its degree of truth on the infinite tables. Equally, that a formula is a two-valued contradiction tells one nothing about its degree of truth on those tables.

The continuum-valued framework legitimizes a vast range of other functors. For example, a non-standard conditional $\rightarrow$ and the corresponding biconditional $\leftrightarrow$ can be defined by

$$(\rightarrow) \qquad [p \rightarrow q] \;= 1 \text{ if } [p] \leqslant [q]$$
$$= 0 \text{ otherwise,}$$
$$(\leftrightarrow) \qquad [p \leftrightarrow q] \;= 1 \text{ if } [p] = [q]$$
$$= 0 \text{ otherwise.}$$

Given the framework of classical mathematics, $p \rightarrow q$ and $p \leftrightarrow q$ are always either perfectly true or perfectly false. $[p \rightarrow q]$ and $[p \leftrightarrow q]$ are not continuous functions of $[p]$ and $[q]$.

A functor > can in turn be defined in terms of the non-standard conditional to mean something like 'more than', as in 'It is raining more than it is snowing'. $p > q$ is equivalent to $\sim(p \rightarrow q)$:

$$(>) \qquad [p > q] = 1 \text{ if } [p] > [q]$$
$$= 0 \text{ otherwise.}$$

When degrees are restricted to 0 and 1, $p > q$ is equivalent to p & $\sim q$.

Continuum-valued semantics permits the definition of simple constructions with no idiomatic equivalent. For example, if α is any real number between 0 and 1, a constant sentence C_α can be defined to have degree of truth α in all circumstances:

$$(C_\alpha) \qquad [C_\alpha] = \alpha.$$

Then $C_\alpha \rightarrow p$ says that p is true to degree at least α, for it is perfectly true if that is so and perfectly false otherwise. For example, $C_{1/2} \rightarrow p$ says that p is at least half-true; $p > C_{1/2}$ says that p is more. than half-true (it is 'on balance true'). Similarly, $p \rightarrow C_\alpha$ says that p is true to degree at most α, for it is perfectly true if that is so and perfectly false otherwise. Thus $p \leftrightarrow C\alpha$ says that p is true to degree exactly α. It is convenient to have a special symbol J_α for that functor:

$$(J_\alpha) \qquad [J_\alpha p] \;= 1 \text{ if } [p] = \alpha$$
$$= 0 \text{ otherwise.}$$

In particular, J_1 is a kind of determinacy operator: $J_1 p$ is perfectly true if p is perfectly true and perfectly false otherwise.

One can also define an 'averaging' operator , for which the degree of truth of a compound sentence is the average of the degrees of truth of its constituent sentences:

($\sqcap$) $[p \sqcap q] = 1/2([p] + [q])$

$\sqcap$ lacks a feature possessed by all the other one- and two-place sentence functors considered here, that the compound is true to degree 0 or 1 when each constituent is true to degree 0 or 1. If *p* is perfectly true and *q* perfectly false, or vice versa, then $p \sqcap q$ is half-true. Thus even when *p* and *q* are perfectly precise, $p \sqcap q$ can have an intermediate degree of truth. This makes $\sqcap$ hard to interpret intuitively. It is somehow the average of conjunction and disjunction, but how can there be such an operator? '2 = 2$\sqcap$2 = 3' is half-true. Any relevant vagueness in it must come from $\sqcap$ itself: perhaps it is a vague mixture of conjunction and disjunction. Now an utterance that was vague between a conjunction and a disjunction would ordinarily be regarded as ill-defined, obscure and confused. Yet $\sqcap$ is perfectly well-defined within the terms of continuum-valued semantics by ($\sqcap$). To accept the general form of continuum-valued semantics is to be committed to the intelligibility of operators such as $\sqcap$. Unfortunately, being committed to the intelligibility of something is not the same as understanding it.[24]

4.9 FUZZY SETS AND FUZZY LOGIC

Lukasiewic's work had no immediate effect on the study of vagueness. However, formally similar proposals were made in the course of later attempts to develop the notion of a *fuzzy set.* As sets are usually conceived, membership of them is an all-or-nothing matter. In effect, the law of excluded middle is assumed. When a set is defined by a vague predicate ('the set of heaps'), this assumption may seem suspect. One alternative is to develop a theory of sets membership of which is a matter of degree. An early such attempt was made by Abraham Kaplan and Hermann Schott in 1951, with applications to empirical science in mind. They measured degree of membership of empirical classes by real numbers between 0 and 1, and defined corresponding notions of intersection, union, complementation and subset. Although Kaplan and Schott's paper fell on stony ground, the electrical engineer Lofti Zadeh had more success fourteen years later with similar ideas, developed independently and with greater mathematical elaboration in his 1965 paper 'Fuzzy sets' and its many successors. Publishing on fuzziness grew rapidly in the 1970s to become a middle-sized but largely self-contained industry, with from 1978 its own journal (the *International Journal of Fuzzy Sets and Systems*).

Zadeh's work was intended to have applications to problems in computing, such as pattern recognition; this may account for its popularity. Consider, for instance, the task of programming a computer to read handwriting: a scrawled shape may fit the patterns 'm', 'n' and 'w' to varying degrees. Again, if human users are to supply a computer-controlled system with vaguely formulated information or instructions ('The station is about a kilometre further on', 'Turn left soon after passing a large school'), the computer needs a framework for handling vagueness, and the theory of fuzzy sets has been presented as a good candidate. One can even buy fuzzy washing-machines. However, all such applications are made within a framework of classical logic and mathematics. For example, a fuzzy set is simply a classical function from a domain to the interval [0, 1]: the function mapping each member of the domain to its degree of membership of the set. Thus the success of such applications, if any, is far from showing the classical framework to be inadequate. Moreover, much of the detail of fuzzy set theory seems to do little work in the applications. Its basic ideas may nevertheless be sketched.

Standard set-theoretic notions are adapted to sets with degrees of membership between 0 and 1. The intersection of X and Y is a set $X \cap Y$ of which a thing's degree of membership is the minimum of its degrees of membership of X and of Y. The union of two sets X and Y is a set $X \cup Y$ of which a thing's degree of membership is the maximum of its degrees of membership of X and of Y. Similarly, if a thing is a member of X to degree α, it is a member of the complement of X to degree $1 - \alpha$ (for a given domain of discourse). These definitions are the set-theoretic equivalent of Lukasiewicz's tables for conjunction, disjunction and negation.

One might expect the conditional and biconditional to correspond to the subset and equality relations on sets. However, natural definitions of the latter pair for fuzzy sets are that $X \subseteq Y$ just in case everything's degree of membership of Y is at least as great as its degree of membership of X and that $X = Y$ just in case everything's degree of membership of Y is exactly the same as its degree of membership of X.

The holding of these relations is not a matter of degree in the same way as membership of X or Y. They correspond to the non-standard conditional and biconditional $\rightarrow$ and . One could define relations like subsethood and equality between fuzzy sets that would correspond better than $\subseteq$ and $=$ to $\supset$ and $\equiv$. However, a difference would remain. Just as conjunctions, disjunctions and negations formed out of sentences in a language are themselves sentences in that language, so are conditionals and biconditionals; but while intersections, unions and complements of subsets of a domain are themselves subsets of that domain, subsethood and equality are not: they are relations between such subsets.[25]

Zadeh's original paper concerns fuzzy sets, not fuzzy logic. It is nevertheless natural to suppose that if membership of the set of heaps is a matter of degree, then so too is the truth of 'That is a heap'. Moreover, fuzzy intersection, union and complementation correspond to continuum-valued conjunction, disjunction and negation. The development of fuzzy logic out of fuzzy set theory was soon initiated by Joseph Goguen.[26]He does not require the truth set to be the interval [0, 1], but allows it to be any set on which a certain kind of abstract mathematical structure is defined; thus degrees of truth need not be numerical. This generalization will be discussed in Section 4.13. The choice of truth set is supposed to depend on the details of the case at hand. For many purposes, [0, 1] is perfectly adequate. Even then, Goguen allows the definitions of the functors to differ from those given above, provided that they meet various structural constraints. He suggests, for example, that the degree of truth of a conjunction be defined as the product of the degrees of truth of its conjuncts. Repetition would then make a logical difference; if 'The patch is dark' were two-thirds true, 'The patch is dark and the patch is dark' would be only four-ninths true. The operations given above seem more natural, and will be used in the rest of this chapter; they meet Goguen's structural constraints. No attempt will be made to describe the often unrewarding details of fuzzy logic, but some main ideas relevant to vagueness will be picked out.

4.10 DEGREE-THEORETIC TREATMENTS OF SORITES PARADOXES

Fuzzy logic permits a smooth treatment of sorites paradoxes. Let p_n stand for 'A man with *n* hairs on his head is bald'. A sorites argument may be laid out thus:

$$\begin{array}{l} p_0 \\ p_0 \supset p_1 \\ p_1 \supset p_2 \\ . \\ . \\ . \\ \underline{p_{99{,}999} \supset p_{100{,}000}} \\ p_{100{,}000} \end{array}$$

The argument reaches its conclusion by 100,000 steps of *modus ponens*. p_0, the first premise, is perfectly true; $p_{100{,}000}$, the conclusion, is assumed to be perfectly false. The degree-theoretic diagnosis is that, as *n* increases from 0 to 100,000, the degree of truth of p_n decreases by imperceptible steps. For simplicity, we can realize this idea by making each p_n true to degree $1 - (n/100{,}000)$ $(0 \leqslant n \leqslant 100{,}000)$. Thus the drop in degree of truth from pn to p_n+1 is just 1/100,000; conversely, it is the increase in the degree to which a man is bald when one hair falls out. Thus in each conditional premise of the argument, the antecedent is truer than the consequent by just 1/100,000. It follows by ($\supset$) that each conditional premise is true to degree 99,999/100,000. At each intermediate step of the argument, *modus ponens* yields a conclusion, P*n*+1, true to degree (99,999 – *n*)/100,000 from premises p_n and $p_n \supset p_n$+1, true to degrees (100,000 – *n*)/100,000 and 99,999/100,000 respectively. The truth of the intermediate conclusions goes down in minute steps. There is nothing like a cut-off point for baldness, for all the steps are equal.[27] The attraction of the conditional premises has been explained, for each is almost perfectly true, and its negation almost perfectly false.

Is *modus ponens* invalid on this account? That depends on the definition of validity. If validity is preservation of perfect truth, then *modus ponens* is valid; if *q* and $q \supset r$ are true to degree 1, so is *r*. In this sense, the sorites argument is valid, but a valid argument can have *almost* perfectly true premises and a perfectly false conclusion. On the other hand, if validity is preservation of degree of truth, then *modus ponens* is not valid; if q and $q \supset r$ are true to degree at least 99,999, r may be true only to degree 99,998. *A fortiori*, the sorites argument is invalid in the latter sense. One can introduce a fuzzy notion of validity: an argument is valid to degree at least α just in case when all its premises are true to degree at least β, then its conclusion is true to degree at least $\beta - (1 - \alpha)$. However, *modus ponens* is not even valid to a high degree in this sense, for it permits a drop of up to 1/2 from premises to conclusion. If q is half-true and r perfectly false, the step of *modus ponens* from q and $q \supset r$ to r has half-true premises and a perfectly false conclusion. Thus *modus ponens* would be at best half-valid, and its apparent validity could not be explained by its near-validity. The definition of validity as preservation of perfect truth therefore seems to fit our intuitions better. The sorites argument is valid, but a small degree of falsity in its conditional premises produces a high degree of falsity in its conclusion.[28]

Is $\supset$ transitive on this account? The conditionals $p_0 \supset p_1, p_1 \supset p_2, \ldots,$ $p_{99,999} \supset p_{100,000}$ count as almost perfectly true, and $p_0 \supset p_{100,000}$ as perfectly false. However, if the former conditionals were perfectly true, so would be the latter. Chaining conditionals together preserves perfect truth but not almost perfect truth.

4.11 COMPARATIVES AND MODIFIERS

Fuzzy semantics has been applied to a variety of constructions in natural languages. If it provides an illuminating account of them, that is an argument in its favour.

Comparatives form a central case, for they were used to introduce the idea of degrees of truth. ‘The patch is dark’ is truer than it was because the patch is darker than it was. The general task is to give the semantics of ‘*F*er than’ in terms of the semantics of the adjective ‘*F*’. How else could one explain the fact that a speaker who has mastered the comparative construction and understands ‘*F*’ is in a position to understand ‘*F*er than’ without further training?[29] If the degree to which ‘*x* is *F*’ is true is the degree to which *x* is *F*, then *x* is *F*er than *y* if and only if ‘*x* is *F*’ is truer than ‘*y* is *F*’. One might therefore attempt to state the truth condition of the comparative sentence thus: ‘*x* is Fer than *y*’ is true if and only if ‘*x* is *F*’ is truer than ‘*y* is *F*’. Now that is not yet quite what is wanted, for if degree of truth is to be the key notion in the recursive semantics, we need to say what *degree* of truth the comparative sentence has. However, there is a simple suggestion: 'x is Fer than y' is perfectly true if 'x is F' is truer than 'y is F', and perfectly false otherwise.[30] The requirement that comparative sentences be either perfectly true or perfectly false is not wholly unnatural. 'Is x darker than y?' is usually a more precise question than 'Is x dark?'. Of course, not even the former is perfectly precise; unclarities may arise when x or y is not uniform in darkness. Such difficulties are discussed in Section 4.12.

Formally, the comparative might be defined in terms of the 'more than' functor > introduced in Section 4.8. 'x is Fer than y' is equivalent to $Fx > Fy$, i.e. 'x is F more than y is F'. Similarly, 'x is at least as F as y' is definable as $F_y \rightarrow F_x$, i.e. as 'If y is F then x is F' with the non-standard construal of the conditional.

Related to the comparative are modifiers such as 'very' and '-ish'. They have been given fuzzy semantics such as the following. When the patch is dark to an intermediate degree α, it is very dark to a degree less than α and darkish to a degree greater than α. When the patch is dark to degree 0, it is also very dark to degree 0, and darkish to degree 0. When it is dark to degree 1, it is also very dark to degree 1, and darkish to degree 1 (if 'darkish' is taken not to exclude 'dark'). To achieve these effects, one can define the degree of truth of 'x is very F' as the square of the degree of truth of 'x is F', and the degree of truth of 'x is Fish' as its square root. On this simple account, 'very Fish' is equivalent to 'F'. Other constructions have quite differently shaped graphs. 'Semi-' has a bell-shaped curve. 'The patch is semi-dark' is perfectly true when 'The patch is dark' is semi-true, and perfectly false when the latter is perfectly true or perfectly false. Corresponding equations are again easy to find. This kind of treatment has been extended to many other adverbial modifiers: 'quite', 'rather', 'somewhat', 'more or less', 'highly', 'fairly', 'slightly', 'extremely' and so on.[31]

The recursive semantics also generates interpretations of arbitrarily complex combinations of such modifiers. When 'The patch is dark' is true to degree x, for example, 'The patch is not very very dark' is true to degree $1 - x^4$. Many of the combinations so interpreted are of dubious standing in natural language: consider 'very not very dark', 'darkishish', 'extremely rather dark' and 'highly dark'. The semantics certainly involves much artificiality, but it does not follow that it is worthless. If the judgements of native speakers confirmed the predictions of the fuzzy semantics as to which arguments in natural language are valid, that would constitute some evidence that the semantics was on the right lines. Its defenders might hope to find a way of removing the artificiality whilst retaining the successes (easier said than done; see Section 4.13). Fuzzy semantics does at least validate some central inferences judged valid by native speakers. It makes '*F*er than' a transitive asymmetric relation, and the argument from '*x* is *F*er than *y*' and '*y* is *F*' to '*x* is *F*' valid. Again, suitable choices of fuzzy semantics will validate the argument from '*x* is more or less the same height as *y*' and '*x* is tall' to '*x* is fairly tall'.

Unfortunately, it is not obvious that the notion of degree of truth appropriate for the semantics of comparatives and other modifiers is also appropriate for the analysis of vagueness. We can make some sense of 'truer than' by using equivalences of the form: '*x* is *F*' is truer than '*y* is *F*' if and only if *x* is *F*er than *y*. Of course, that form is relevant only to a very limited range of comparisons. It does not tell us when 'Snow is white' is truer than 'Snow is cold', nor does it apply to complex sentences. For full generality, we need equivalences of the form: '*p*' is truer than '*q*' if and only if *p* more than *q*. Let it be granted that such equivalences do give us some grip on 'truer than'. What has it to do with vagueness?

Let *x* be the tallest person in the world, and *y* the second tallest. Both are clearly tall (people). Neither is remotely a borderline case for 'tall'. '*x* is tall' and '*y* is tall' are quite straightforwardly true. Nevertheless, *x* is taller than *y*; by the equivalence above, '*x* is tall' is truer than '*y* is tall'. Thus although '*y* is tall' is clearly and straightforwardly true, it has a less than maximal degree of truth.[32] To change the example, an angle is acute if and only if it is less than a right angle. Thus both 30° and 60° are clearly acute; both '30° is acute' and '60° is acute' are clearly true. 'Acute' is precise in all relevant respects. Nevertheless, 30° is acuter than 60°; by the equivalence above, '30° is acute' is truer than '60° is acute'. If such uses of 'truer' make sense, their connection with the phenomenon of vagueness is quite unclear.

Very little is needed to evoke comparative judgements. We can stipulate that the jane months are January, February, March, April, May and June. Which is janer, April or June? The obvious answer is April, presumably because it is further than June from the non-jane months on the salient ordering of the months. No vagueness in ‘jane’ was required for this result. Such comparative judgements can be evoked by any predicate associated with a dimension of similarity, however precise the predicate.[33]

If the notion of degree of truth is explained in comparative terms, then the occurrence of degrees of truth between perfect truth and perfect falsity in no way implies the occurrence of vagueness. Something can be clearly true or clearly false without being perfectly true or perfectly false in the relevant sense. This makes trouble for the degree theorist’s account of sorites paradoxes. On that account, any class of sentences capable of the appropriate range of intermediate degrees of truth can be used to construct a sorites paradox. If some sentences are capable of those degrees without relevant vagueness, then it should be possible to use them to construct sorites paradoxes without relevant vagueness. ‘Acute’ should be sorites-susceptible, for instance. But an argument of sorites form for ‘acute’ would not be genuinely paradoxical; one could easily detect the false step. An account of sorites paradoxes applicable to the non-paradox for ‘acute’ has failed to locate the source of the paradoxes. Thus the degree theorist’s account of sorites paradoxes is subverted by the attempt to explain degrees of truth in comparative terms. How else could they be explained?

The semantics of comparatives has less to do with the problem of vagueness than one might suppose.[34] The best prospects for a theory of degrees depend on abandoning the attempt to use it as a theory of vagueness.

For purposes of exposition, Sections 4.12–4.15 will ignore the problem just raised: not because it is not important, but because the problems they consider are independent of it.

4.12 VAGUE DEGREES OF TRUTH

The analysis of vagueness in terms of continuum-valued logic faces a further problem: the phenomenon of higher-order vagueness.

It is important to be clear about the nature of the difficulty. Consider the statement:

($\#_1$) 'It is wet' is true to a degree greater than 0.729.

Let the context of ($\#_1$) be a casual remark about the weather. The problem is not that ($\#_1$) is too precise to describe such a situation correctly. Rather, ($\#_1$) is extremely vague. In many contexts it is neither clearly true nor clearly false. Attempts to decide it can founder in just the way characteristic of attempts to decide ordinary vague statements, such as 'It is wet' in borderline cases. The mathematical terms in ($\#_1$) may be precise, but the notion of the degree of truth of a sentence is not a mathematical one. It represents an empirically determined mapping from sentences in contexts to real numbers. Even if statistical surveys of native speaker judgements were relevant to deciding ($\#_1$), the results would be vague. It would often be unclear whom to include in the survey, and how to classify the responses. The problem is that the vagueness of ($\#_1$) goes unacknowledged.

The problem would not be solved by the use of non-numerical degrees of truth. Numerical degrees are problematic, but for different reasons (Section 4.13). Consider in place of ($\#_1$) a purely comparative statement:

($\#_2$) 'It is wet' is truer than 'It is cold'.

($\#_2$) is vague in much the same way as ($\#_1$). In many contexts it is neither clearly true nor clearly false, attempts to decide it can founder in just the way characteristic of attempts to decide ordinary vague statements in borderline cases, and so on. What needs to be acknowledged is the vagueness of both ($\#_1$) and ($\#_2$).

Why should the vagueness of ($\#_1$) and ($\#_2$) be hard to acknowledge? If a vague language requires a continuum-valued semantics, that should apply in particular to a vague meta-language. The vague meta-language will in turn have a vague meta-meta-language, with a continuum-valued semantics, and so on all the way up the hierarchy of meta-languages. The details would no doubt be very complex, but is there any difficulty in principle?

There is a problem. The many-valued semantics invalidates classical logic. Thus if the meta-language is to be given a many-valued semantics, classical reasoning is not unrestrictedly valid in the meta-language. The problem is not just notional. Consider, for example, a system with the conditional $\rightarrow$. If classical logic is valid in the meta-language, one may reason as follows. $\rightarrow$ is completely defined by the rule ($\rightarrow$) that $p \rightarrow q$ is perfectly true if q is at least as true as p and perfectly false if not; one or other condition holds by the law of excluded middle. Thus $p \rightarrow q$ is either perfectly true or perfectly false, so by (~) and ($\vee$) $(p \rightarrow q) \vee \sim (p \rightarrow q)$ is perfectly true. Although use of excluded middle in the meta-language does not enable one to establish the unrestricted validity of excluded middle in the object-language, this specific use of the former does enable one to establish the validity of a special case of the latter. An example of the meta-linguistic assumption for a natural language would be

(#$_3$) Either 'It is wet' is at least as true as 'It is cold' or 'It is wet' is not at least as true as 'It is cold'.

The argument would use (#$_3$) to validate this object-language disjunction:

(#$_4$) Either it is at least as wet as it is cold or it is not at least as wet as it is cold.

The problem is that if the original objection to the law of excluded middle is cogent at all, it applies just as much to (#$_3$) and (#$_4$) as to other instances. On the many-valued approach, a simple instance of excluded middle such as 'It is wet or it is not wet' is invalidated by cases in which it is neither clearly wet nor clearly not wet. But sometimes it is neither clearly at least as wet as it is cold nor clearly not at least as wet as it is cold, and 'It is wet' is neither clearly at least as true as 'It is cold' nor clearly not at least as true as 'It is wet'. By parity of reasoning, such cases should invalidate (#$_3$) and (#$_4$). There is no *more* reason to think that one or other disjunct, we know not which, is perfectly true in this case than in any other.[35]

If the degree of truth required of the theorist's assertions in the meta-language were lowered from perfect truth to half-truth, (#$_3$) would be assertible. However, this move would not permit the unrestricted use of classical logic in the meta-language, for it would debar the theorist from unrestricted use of (for example) disjunctive syllogism: if $p \vee q$ and $\sim p$ are half-true, it does not follow that q is at least half-true.

Defenders of continuum-valued logic have been reluctant to acknowledge higher-order vagueness and abandon the use of classical logic in the meta-language.[36] Of course, one can study precise mappings of formulas to real numbers as a purely mathematical exercise, and a classical meta-logic will then be quite unobjectionable, but the results cannot be assumed to apply to the vague mappings characteristic of empirical semantics. More to the point, a degree theorist could regard the use of a classical meta-logic in giving a continuum-valued semantics for a natural language as a convenient over-simplification. The resulting description would not be completely correct, on this view, but might be truer than that given by two-valued semantics. What has just been seen is that the description would be incorrect in its account of the logic of the natural language. To do better, degree theorists must use a non-classical meta-logic.

On the degree-theoretic account, what is an appropriate logic for a vague language? It should have at least this feature: when combined in the meta-language with an appropriate degree-theoretic semantics for the object-language, it should permit one to prove its validity as a logic for the object-language. This constraint is not perfectly precise, for it is not always clear when the logic of the meta-language can be treated as 'the same' as the logic of object-language. However, the constraint is not vacuous either, for classical logic clearly fails it. If one combines a classical logic in the meta-language with a continuum-valued semantics for the object-language, one can prove that classical logic is not valid for the object-language. Unfortunately, it is not clear what logics do meet the constraint. One could devise an *ad hoc* logic to meet it, with a restricted version of the law of excluded middle corresponding to the assumption that all vagueness is first order, but such an assumption has just been seen to be unmotivated.[37]

In practice, research on many-valued logic uses a classical meta-logic, and therefore yields no answer to our question. It may well be that a logic weak enough to be consistent with the philosophical motivation for applying many-valued semantics to vague languages would be so weak as to make the derivation of interesting meta-logical results impossible. In any case, it is just an illusion that many-valued logic constitutes a well-motivated and rigorously worked out theory of vagueness. The rigorous work on many-valued logic has embodied assumptions inconsistent with any cogent philosophical motivation from the problem of vagueness.

4.13 NON-NUMERICAL DEGREES OF TRUTH

The use of numerical degrees of truth may appear to be a denial of higher-order vagueness, for numbers are associated with precision. However, the appearance is deceptive. A degree theorist can and should regard the assignment of numerical degrees of truth to sentences of natural language as a vague matter. If the meta-language requires a non-classical logic, the degree theorist will concede that anyway, whether or not degrees are numerical.

The details would need to be worked out with care. Let $[p]$ be the numerical degree of truth of a vague sentence p of natural language. The degree theorist may hold that in a particular borderline case neither '$0 \leqslant [p] \leqslant 1/2$' nor '$1/2 \leqslant [p] \leqslant 1$' is perfectly true; by the many-valued account of disjunction, the disjunction of these two sentences is itself not perfectly true. Presumably, however, '$0 \leqslant [p] \leqslant 1$' is perfectly true, and therefore not equivalent to the preceding disjunction. Although such an account threatens to be formidably complex, that is not the same thing as incoherence.[38]

Numerical degrees face objections of a different kind. Comparisons often have several dimensions. To take a schematic example, suppose that intelligence has both spatial and verbal factors. If x has more spatial intelligence than y but y has more verbal intelligence than x, then 'x is intelligent' may be held to be truer than 'y is intelligent' in one respect but less true in another. Moreover, this might be held to be a feature of the degrees to which x and y are intelligent: each is in some respect greater than the other. How can two real numbers be each in some respect greater than the other? On this view, degrees are needed that preserve the independence of different dimensions, rather than lumping them together by an arbitrary assignment of weights. The problem lies less in the use of real numbers as degrees than in the use of the standard ordering of real numbers as the only ordering of degrees.[39]

The simplest way to allow multi-dimensional degrees is to treat them as sequences of real numbers between 0 and 1, each representing a different dimension. In the example above, the first number might represent verbal intelligence, the second spatial intelligence and so on. The degree $<\beta 1, \beta 2, \ldots, \beta n>$ would be at least as great as the degree $<\alpha_1$,

$\alpha_2, \ldots, \alpha_n$> just in case $\alpha i \leqslant \beta i$ for each i between 1 and n. Thus neither <0, 1, . . . > nor <1, 0, . . . > would be at least as great as the other. Other notions are generalized in the same way, component by component. <1, 1, . . . , 1> is perfect truth in all respects, <0, 0, . . . , 0> perfect falsity in all respects. If p is true to degree $<\alpha_1, \alpha_2, \ldots, \alpha_n>$ and q to degree $<\beta_1, \beta_2, \ldots, \beta_n>$, p & q is true to degree $<\min\{\alpha_1, \beta_1\}, \min\{\alpha_2, \beta_2\}, \ldots, \min\{\alpha_n, \beta_n\}>$, $p \vee q$ to degree $<\max\{\alpha_1, \beta_1\}, \max\{\alpha_2, \beta_2\}, \ldots, \max\{\beta_n, \beta_n\}>$, $p \supset q$ to degree $<1 + \min\{\alpha_1, \beta_1\} - \alpha_1, 1 + \min\{\alpha_2, \beta_2\} - \alpha_2, \ldots, 1 + \min\{\alpha_n, \beta_n\} - \alpha_n>$ and $\sim p$ to degree $<1 - \alpha_1, 1 - \alpha_2, \ldots, 1 - \alpha_n>$.[40]

That model is not altogether satisfactory. It suggests, for example, that one component of a degree of truth represents the dimension of spatial intelligence. But unless our language permits us to speak of nothing but intelligence, it will contain many sentences to which that dimension is wholly irrelevant. The degree to which 'It is wet' is true has no component for spatial intelligence. When that sentence is incomparable in degree of truth with 'She is intelligent', it is not because one is truer than the other in a first respect and less true in a second; they are just incommensurable. A more radical generalization is needed, one that does not rely on numbers at all. Much of Goguen's work is indeed at the required level of generality.[41]

A natural idea is to take as primitive a notion of non-numerical degrees of truth and of an ordering relation $\leqslant$ on them. For the judgement that this is darker than that does not seem to depend on independent measurement of the darkness of this and of that, and the point extends to the judgement that 'This is dark' is truer than 'That is dark'. Indeed, degrees of truth were introduced by reference to just such purely comparative judgements. $\leqslant$ would be reflexive, anti-symmetric and transitive but not connected: for some degrees α and β, neither $\alpha \leqslant \beta$ nor $\beta \leqslant \alpha$ would hold. '1' and '0' can till be used as names of perfect truth and perfect falsity, on a non-numerical reading. Conjunction and disjunction might be defined much as before:

($\&_4$) $\alpha \leqslant [p \,\&\, q]$ just in case both $\alpha \leqslant [p]$ and $\alpha \leqslant [q]$;

($\vee_4$) $[p \vee q] \leqslant \alpha$ just in case both $[p] \leqslant \alpha$ and $[q] \leqslant \alpha$.

These clauses specify unique degrees of truth for conjunctions and disjunctions, on the assumption that such greatest lower bounds and least upper bounds do exist. The definitions of $\rightarrow$ and $\leftrightarrow$ can also be given a non-numerical reading.

Negation and the conditional are more difficult cases, for they were defined by means of the numerical operation of subtraction. If *p* is true to non-numerical degree α, no sense has been given to saying that ~*p* is true to degree $1 - \alpha$. One might postulate an operation N on non-numerical degrees and set ~*p* true to degree Nα. N would need some appropriate features, such as:

(N_1) $N\alpha \leqslant N\beta$ just in case $\beta \leqslant \alpha$,

(N_2) $NN\alpha = \alpha$.

However, (N_1) and (N_2) do not specify a unique operation in the numerical case; there is no reason to think that they secure uniqueness in the non-numerical case either. The point holds whatever constraints in terms of $\leqslant$ are added to (N_1) and (N_2).[42]To postulate an operation is not to specify one. To define an analogue of $\supset$ for non-numerical degrees is equally difficult. Goguen's work does not solve the problem; he specifies classes of operations with appropriate features, but he does not say *what* member of the class defines non-numerical degrees of truth for negation or implication in a natural language.

One might hope to solve the problem by employing the resources of the object-language within the meta-language. For example, one could say that the negation of a sentence is true if and only if that sentence is not true; more formally:

($\sim_1$) $T(\text{`}{\sim}p\text{'}) \leftrightarrow {\sim}T(\text{`}p\text{'})$.

Here predications of the truth predicate *T* themselves admit of degrees. By definition of $\leftrightarrow$, ($\sim_1$) is perfectly true if *T*('~*p*') and ~*T*('*p*') are true to the same degree as each other; (~*1*) is perfectly false otherwise.[43] Thus ($\sim_1$) equates the degree to which ~*p* is true with the degree to which *p* is not true. Similar homophonic accounts might be given of the other operators.

The trouble with ($\sim_1$) is that it completely fails to mesh with the metalinguistic notions used in the degree-theoretic account of vagueness. For example, the analysis of the sorites paradox speaks of small decreases in degree of truth, a way of speaking about truth quite different from that employed in ($\sim_1$).[44] If such an analysis is to be grounded in a semantic theory of the language to which it is applied, the terms of the analysis must figure in the theory. At the very least, the theory should specify the conditions under which a given sentence *p* is true to a given degree α. Presumably, the conditions under which $\sim p$ is true to degree α should be specified in terms of the conditions under which p is true to some degree. Nothing in the form of ($\sim_1$) even requires $[\sim p]$ to be a function of $[p]$, for it does not say how the degrees of truth of the sentences flanking are calculated.

One might try to specify the degree to which $\sim p$ is true by using negation in the meta-language:

($\sim_2$) $\quad [\sim p] = \alpha$ if and only if $[p] \neq \alpha$.

However, a moment's reflection shows that the degree theorist must reject ($\sim_2$) as incoherent. For by ($\sim_2$) $[\sim p] = \alpha$ for every degree α other than $[p]$, so there could be at most two degrees. It would not help to replace = in ($\sim_2$) by $\leqslant$:

($\sim_3$) $\quad [\sim p] \leqslant \alpha$ if and only if $[p] \leqslant \alpha$.

For ($\sim_3$) implies that $[\sim p] \leqslant 1$ if and only if $[p] \leqslant 1$, which is impossible because both $[\sim p] \leqslant 1$ and $[p] \leqslant 1$, by definition of 1. Similar proposals meet similar fates.

One could define a special negation $\neg$ as follows:

($\neg$) $\quad [\neg p] = 1$ if $[p] \neq 1$
$\quad\quad = 0$ otherwise.

However, $\neg$ lacks the property that the degree of truth of $\neg p$ varies inversely with the degree of truth of *p*; the corresponding function on degrees of truth does not satisfy (N_1). Thus $\neg$ is not a generalization of $\sim$; for many purposes it is not what the degree theorist seeks.

Zadeh has attempted to develop a theory of non-numerical 'linguistic' truth values. These are terms such as *true*, *false*, *not true*, *very true*, *not very true*, *very not true*, and *not very true and not very false*. However, what the theory comes to is a numerical fuzzy semantics for such terms (not all of which are obviously grammatical). For example, he suggests that if *p* is true to degree 0.6, then '*p* is true' might be true to degree 0.3. It thus cannot solve the problem of constructing a degree- theoretic semantics without reliance on numerical degrees.[45]

The generalization of many-valued semantics to non-numerical degrees of truth remains highly problematic. It is easy to specify the kind of structure that might be desirable for such a semantics; the difficulty is in specifying a plausible instance of that structure. Degree theorists have done little to meet the difficulty.

Should degree theorists fall back on numerical degrees? After all, if one makes a statistical survey of native speakers, to find out what proportion of them assent to a given sentence in a given context, the result will be a real number between 0 and 1. It may be unclear just what the right number is, but second-order vagueness has already been accepted. It would follow that a sentence is perfectly true just in case every native speaker assents to it. As it stands, that claim is obviously false. Ignorance and error ensure that unanimity is neither necessary nor sufficient for truth, even when vagueness is not in question. The degree theorist may therefore be driven to speak of the proportion of ideally rational native speakers who would assent in epistemically ideal conditions. Such a theory involves grave commitments. Even if the notions involved make good sense, it is not obvious that for every sentence in every context there are epistemically ideal conditions in which every fact relevant to the truth of that sentence in the original context can be known. Perhaps some matters are essentially unknowable, and perhaps some of them are matters of degree. Degree theorists have not shown it to be otherwise. Moreover, the numbers so produced would not satisfy degree-functionality. If 53 per cent of ideally rational speakers assent to *p* and 29 per cent to ~*p*, it does not follow that 29 per cent assent to *p* & ~*p*, as (&) would require.

The conclusion can be put harshly: semantic theories using numerical degrees of truth are implausible; semantic theories using non-numerical degrees are inchoate.

4.14 DEGREE-FUNCTIONALITY

The central assumption of the degree-theoretic approach to vagueness as considered in this chapter is that the degree of truth of various compounds is a function of the degrees of truth of their components.

Although the objections to this claim of degree-functionality are not new, their force has been insufficiently appreciated.[46]

Conjunction may be taken first. Suppose that p is true to the same degree as q. Thus the first and second conjuncts of p & q match the first and second conjuncts of p & p respectively in degree of truth. By generalized truth-functionality, it follows that p & q is true to the same degree as p & p. Since p & p is true to the same degree as p, p & q is true to the same degree as p. Now imagine someone drifting off to sleep. The sentences 'He is awake' and 'He is asleep' are vague. According to the degree theorist, as the former falls in degree of truth, the latter rises. At some point they have the same degree of truth, an intermediate one. By what has just been argued, the conjunction 'He is awake and he is asleep' also has that intermediate degree of truth. But how can that be? Waking and sleep by definition exclude each other. 'He is awake and he is asleep' has no chance at all of being true. Our man is not in an unclear area between the cases in which the conjunction is true and those in which it is false, for there are no cases of the former kind. If intermediate degrees of truth are a matter of vagueness, they characterize cases in which a sentence is neither clearly correct nor clearly incorrect. Since the conjunction in question is clearly incorrect, it should not have an intermediate degree of truth. It is clearly incorrect, although neither conjunct is; one must be careful to distinguish what can be said of the conjunction from what can be said of each conjunct.[47] Thus degree-functionality fails for conjunction.

The same point can be made with 'He is not awake' in place of 'He is asleep'. At some point 'He is awake' is supposed to be half-true, so 'He is not awake' will be half-true too. Then 'He is awake and he is not awake' will count as half-true. How can an explicit contradiction be true to any degree other than 0?

Intuitions can be confused by the idiomatic use of contradictions such as 'He is and he isn't' to describe borderline cases. Another example may therefore help. Consider the dying moments of James I of England, who was James VI of Scotland. In the circumstances, his death constituted the accession of Charles I to both thrones. James's death was not absolutely instantaneous; its timing was very slightly vague. At some point 'James is alive' and 'James is dead' had the same intermediate degree of truth, according to standard degree theory. Since 'James is King of England' was true to the same degree as 'James is alive' and 'Charles is King of Scotland' to the same degree as 'James is dead', all these sentences had the same degree of truth. But then 'James is King of England and Charles is King of Scotland' had that intermediate degree of truth too, as did 'James is King of England and James is not King of Scotland'. That is absurd. Those sentences were true to degree 0, for there was clearly no difference between England and Scotland in the identity of their king.

The point can still be made if one assumes only approximate equality in degree of truth between the conjuncts, for the degree theorist presumably holds that a small difference in the degree of truth of a conjunct makes only a small difference to the degree of truth of a conjunction. Thus if q is approximately equal to p in degree of truth, then p & q is approximately equal to p & p, and therefore to p, in degree of truth. Suppose, for example, that 'n grains make a heap' is roughly half-true. Then if contextual factors are held constant, '$n + 1$ grains make a heap' will be only slightly truer, so '$n + 1$ grains do not make a heap' will be nearly half-true. Thus 'n grains make a heap and $n + 1$ grains do not make a heap' will count as nearly half-true. That too is absurd; there is no chance at all that n grains are sufficient and $n + 1$ insufficient. Degree-functionality forces the conjunction to be treated as though its application were vague. The degree of truth of a conjunction is not determined by the degrees of truth of its conjuncts.

The case of disjunction is formally similar, although some find the implications of generalized truth-functionality less absurd in this case. If p is true to the same degree as q, it follows by generalized truth-functionality that $p \vee q$ is true to the same degree as $p \vee p$, and therefore as p. Thus if 'He is awake' and 'He is asleep' have the same (or approximately the same) intermediate degree of truth, so has 'He is awake or he is asleep', even if waking and sleep are regarded as complementary 3.n 'He is awake or he is dead', however healthy he is. To vary the example, consider Marc Bloch's account of the mediaeval understanding of the distinction between liberty and servitude as 'a line which was often uncertain and even variable according to the bias of the time or of the class, but one of whose essential characteristics was precisely that of having never allowed that marginal zone which the name of half-liberty suggests with such tiresome persistence'.[48] The marginal zone is just what the degree-theoretic account requires. It does not allow the possibility that someone is clearly either free or unfree without being either clearly free or clearly unfree. To allow that possibility, one must drop the assumption of degree-functionality for disjunction.

The case of the conditional is hardly different. If p is true to the same degree as q, it follows by generalized truth-functionality that $p \supset q$ is true to the same degree as $p \supset p$. Since the latter should count as perfectly true, so should the former. On that treatment, the following conditionals will count as perfectly true in the circumstances envisaged above: 'If he is awake then he is asleep'; 'If he is awake then he is not awake'; 'If James is King of Scotland then Charles is King of England'; and so on. For similar reasons, most degree-theoretic treatments will count 'If n grains make a heap then $n + 1$ grains do not make a heap' as almost perfectly true when its antecedent is about half-true. That is absurd. Degree-functionality fails for the conditional too.

To deny degree-functionality is not to deny that truth is a matter of degree. Conjunctions, disjunctions and conditionals might be true to degrees not determined by the degrees of truth of their components. But the failure of degree-functionality does make it hard to see how degrees could be central to semantic theory. For some kind of truth-functionality surely is what is central to the semantics of the most basic logical operators, and it is not to be described in terms of degrees. Moreover, the introduction of degrees by means of comparatives leaves their connection with vagueness obscure (Section 4.11); formal work on the subject is hard to adapt to higher-order vagueness (Section 4.12) and to non-numerical orderings of degrees (Section 4.13). The attempt to adapt truth-conditional semantics to vague languages by substituting degrees of truth for truth-values has ended in failure.

4.15 APPENDIX: AXIOMATIZATIONS OF CONTINUUM-VALUED LOGIC

One would like a sound and complete system of axioms and rules of inference within which all and only the arguments validated by the continuum-valued semantics could be derived. Whether this is possible depends on both the vocabulary of the formal language and the definition of validity. For example, one may take Lukasiewicz's infinitely valued logic, with 1 as the only designated value and $\supset$ and ~ as the only primitive operators ($A \vee B$ may be defined as $(A \supset B) \supset B$ and A & B as $\sim(\sim A \vee \sim B)$). Then from a few axioms one can derive all and only valid formulas by the rules of *modus ponens* and substitution.[49] However, if the universal and existential quantifiers are added to the language, with many-place predicates to match, then the set of valid formulas is not even recursively enumerable: it cannot be generated by any mechanical procedure.[50] In particular, it cannot be generated by any formal system with a mechanical test for what counts as a proof. Since the quantifiers are central to our thought, this is a serious limitation.

Even in propositional logic, the introduction of new operators causes problems. Consider the constants C_α. They form an uncountable infinity, for α may be any real number between 0 and 1; thus no language capable of being written with a finite alphabet can contain them all. Such a language could contain J_α for each rational number between 0 and 1. It may also be assumed to contain $\rightarrow$. But then a new difficulty arises: validity will not be compact. That is, there will be valid arguments with infinitely many premises whose conclusion is not a valid consequence of any finite subset of those premises. An example is the argument with premises $C_{1/2} \rightarrow p$, $C_{3/4} \rightarrow p$, $C_{7/8} \rightarrow p, \ldots$ and conclusion $C_1 \rightarrow p$. Its premises say that p is true to at least degree 1/2, and to degree at least 3/4, and to degree at least 7/8, . . . , while its conclusion says that p is true to at least degree 1. The conclusion follows from the premises, but not from any finite subset of them. Yet if proofs can be mechanically certified as such, they must be finite objects, and therefore make use of only finitely many premises. For such a proof system to permit the derivation of all and only the valid arguments, validity must be compact. Thus there is no sound and complete system of proofs for arguments in a propositional language containing $C_{1/2}$, $C_{3/4}$, $C_{7/8}, \ldots, C_1$ and $\rightarrow$.[51]

Soundness and completeness results have been published for some systems of fuzzy logic in which provability as well as truth is a matter of degree.[52] Their fundamental proof-theoretic notion is not that the formula A is provable,$\vdash A$, but that A is provable to at least degree α,$\vdash\alpha$ A. A notion of provability to a certain degree from a set of assumptions can also be introduced. *Modus ponens*, for example, then holds in the form: if $\vdash_\alpha A$ and $\vdash_\beta A \supset B$ then $\vdash_{\max\{0,\alpha+\beta-1\}} B$. A soundness theorem (for single formulas) will state that if A is provable to at least degree α then A is true in every model to at least degree α a completeness theorem will state the converse. More ordinary notions of soundness and completeness emerge as the special case in which $\alpha = 1$. The results mentioned in the previous paragraphs therefore limit what can be done on these lines. Sound and complete proof systems can be provided only for fragments of the infinitevalued system based on the quantifiers and functors so far defined.

How great a problem do these limitations pose to continuumvalued semantics? At a theoretical level, they make it less attractive without constituting a decisive objection. There is no *a priori* guarantee that all valid arguments involving a given set of logical concepts can be captured within a formal system. For example, our understanding of the natural numbers essentially involves the second-order concept of all properties of natural numbers as well as the first-order concept of all natural numbers, and is thus appropriately formulated within a system of second-order logic.[53] Yet there can be no sound and complete proof system for second-order logic. A logic with a sound and complete proof system – such as classical first-order predicate logic – is in that respect simpler, and simplicity is a genuine theoretical advantage. Sometimes, however, the truth is complex.

At a practical level, the situation is similar. A sound and complete proof system allows one to find out that arguments are valid. But even if no such system is possible, one may still be able to construct a sound system that is complete for most practical purposes: all provable arguments are valid and the only valid arguments that are not provable are very *recherché* ones. Moreover, the vocabulary involved in a given application may be so limited that a sound and complete proof system can be provided for the relevant fragment of continuum-valued logic. For example, the J_α operators specify exact degrees of truth; since ordinary speakers rarely do that, formalization of their reasoning can dispense with the J_α operators in the object-language.[54]

It has been suggested that, since human powers of discrimination are finite, only finitely many degrees of truth need be distinguished.[55] One might therefore consider Lukasiewicz's $(n + 1)$-valued predicate logic, in which degrees of truth are restricted to the rational numbers $0, 1/n, 2/n, \ldots, (n - 1)/n, 1$. These logics do have sound and complete proof systems.[56]However, the proposal does not seem theoretically justified. No specific finite limit to human powers of discrimination is a matter of logic, a limit on what degrees of truth are logically possible; none should therefore make any difference to validity. Moreover, even actual changes in a sentence's degree of truth need not be discriminable by speakers of the language. If 'The branch is long' is true to the degree that the branch is long, that degree of truth is subject to changes indiscriminable by us. At best, the finitely many-valued Lukasiewicz logics may be regarded as approximations to the continuum-valued logic, convenient for some practical purposes.[57]

Chapter 5

Supervaluations

5.1 INCOMPLETE MEANINGS

The problem of vagueness is often conceived as the problem of generalizing a formal theory of meaning applicable only to precise languages to a formal theory of meaning applicable to vague languages too. A Procrustean method of generalization would be to make the vague language precise, then apply the original theory. It is open to several obvious objections. First, a vague language can be made precise in more than one way. Second, the 'can' is only in principle; in practice we cannot make our vague language fully precise in even one way. Third, if a vague language is made precise, its expressions change in meaning, so an accurate semantic description of the precise language is inaccurate as a description of the vague one. These objections seem at first sight to have little in common with each other. However, one line of thought suggests a revision of the Procrustean method that promises to answer them all.

Perhaps the vagueness of a language consists in its capacity in principle to be made precise in more than one way. Not every substitution of precise meanings for vague ones counts as making the language precise, of course. Rather, vague meanings are conceived as incomplete specifications of reference. To make the language precise is to complete these specifications without contradicting anything in their original content.[1] For example, the meaning of 'heap' and the non-linguistic facts are supposed to determine of some things that they are heaps, of others that they are not heaps, and of still others to leave the matter open. The clear cases are those of the first kind, the clear non-cases those of the second, and the borderline cases those of the third. To make ‘heap’ precise is to assign it a meaning that makes it true of the clear cases, false of the clear non-cases, and either true or false of the borderline cases. Such a sharpening must also respect any systematic constraints built into the meaning of the term. For instance, if x and y are borderline cases of ‘heap’ whose only relevant difference is that x contains one grain more than y, then the vague meaning of ‘heap’ already ensures that if y is a heap then so too is x. Thus some sharpenings of ‘heap’ will be true of both x and y, some will be true of x and not of y, and some will be true of neither, but none will be true of y yet not of x. The original vague meaning of ‘heap’ is reflected not in any one sharpening but in the class of all its sharpenings.

What is needed is a method that collects many semantic descriptions of the language as made precise in different ways into one semantic description of the original vague language. The first objection above would be answered by considering all ways of making the vague language precise. The second objection would be answered by considering them collectively, without a futile attempt to specify them individually. The third objection would be answered because the incompleteness of a vague meaning is mirrored in the variety of its completions. A method of this kind is the method of *supervaluations*.

5.2 ORIGINS

The method of supervaluations and its application to vagueness seem to have originated in the philosophy of science. It was commonly supposed in the 1950s that scientific vocabulary could be divided into the observational and the theoretical. Observational terms drew their meaning from connections to experience defined by ostension; theoretical terms drew their meaning from connections to observational terms and with each other, as defined by a scientific theory. Vagueness seeped into the observational terms, for it was a commonplace that an ostensive definition using a finite number of particular experiences could not make clear the application of a term for all possible future experience. More important, even if the observational terms were treated as perfectly precise, and a scientific theory were specified to define the connections of the theoretical terms to them and with each other, the meanings of the theoretical terms would not thereby be fixed. Theoretical vocabulary was at best partially defined in terms of observational vocabulary: if it were totally defined in observational terms, it would not after all be theoretical. More than one interpretation of the theoretical terms would respect all the theoretical and ostensive connections.[2] Call any such interpretation *admissible.* On the view just sketched, all admissible interpretations are equally good. Although the plurality of admissible interpretations of the theoretical terms has a different source from the vagueness of the observational ones, both could be taken as forms of indeterminacy in meaning. A natural treatment of theoretical indeterminacy could then be extended to observational vagueness, by a generalization of admissibility to any interpretation meeting all the constraints on meaning.[3]

An admissible interpretation is precise.[4] Each statement in the language is either true on the interpretation or false on it. Call the corresponding assignment of truth-values to statements an admissible *valuation.* Now consider a scientific statement made using theoretical terms. Is it true or false? If each admissible valuation makes it true, then we can say without qualification that it is true. Similarly, if each admissible valuation makes it false, then we can say without qualification that it is false. But if some admissible valuations make it true while others make it false, then neither answer will do, for the interpretations backing one are no better than those backing the other. In this case we seem driven to say that the statement is neither true nor false, even though each admissible valuation makes it either true or false. The *supervaluation* is the assignment of truth to the statements true on all admissible valuations, of falsity to the statements false on all admissible valuations, and of neither to the rest. Few would now endorse the conception of science in which the method of supervaluations originated, for it seems to neglect both the dependence of observation on theory and the dependence of meaning on the actual nature of what is in the environment. Nevertheless, the treatment of indeterminacy it suggested might be the right one for vagueness.

Call a statement *supertrue* if it is true on all admissible interpretations and *superfalse* if it is false on all admissible interpretations. The supervaluationist claims that truth is supertruth and falsity is superfalsity.

The idea of supervaluations, although not the word, is used in Henryk Mehlberg's *The Reach of Science* (1958), and applied to vagueness. He defines a term as vague 'if it can be understood in several ways without being misunderstood'.[5] A statement including vague terms 'is true (or false, as the case may be) if it remains true (or false) under every admissible interpretation of the vague terms it contains'.[6] A non-scientific example of a vague term is 'Toronto', for the spatio-temporal boundaries of its denotation can be admissibly drawn in more than one way. Since 'Toronto is in Canada' is true on each admissible interpretation, it is true. Since 'Toronto is in Europe' is false on each admissible interpretation, it is false. Since 'The number of trees in Toronto is even' is true on some admissible interpretations and not on others, it is neither true nor false.[7]

Much of the interest of the method of supervaluations lies in its treatment of compound statements. Consider, for example, the statement 'The number of trees in Toronto is either odd or even'. It is true on each admissible interpretation, and therefore supertrue. Yet neither 'The number of trees in Toronto is odd' nor 'The number of trees in Toronto is even' is supertrue, for each is false on some admissible interpretations. Since truth is supertruth, according to Mehlberg, a true disjunction may have no true disjunct.[8] He distinguishes between what he calls the logical and meta-logical laws of excluded middle. The logical law of excluded middle is the schema '*A* or not *A*' in the object-language under study; it is now known just as the law of excluded middle. The meta-logical law of excluded middle is the meta-linguistic principle that any statement '*A*' in the object-language is either true or false; it is now known as the principle of bivalence. The supervaluationist denies the meta-logical law but accepts the logical law of excluded middle.[9] The statement '*A* or not *A*' is true on each admissible interpretation, and therefore true, even if '*A*' is true on some admissible valuations and false on others, and therefore neither true nor false, in which case 'Not *A*' is also neither true nor false. In such a case, '*A* or not *A*' is a true disjunction without a true disjunct.

Mehlberg points out that the admissibility of an interpretation cannot be assessed term by term.[10] Consider, for example, the statement 'If Eve is a woman then Eve is an adult', when Eve is a female human on the borderline of adulthood. The statement is clearly true; a woman is an adult female human. Now 'Eve is a woman' is true on some admissible interpretations of ‘woman’ and ‘Eve is an adult’ false on some admissible interpretations of ‘adult’. If the result of combining these interpretations were admissible, then the conditional would have a true antecedent and false consequent on at least one admissible interpretation, and would therefore not be supertrue. The result of combining the interpretations must therefore be inadmissible. One may interpret ‘Eve is a woman’ and ‘Eve is an adult’ as both true, and one may interpret them as both false, but one must not interpret the former as true and the latter as false. The semantic rules for the logical constants achieve a similar effect. On some admissible interpretations ‘Eve is a woman’ is true and ‘Eve is not a woman’ false; on some it is the other way round; but on none do the two statements have the same truth-value.

Mehlberg’s discussion was informal. A more systematic and rigorous treatment of supervaluations (under that name) was given by Bas van Fraassen in the 1960s. However, van Fraassen did not apply the method to the problem of vagueness; he was seeking a semantic treatment of names which lacked a reference and of self-referential sentences such as those involved in the Liar paradox.[11] The application to vagueness was worked out in detail by a number of writers in the 1970s: Michael Dummett, Kit Fine, Hans Kamp, David Lewis, Marian Przelecki.[12]

5.3 LOGIC AND SEMANTICS

A striking feature of supervaluations is the failure of truth-functionality for compound statements. In particular, the standing of a conjunction or disjunction as true, false or neither is not determined by the standings of its conjuncts or disjuncts. Suppose, as before, that ‘Eve is an adult’ and ‘Eve is not an adult’ are neither true nor false. As already noted, ‘Eve is an adult or Eve is not an adult’ is true; but the pleonastic ‘Eve is an adult or Eve is an adult’ is neither true nor false, for it is equivalent simply to ‘Eve is an adult’. Yet the two disjunctions are indistinguishable in terms of the semantic standings of their components; each has two disjuncts that are neither true nor false. Similarly, the two conjunctions ‘Eve is an adult and Eve is not an adult’ and ‘Eve is an adult and Eve is an adult’ are indistinguishable in terms of the semantic standings of their components, each having two conjuncts that are neither true nor false; yet the former conjunction is false and the latter neither true nor false. Contrast a many-valued approach, on which the degree of truth of a conjunction is a function of the degrees of truth of its conjuncts (see also Section 4.14). When ‘Eve is an adult’ has the same degree of truth as its negation, such degree-functionality forces the merely repetitious ‘Eve is an adult and Eve is an adult’ to have as low a degree of truth as the self-contradictory ‘Eve is an adult and Eve is not an adult’. In those circumstances, it also forces ‘Eve is an adult or Eve is an adult’ to have as high a degree of truth as ‘Eve is an adult or Eve is not an adult’. Again, the supervaluation makes ‘Eve is not an adult or Eve is a woman’ true and ‘Eve is an adult and Eve is not a woman’ false; the degree-functional approach assigns them both an intermediate degree of truth.[13]

There is a corresponding difference for conditionals. In the envisaged circumstances, the degree-functional approach cannot distinguish between the obvious ‘Eve is a woman if and only if Eve is an adult’ and the silly ‘Eve is a woman if and only if Eve is not an adult’, since there is no difference in degree of truth on either side of the ‘if and only if’. But each admissible valuation either makes ‘Eve is a woman’ and ‘Eve is an adult’ true and ‘Eve is not an adult’ false or makes ‘Eve is a woman’ and ‘Eve is an adult’ false and ‘Eve is not an adult’ true, so the supervaluation makes ‘Eve is a woman if and only if Eve is an adult’ true and ‘Eve is a woman if and only if Eve is not an adult’ false.

The differences above between the two approaches tell heavily in favour of supervaluations.[14] They are sensitive to intuitively significant distinctions obliterated by degree-functionality. Parallel arguments show the superiority of supervaluations to other forms of many-valued logic. A three-valued logic, for example, based on a classification of sentences as true, false or neither, fares just as badly as one based on a classification according to degree of truth.

What effect have supervaluations on logic? To answer the question, we must first settle on an account of validity appropriate to supervaluations. It might be suggested that if the condition for a sentence to be true is that every admissible valuation makes it true, then the analogous condition for an argument to preserve truth is that every admissible valuation that makes its premises true also makes its conclusion true. Since validity is necessary preservation of truth, an argument is valid just in case necessarily every admissible valuation that makes its premises true also makes its conclusion true. This property may be called *local validity.* The problem for supervaluationists is that supertruth plays no role in the definition of local validity. Yet they identify truth with supertruth; since validity is necessary preservation of truth, they should identify it with necessary preservation of supertruth. That amounts to an alternative definition, on which an argument is valid just in case necessarily if every admissible valuation makes its premises true then every admissible valuation makes its conclusion true, in other words, necessarily if its premises are supertrue then its conclusion is also supertrue. The latter property may be called *global validity.* An admissible valuation on which '*A*' is true and '*B*' false automatically makes the argument from '*A*' to '*B*' not locally valid, but the argument might still be globally valid, for if '*A*' is false on another admissible valuation, this is not a case in which '*A*' but not '*B*' is supertrue. It is obvious that a locally valid argument is also globally valid. In many languages, the converse also holds: any globally valid argument is also locally valid. Indeed, in any language the converse holds for an argument without premises. But examples will be given later of other arguments that are globally but not locally valid. For the reason given above, we may work with the assumption that supervaluationists identify validity with global validity. From time to time the consequences of the alternative identification with local validity will also be noted.[15]

In simple cases, global validity coincides with classical validity, as it does with local validity. Consider, for example, any argument containing no constants other than negation, conjunction, disjunction, material implication, identity and the universal and existential quantifiers. The premises and conclusion may contain propositional or predicate variables, which can have any interpretation appropriate to their syntactic category. If the argument is classically valid, then since admissible valuations are classical, any admissible valuation that makes the premises true also makes the conclusion true; thus the argument is locally valid, and therefore globally valid. Conversely, it can be shown that if it is not classically valid, then there is an assignment of precise values to its variables on which its premises are true and its conclusion false, so it is neither locally nor globally valid. Within this language, local, global and classical validity are equivalent. In particular, a single formula as the conclusion of an argument without premises is globally valid if and only if it is classically valid, whence the global validity of the law of excluded middle. Thus supervaluationism seems to inherit the power of classical logic.

Supervaluations validate classical predicate logic, but they also enable it to be extended. They make room for a new operator 'definitely' to express supertruth in the object-language: 'Definitely *A*' is true if and only if '*A*' is supertrue. For example, to say that something is neither definitely a heap nor definitely not a heap is to say that it is a borderline case. 'Definitely' can be given a formal semantics very like the possible worlds semantics for the modal operator 'necessarily'. For simplicity, the analogy will first be developed by reference to the simple modal system S5. Account is taken of further complications in Section 5.6. The aim of the formal semantics is to define in mathematical terms a set of models such that an argument is valid if and only if it preserves truth in every model in the set, for that will provide us with a precise standard of validity.

A model for S5 is a *structure* containing a number of objects, which may be thought of as *possible worlds*. Each world in such a structure is associated with a valuation: sentences are true or false at worlds. 'Not *A*' is true at a world if and only if '*A*' is not true at that world, '*A* and *B*' is true at a world if and only if '*A*' is true at that world and '*B*' is true at that world, and so on. In contrast, the truth-value of 'Necessarily *A*' at a world in a structure depends on its truth-values at all the worlds in the structure, not just that one; 'Necessarily *A*' is true at a world in a structure if and only if *A* is true at every world in the structure. By analogy, a model for the language with 'definitely' is a *space* containing a number of *points*, which may be thought of as admissible interpretations. Each point in such a space is associated with a valuation: sentences are true or false at points. 'Not *A*' is true at a point if and only if '*A*' is not true at that point, '*A* and *B*' is true at a point if and only if '*A*' is true at that point and '*B*' is true at that point, and so on. The truth-value of 'Definitely *A*' at a point in a space depends on its truth-values at all the points in the space, not just that one; 'Definitely *A*' is true at a point in a space if and only if '*A*' is true at every point in the space.

A formula in the language of 'necessarily' is valid in the S5 semantics if and only if it is true at every world in every structure. Thus every instance of the axiom schema 'If necessarily if *A* then *B*, and necessarily *A*, then necessarily *B*' is valid in S5, for when 'If *A* then *B*' and '*A*' are both true at every world in a structure, so too is '*B*'. Moreover, if '*A*' is valid in S5, then it is true at every world in any structure, so 'Necessarily *A*' is true at every world in any structure, so 'Necessarily *A*' is valid in S5 too (the rule of necessitation). Quite generally, the logical consequences of necessary truths are themselves necessary truths. 'If necessarily *A* then *A*' (known as the T schema) is also valid in S5, for if '*A*' is true at every world in a structure, then it is true at any world in that structure. What necessarily holds, holds. The distinctive S5 axiom schema is 'If not necessarily *A* then necessarily not necessarily *A*'; it is valid in S5 because the semantics makes the addition of 'necessarily' to 'Not necessarily A' vacuous. For similar reasons, 'If necessarily *A* then necessarily necessarily *A*' (known as the S4 schema) is also valid in S5. According to S5, it cannot be contingent whether something is necessary.

By analogy, a formula in the language of 'definitely' is valid if and only if it is true at every point in every space. Thus 'If definitely if *A* then *B*, and definitely *A*, then definitely *B*' is valid. If '*A*' is valid, then 'Definitely *A*' is valid too. Quite generally, the logical consequences of definite truths are themselves definite truths. The T schema 'If definitely *A* then *A*' is valid. What definitely holds, holds. The S5 schema, 'If not definitely *A* then definitely not definitely *A*' is valid, as is the S4 schema 'If definitely *A* then definitely definitely *A*'. On this semantics, it cannot be indefinite whether something is definite.

Validity has so far been considered only for single formulas. However, the more important notion is of validity for arguments. Here the analogy between 'definitely' and 'necessarily' begins to break down. An argument is valid in S5 if and only if in any structure, if its premises are true at a given world, then so is its conclusion. For 'definitely', one might analogously define an argument as valid if and only if in any space, if its premises are true at a given point, then so is its conclusion. However, this is the formal analogue of *local* validity. It was argued above that supervaluationists should identify validity with *global* validity. They should therefore use its formal analogue: an argument is (globally) valid just in case in any space, if the premises are true at *every* point, then so is the conclusion.

For single formulas, local and global validity coincide, so the analogy between 'definitely' and 'necessarily' remains. The disanalogies emerge for arguments with at least one premise. If '*A*' is an atomic formula, the inference from '*A*' to 'Necessarily *A*' is not valid in S5, for '*A*' can be true at one world in a structure and false at another world in the same structure, so that 'Necessarily *A*' is false at the former world, and validity in S5 requires each world to preserve truth. The converse inference from 'Necessarily *A*' to '*A*' is of course valid in S5. For 'definitely', in contrast, global validity merely requires each space to preserve supertruth (truth at every point), so the inference from '*A*' to 'Definitely *A*' is globally (but not locally) valid, for if '*A*' is true at every point in a space, then so is 'Definitely *A*'. The converse inference from 'Definitely *A*' to '*A*' is of course globally valid.

Since '*A*' and 'Definitely *A*' are interderivable, one might expect 'definitely' to be a redundant operator. It is not, however, for 'If *p* then definitely *p*' is not globally valid, where '*p*' is an atomic formula. If '*p*' is true at some but not all points in a space, then the conditional has a true antecedent and false consequent.[16] For similar reasons, the inference from 'Not definitely *p*' to 'Not *p*' is not globally valid. Such examples can be shown to involve breakdowns of the classical rules of contraposition, conditional proof, argument by cases and *reductio ad absurdum* in the supervaluationist logic of 'definitely'. This is in a sense a violation of classical propositional logic, but at the level of inference rules permitting transitions from arguments to arguments rather than from formulas to formulas. The cases are as follows.

(a) *Contraposition* In classical logic, if one can validly argue from '*A*' and auxiliary premises to '*B*', then one can validly argue from 'Not *B*' and the auxiliary premises to 'Not *A*'. Contraposition does not hold in the supervaluationist logic, for although the inference from '*p*' to 'Definitely *p*' is globally valid, the inference from 'Not definitely *p*' to 'Not *p*' is not globally valid.[17]

(b) *Conditional proof* (*the deduction theorem*) This is the standard way of reaching conditional conclusions. In classical logic, if one can validly argue from '*A*' and auxiliary premises to '*B*', then one can validly argue from the auxiliary premises alone to 'If *A* then *B*'. Conditional proof does not hold in the supervaluationist logic, for although the inference from 'p' to 'Definitely p' is globally valid, the unpremised conclusion 'If p then definitely p' is not globally valid.[18]

(c) *Argument by cases* (*or-elimination*) This is the standard way of using disjunctive premises. In classical logic, if one can validly argue from 'A' and auxiliary premises to 'C', and from 'B' and auxiliary premises to 'C', then one can validly argue from 'A or B' and the combined auxiliary premises to 'C'. Argument by cases does not hold in the supervaluationist logic, for although the inference from 'p' to 'Definitely p or definitely not p' is globally valid, as is that from 'Not p' to the same conclusion, the inference from 'p or not p' to 'Definitely p or definitely not p' is not globally valid.

(d) *Reductio ad absurdum* This is the standard way of reaching negative conclusions. In classical logic, if one can validly argue from 'A' and auxiliary premises to 'B', and from 'A' and auxiliary premises to 'Not B', then one can validly argue from the combined auxiliary premises to 'Not A'. *Reductio ad absurdum* does not hold in the supervaluationist logic, for the inference from 'p and not definitely p' to 'Definitely p' is globally valid, as is that from the same premise to 'Not definitely p', but the unpremised conclusion 'Not: p and not definitely p' is not globally valid (it is equivalent to 'If p then definitely p').[19]

Conditional proof, argument by cases and *reductio ad absurdum* play a vital role in systems of natural deduction, the formal systems closest to our informal deductions. They are the rules by which premises are discharged, i.e. by which categorical conclusions can be drawn on the basis of hypothetical reasoning. Contraposition is another very natural deductive move. Thus supervaluations invalidate our natural mode of deductive thinking. The examples so far have all involved the 'definitely' operator. If we had to exercise caution only when using this special operator, then our deductive style might not be very much cramped. However, supervaluationists have naturally tried to use their semantic apparatus to explain other locutions. If their attempts succeed, our language will be riddled with counterexamples to the four rules (see Section 5.5). In order to restore classical logic, supervaluationists might switch to the alternative definition of validity as local validity. That would restore classical logic at the expense of endangering the identification of truth with supertruth, for validity would no longer be identified with the preservation of supertruth. The gravity of the danger will emerge in Section 5.7.

5.4 THE ELUSIVENESS OF SUPERTRUTH

According to supervaluationism, '*p* or *q*' is sometimes true when no answer to the question 'Which?' is true. For similar reasons, 'Something is *F*' is sometimes true when no answer to the question 'Which thing is *F*?' is true. In this sense supertruth is elusive.

The most dramatic examples occur in sorites paradoxes. Consider the Heap. Any admissible valuation has a cut-off number k for 'heap', more than one and less than ten thousand. On the valuation, 'n grains make a heap' is true if n is more than k and false otherwise; in particular, '$k + 1$ grains make a heap' is true and 'k grains make a heap' false, so '$k + 1$ grains make a heap and k grains do not make a heap' is true, so 'For some n, $n + 1$ grains make a heap and n grains do not make a heap' is true. Since the existential generalization is true on each admissible valuation, it is supertrue. Yet no answer to the question 'For which n do $n + 1$ grains make a heap and n grains not make a heap?' is supertrue, for not all admissible valuations have the same cut-off number. The supervaluational response to the sorites argument from 'For all n, if $n + 1$ grains make a heap then n grains make a heap' and 'Ten thousand grains make a heap' to 'One grain makes a heap' is now clear. The argument is classically valid, and therefore globally (and locally) valid. The minor premise 'Ten thousand grains make a heap' is supertrue, and the conclusion 'One grain makes a heap' superfalse. However, each admissible valuation has a counterexample to the major premise, although it is not the same in each case; thus the major premise is superfalse. In short, the argument is valid but unsound.[20]

The supervaluational treatment of the sorites argument is formally elegant. The question is whether it defuses the intuitive backing for the major premise. Many people have found the major premise plausible just because it seemed to them that there could not be a number n such that $n + 1$ grains make a heap and n do not. Supervaluationism makes the very claim that they find incredible. Nor should the supervaluationist say that the claim does not mean what they think it means. The point of the enterprise is to give semantic descriptions of vague sentences as we actually use them. If supervaluationism delivers a meaning for the existential claim other than its ordinary one, the enterprise fails.

The supervaluationist must insist that the sentence ‘For some *n*, *n* + 1 grains make a heap and n grains do not make a heap’ is true in its ordinary sense, and use the apparatus of supervaluations to explain away appearances to the contrary. Usually, a true existential generalization has a true instance; we note that no sentence of the form ‘*n* + 1 grains make a heap and n grains do not make a heap’ is true, and naturally tend to assume that the corresponding existential generalization is not true either. In effect, the explanation is that we ignore vagueness, making semantic assumptions appropriate only if ‘heap’ were not vague. The trouble with the explanation is that it assumes that we do not ignore vagueness at a different point. It is precisely because we have noticed the vagueness of ‘heap’ that we doubt that anything of the form ‘*n* + 1 grains make a heap and *n* grains do not make a heap’ can be true. Perhaps we are so confused that we notice the vagueness and ignore its consequences. However, there is at least a suspicion of a mismatch between supertruth and truth in the ordinary sense. This suspicion will be confirmed in Section 5.7.[21]

5.5 SUPERVALUATIONAL DEGREES OF TRUTH

The idea of degrees of truth tends to be associated with the assumption that the degree of truth of a complex sentence is a function of the degrees of truth of its components, and in particular with many-valued logic. However, Lewis and Kamp showed that it can be understood in terms of supervaluations too.[22] As a first try, suppose that ‘*B*’ is true on every admissible interpretation on which ‘*A*’ is true; in other words, the material conditional ‘If *A* then *B*’ is supertrue. One might then say that ‘*B*’ is *at least as true as* ‘*A*’. If ‘*B*’ is at least as true as ‘*A*’ but ‘*A*’ is not at least as true as ‘*B*’, then one might say that ‘*B*’ is *truer than* ‘*A*’. For example, different admissible interpretations will set different precise standards for counting as ‘young’. However, if Adam was born before Eve, then she will count as ‘young’ by any standard by which he does.

On any admissible interpretation on which 'Adam is young' is true, 'Eve is young' is true too; thus 'Eve is young' is at least as true as 'Adam is young'. Correspondingly, the conditional 'If Adam is young then Eve is young' is supertrue. Assume that there is a reasonable standard by which Eve counts as 'young' and Adam does not. Then 'Adam is young' is not at least as true as 'Eve is young', so the latter is truer than the former.

That first attempt is rather crude. For example, one might sometimes wish to say that 'Eve is young' and 'Eve is not young' are equally true, in the sense that each is at least as true as the other. According to the definitions above, that is impossible, for '*A*' and '*B*' are as true as each other if and only if they are true on exactly the same admissible interpretations, which contradictories never are (given that there is at least one admissible interpretation). Nevertheless, 'Eve is young' and 'Eve is not young' might be thought to be equally true in the sense that the set of admissible interpretations on which the former is true and the set on which the latter is are equally 'large'. This idea presupposes a measure of the size of sets of admissible interpretations. Such a measure might boldly, or rashly, be postulated.

Once a notion of 'truer than' is in place, it can be used to formulate a semantic treatment of comparative adjectives. The guiding principle is that '*a* is *F*er than *b*' is true if and only if '*a* is *F*' is truer than '*b* is *F*'. For example, 'Eve is younger than Adam' is true if and only if 'Eve is young' is truer than 'Adam is young'. More precisely, '*a* is *F*er [more *F*] than *b*' is true on an admissible valuation if and only if '*a* is *F*' is truer than '*b* is *F*'. A corresponding account can be given of 'at least as'. '*a* is at least as *F* as *b*' is true on an admissible valuation if and only if '*a* is *F*' is at least as true as '*b* is *F*'.[23] Lewis and Kamp extended the treatment of comparatives to modifiers such as 'rather' and 'in a sense', as in 'rather clever' and 'clever in a sense'.[24]

Semantic treatments of the kind above face a number of related problems that seem to be caused by the use of a fixed class of admissible interpretations. They may be compared with the similar problems faced by degree-theoretic treatments of such constructions within a framework of many-valued logic (Section 4.11).

(a) Since 'truer than' is defined in terms of *all* admissible valuations, the truth value of '*a* is *F*er than *b*' should be stable from one admissible valuation to another. Thus the semantics predicts that comparatives and related terms should be absolutely precise. Now 'taller' does indeed seem more precise than 'tall'. But it does not seem perfectly precise; stoops and curly scalps may produce borderline cases even for it. A comparative such as 'more intelligent' is notably vague.

(b) If 'truer than' is defined in terms of admissible valuations, '*A*' is not truer than '*B*' when '*B*' is true on every admissible valuation; so if '*A*' is truer than '*B*', '*B*' is not (super)true. Since the truth of 'David is braver than Saul' requires 'David is brave' to be truer than 'Saul is brave', it is incompatible with the truth of 'Saul is brave'. Thus 'Saul is brave, but David is braver than Saul' cannot be true. That is absurd. The brave are not all equally brave.[25]

(c) Consider 'acute' as an adjective of angles. It is precise, for '*a* is acute' is true if *a* is less than a right angle, and false otherwise. 'An angle of 60° is acute' is true, and therefore true on every admissible valuation. Nevertheless, an angle of 30° is more acute than an angle of 60°.

(d) One would expect that a good semantics of comparatives would extend to a modifier such as 'very'. Now if 'very' is treated by means of admissible valuations, it should be sufficient for the truth of '*a* is very *F*' that '*a* is *F*' is true on every admissible valuation. But then 'That is definitely dark blue' should entail 'That is very dark blue', which it does not.

The moral of (a) is presumably that 'admissible' is itself a vague notion. The resolution of the problem therefore depends on an adequate supervaluationist treatment of higher-order vagueness. Such a treatment may also provide the means to resolve (b)–(d).[26] More generally, one of the chief challenges to supervaluationism is to make proper room for higher-order vagueness. We must therefore give a supervaluationist account of the phenomenon, before briefly returning to (a)–(d).

5.6 SUPERVALUATIONS AND HIGHER-ORDER VAGUENESS

A supervaluation divides the sentences of a language into three classes: the true, the false, and the neither true nor false. The comprehensiveness of the classification is plausible enough in some applications of the method. Consider, for example, the view that the future is open except in so far as it is determined by laws of nature and the present state of the universe. On this view, a future tense statement is now true if it is true of each possible future, false if it is false of each possible future, and neither true nor false otherwise. Since each possible future corresponds to a bivalent valuation, the method of supervaluations is appropriate. Whatever its merits, this view makes the threefold division quite natural. Vagueness is a different matter. If it is hopeless to look for the first red shade in a sorites series from orange to red, it is equally hopeless to look for the first shade which can truly be called 'red' (try). The idea that our rough-and-ready use of vague terms does not determine hidden boundaries tells just as much against a pair of hidden second-order boundaries between the true and the neither true nor false and between the latter and the false as it does against one hidden first-order boundary. In supervaluationist terms, the admissibility of a valuation is itself a vague notion.[27]

Second-order vagueness shows itself in an object-language with a 'definitely' operator. If vagueness were only first order, 'Definitely *A*' would be precise, so 'Either definitely definitely *A* or definitely not definitely *A*' would hold. But 'Definitely *A*' has borderline cases, for otherwise there would be sharp second-order boundaries. Since borderline cases for '*A*' are counterexamples to the schema 'Either definitely *A* or definitely not *A*', borderline cases for 'Definitely *A*' are counterexamples to the schema 'Either definitely definitely *A* or definitely not definitely *A*'. That schema entails the S5 schema 'If not definitely *A* then definitely not definitely *A*', for 'If not definitely *A* then not definitely definitely *A*' is uncontroversial.[28] It follows that the simple form of supervaluationist semantics described in Section 5.3 is inadequate, for it validates the S5 schema.

Third-order vagueness is equally possible. If it were not, everything would be either definitely definitely definitely red or definitely not definitely definitely red: but where would the break come? The point extends to any order. If $(n + 1)$th-order vagueness were not possible, everything would be either definitely definitely^n red or definitely not definitely^n red. The idea that our rough-and-ready use of a vague term does not determine hidden boundaries tells against a hidden boundary between the extension of the first disjunct and the extension of the second. Even if high orders of vagueness are somehow ruled out by empirical factors, that would not entitle the logician to treat them as impossible. Thus supervaluationist logic should admit at least all finite orders of vagueness.[29]

Supervaluationists often regard admissibility as consistency with the semantic rules of the language. If the rules decide a case, then an admissible interpretation decides it in the same way; it may decide a case when they do not. Since consistency is a matter of logic, admissibility looks as though it should be a precise concept.[30] Higher-order vagueness shows this picture to be misleading. A vague meaning is not like a partial definition in mathematics, formulated in precise terms but not covering all cases. If a vague term is governed by semantic rules, then they are formulated in equally vague terms. Moreover, it is not plausible that the limits of higher-order vagueness are laid down by a hierarchy of higher-order semantic rules. When admissibility is not pictured as consistency with semantic rules, supervaluationism becomes a less inviting approach. Nevertheless, it can adapt to higher-order vagueness. Admissibility might be conceived as a matter of *reasonableness.* An interpretation is reasonable if it does not license misuses of the language (from the standpoint of an ordinary understanding of it).[31] The concept of a misuse is obviously and essentially vague.

The formal semantics for 'definitely' in Section 5.3 validated the S5 axiom. A new formal semantics is therefore required to make room for higher-order vagueness. A similar problem arises in modal logic with the semantics for systems weaker than S5. There, the informal idea is that the possibility of a world is itself a contingent matter. The formal trick is to introduce a relation of accessibility between worlds in a structure; 'Necessarily A' is true at a world w if and only if 'A' is true at every world accessible from w. Analogously, one can introduce a relation of admitting between points in a space; 'Definitely A' is true at a point s if and only if 'A' is true at every point admitted by s. The informal idea is that the admissibility of an interpretation is itself a matter for interpretation. Each interpretation makes its own ruling as to which interpretations are admissible. Formally, a point determines both a bivalent valuation and a set of admitted points. Every point should admit itself; were admitting not a reflexive relation, 'A' might be true at every point admitted by a point s yet not at s itself, in which case 'If definitely A then A' would be false at the point. An interpretation should regard at least itself as reasonable.

In order to accommodate higher-order vagueness, admitting is allowed to be non-transitive. Informally, an interpretation might admit just those interpretations that are reasonable by its lights, because they do not differ from it by too much. If you regard me as reasonable and I regard him as reasonable, you may not regard him as reasonable, for the difference between you and him may be too much even if neither the difference between you and me nor that between me and him is too much. Suppose that a point s admits a point t, which admits a point u, but s does not admit u, and 'A' is true at every point admitted by s but not at u. Then 'Definitely A' is true at s but not at t, so 'Definitely definitely A' is false at s, and so the S4 principle 'If definitely A then definitely definitely A' is false at s. By similar reasoning, the S5 principle 'If not definitely A then definitely not definitely A' also fails. Indeed, in the absence of special restrictions on admitting, the valid formulas are just those corresponding to the theorems of the weak modal logic T, which are just those derivable from the principles other than the S4 and S5 schemata listed in Section 5.3 as valid for 'definitely'.[32]

The interpretation dependence of admissibility is exactly analogous to the contingency of possibility in modal logics weaker than S5. What supervaluationism adds is a conception of admissibility. The conception needs to be worked out with some care. Interpretations specify lists of admissible interpretations, including themselves; such interpretations might be suspected of vicious circularity. Fortunately, the circularity can be avoided. A 0-level interpretation makes precise those expressions of the language that, unlike 'definitely', do not have to do with admissibility. A 1-level interpretation specifies which 0-level interpretations are admissible. More generally, an $(i + 1)$-level interpretation specifies which i-level interpretations are admissible. An ω-level interpretation is an infinite sequence $s_0s_1s_2\ldots$, where each s_i is an i-level interpretation and each s_{i+1} specifies that s_i is admissible. An ω-level interpretation $s_0s_1s_2\ldots$ admits an ω-level interpretation $t_0t_1t_2\ldots$ if and only if each s_{i+1} admits t_i; in non-relational terms, $t_0t_1t_2\ldots$ is admissible if and only if each t_i is admissible. A point in a space is the formal analogue of an ω-level interpretation. Thus every point admits itself, but the definitions can be shown to impose no further constraint on the structure of admitting. Further constraints could be added, but they will not be considered here.

An objection might be raised to the foregoing account of higher-order vagueness. Define 'Definitely* *A*' to mean the infinite conjunction: *A* and definitely *A* and definitely definitely *A* and . . . . The definition guarantees that if definitely* *A* then definitely definitely* *A* and indeed definitely* definitely* *A*. In terms of the formal semantics, let a point *s* admit* a point *t* if and only if either *s* admits *t*, or *s* admits a point that admits *t*, or *s* admits a point that admits a point that admits *t*, or . . . . 'Definitely* *A*' is true at a point *s* if and only if '*A*' is true at every point that *s* admits*. In technical terms, admitting* is the ancestral of admitting; it is automatically transitive, even though admitting is not. The supervaluationist approach can now be applied in terms of admissibility* rather than admissibility. Since the strict notion 'definitely*' obeys an S4 axiom, higher-order vagueness disappears. It turns out to be a surface phenomenon, reflecting our use of an unnecessarily lax standard of definiteness. According to the objection, the supervaluationist cannot recognize higher-order vagueness as a deep phenomenon. The point might be used against supervaluationism; it might be used against the claim that higher-order vagueness is a deep phenomenon.[33]

The supervaluationist may insist that even 'definitely*' is vague. Its vagueness is not manifest in its failure to obey some principle stated in its own terms, but that is just to say that it cannot be used to measure its own vagueness. It is like a cloud said to have an exact length because it is exactly as long as itself. A new operator 'definitely!' is needed to express the vagueness of 'definitely*' in the failure of the principle 'If definitely* *A* then definitely! definitely* *A*'. The vagueness of 'definitely*' corresponds to a vagueness in the meta-language for the original language of 'definitely' on which there had been no need to reflect before 'definitely*' was introduced. The meta-language for the new language of 'definitely!' may in turn harbour vagueness whose expression will require yet another operator 'definitely!!'. The process may have no natural end.

An alternative supervaluationist reply is that 'definitely*' is precise, but imposes a condition that hardly any sentences meet. In a sorites series of men from tall to short, there are more tall men than definitely tall men, more definitely tall men than definitely definitely tall men, and so on. Each iteration of ‘definitely’ reduces the number until none is left. Since the series is finite, such a point will be reached. Thus no man is definitely* tall. This conclusion does not conflict with common sense; it is not disputed that many men are definitely tall. In loose talk we may use repetition for mere emphasis, losing sight of the fact that ‘definitely definitely tall’ is stronger than ‘definitely tall’ in content as well as tone, but it remains a fact.

On the former reply, the definitely* tall men fade into the not definitely* tall men. On the latter, there are no definitely* tall men. In either case there is no sharp line between definitely* tall men and not definitely* tall men. More generally, the two replies agree that there is no sharp line between two phenomena, the perfectly straightforward application of a term and its less than perfectly straightforward application. This conclusion is anyway forced by the view that vagueness does not usually involve hidden boundaries. For a sharp boundary between the perfectly straightforward applications of a vague term and its less than perfectly straightforward applications would usually be hidden, as one can ascertain by trying to find it.

The difficulty comes out in Kit Fine’s suggestion ‘Anything that smacks of being a borderline case is treated as a clear borderline case’.[34] Suppose that the proposal succeeds in drawing a sharp line around the borderline cases. Then there are non-borderline cases very close to the line; but they will smack of being borderline cases, being reminiscent in appearance of cases just the other side of the line, and so count as clear borderline cases after all, which is a contradiction. Thus the proposal does not succeed. Fine’s ruling extends the area of borderline cases, which extends the area of cases that smack of being borderline cases, which extends the area of borderline cases, which extends . . . ; the process has no stable limit short of including all cases. There is also the simple point that a case to which a term perfectly straightforwardly applies might smack of being a borderline case; even the wholly innocent can incur suspicion.

The term ‘perfectly straightforward application of a term’ is itself vague. Not even iterating the supervaluationist construction into the transfinite will give it a precise sense. There is no good reason to treat its vagueness differently from that of other terms. If their vagueness involves indeterminacy, then so does its. Supervaluationism cannot eliminate higher-order vagueness. It must conduct its business in a vague meta-language.

5.7 TRUTH AND SUPERTRUTH

In acknowledging higher-order vagueness, the supervaluationist acknowledges the vagueness of the concept of supertruth. Supertruth is truth on all admissible valuations, and the concept of admissibility is vague. This point indirectly calls into question the supervaluationist equation of truth with supertruth.[35]

Truth is standardly assumed to have the disquotational property to which Tarski drew attention. 'Cascais is in Portugal' is true if and only if Cascais is in Portugal. More generally: '*A*' is true if and only if *A*. Here '*A*' may be replaced by a sentence of the object-language under study; a truth predicate for the object-language has been added to that language to extend it to a meta-language for it. The 'if and only if' is just the material biconditional. How much more there is to the concept of truth than the disquotational property is far from clear, but in most contexts truth is assumed to be at least disquotational, whatever else it is or is not.

Supertruth is not disquotational. If it were, then the supervaluationist would be forced to admit bivalence. Consider any sentence '*A*'. By supervaluationist logic, either *A* or not *A*. Suppose that supertruth is disquotational. Thus '*A*' is supertrue if and only if *A* and 'Not *A*' is supertrue if and only if not *A*. It would then follow, by more supervaluationist logic, that either '*A*' is supertrue or 'Not *A*' is supertrue; in the latter case, '*A*' is superfalse. In order to allow vague sentences in borderline cases to be neither supertrue nor superfalse, the supervaluationist must deny that supertruth is disquotational. Indeed, this is just to deny the meta-linguistic equivalent of the claim that 'definitely' is a redundant operator, which the supervaluationist has already denied.

The supervaluationist did allow the statement that definitely *A* to entail and be entailed by the statement that *A*.[36] In the same way, the supervaluationist may allow the statement that '*A*' is supertrue to entail and be entailed by the statement that *A*. Were 'if and only if' to be used for mutual entailment, the disquotational schema would have a reading acceptable to the supervaluationist. It is not Tarski's reading, on which 'if and only if' is the material biconditional. More important, the mutual entailment reading fails to capture the disquotational idea. If the truth predicate really does have the effect of stripping off quotation marks, then the material biconditional that '*A*' is true if and only if *A* strips down to the tautology that *A* if and only if *A*. The supervaluationist denies that supertruth behaves like that; the availability of the mutual entailment reading is an irrelevance.

A disquotational form of truth can be introduced within the supervaluationist framework. Add quotation marks and a predicate 'true$_T$' of object-language sentences to the object-language, and let '"*A*" is true$_T$' be true on an interpretation if and only if *A* is true on that interpretation. The supervaluationist allows that either '*A*' is true$_T$ or 'Not *A*' is true$_T$, for this is to allow no more than that either *A* or not *A*. In Fine's phrase, the vagueness of 'trues$_T$' waxes and wanes with the vagueness of the given sentence. He suggests that 'true$_T$' is conceptually prior to 'supertrue', for 'supertrue' is definable in terms of 'true$_T$' and 'definitely' – a sentence is supertrue just in case it is definitely true$_T$ – and no reverse definition is possible.[37]

Truth$_T$ is disquotational; supertruth is not. In order of definition, truth$_T$ is primary; supertruth is secondary. Why then does the supervaluationist identify ordinary truth with supertruth rather than with truth$_T$? The idea seems to be that truth$_T$ is not a determinate condition, and therefore has no place in an objective semantics. Truth$_T$ is disqualified because not every sentence is either definitely true$_T$ or definitely not true$_T$. But this disqualification rests on the hopeless demand for a precise meta-language. Once higher-order vagueness is recognized, it disqualifies supertruth just as first-order vagueness disqualifies truth$_T$; not every sentence is either definitely supertrue or definitely not supertrue.[38] There is no more reason to equate ordinary truth with supertruth, definite truth$_T$, than with definite definite truth$_T$. There is more reason to identify it with truth$_T$. Truth$_T$ is vague, but so is any notion of truth we can grasp. Perhaps the ordinary concept of truth *should* match the vagueness of the sentences to which it is applied.

Once the supposed advantages of supertruth are seen to be illusory, it becomes overwhelmingly plausible to equate ordinary truth with the property that meets Tarski's disquotational condition, truth$_T$.[39] Even in a borderline case it is allowed that a vague sentence or its negation is true$_T$; thus it is either true or false in the ordinary sense. Vague sentences are not counter-examples to bivalence. Moreover, if truth is truth$_T$ rather than supertruth, then validity is a matter of preserving truth$_T$. This immediately restores the classically valid patterns of reasoning that must be abandoned if validity is a matter of preserving supertruth: contraposition, conditional proof, argument by cases, *reductio ad absurdum.*[40]The logic of 'definitely' ceases to be distinctive, becoming isomorphic in both theorems and rules of inference to a weak modal logic. What then remains of supervaluationism?

There remains the 'definitely' operator, with its semantics of admissible interpretations. However, this apparatus has lost its privileged connection with the concept of truth. Of any admissible valuation, we can ask whether it assigns truth to all and only the true sentences of the language and falsity to all and only the false ones. At most one valuation has that property. But then any other valuation will assign truth-values incorrectly, so how can it be admissible? It might be replied that no interpretation is definitely the one with the desirable property. Once definiteness has been separated from truth, that reply is without force. If an interpretation does have the desirable property, why should it matter if it does not definitely have it? Indeed, the reply is in danger of losing its sense as well as its force. If we cannot grasp the concept of definiteness by means of the concept of truth, can we grasp it at all? No illuminating analysis of 'definitely' is in prospect. Even if we grasp the concept as primitive, why suppose it to be philosophically significant?

One can make sense of the supervaluationist apparatus if one assumes that an interpretation *s* admits an interpretation *t* just in case if *s* were correct then speakers of the language could not know *t* to be incorrect. On this view, 'definitely' means something like 'knowably'. Just one interpretation is correct, but speakers of the language cannot know all others to be incorrect. Vagueness is an epistemic phenomenon. But that is not the supervaluationist view. Of supervaluationism, nothing remains articulate.

Chapter 6

Nihilism

6.1 DESPAIR

Sorites paradoxes appear to show that vague expressions are empty; any vaguely drawn distinction is subverted. Most philosophies of vagueness postulate a reality at odds with this appearance; vague expressions are assigned meanings to which sorites reasoning is not faithful. Frege took a contrary view: no genuine distinction can be vaguely drawn; since vague expressions are not properly meaningful, there is nothing for sorites reasoning to betray; they are empty. Such a view is *nihilist.*

Is nihilism tenable? Is it compelling? If it is not tenable, one must hope that it is not compelling. If it is tenable, it will be compelling only if no alternative can be made to work, for it is a desperate view. Almost every expression of our language is vague enough to be sorites-susceptible, and so empty by nihilist standards. Since the alternatives to nihilism are discussed in other chapters, the main question to be addressed here is its tenability.

If some vague expressions are non-empty, it does not follow that all are. A local nihilism may be tenable or compelling even if global nihilism is not.[1] To classify all vague expressions as empty may amount to intellectual suicide; perhaps we can live with such a classification of a restricted class of them. A local nihilism may be less desperate than the global doctrine. We might find it compelling, not as a last resort, but in response to features specific to expressions of the restricted class. If a term can be vague through either incompleteness or inconsistency in the rules governing its use, then perhaps only inconsistency makes it empty. The most salient case for local nihilism concerns observational terms. Their dependence on our limited powers of discrimination has been held to involve inconsistency. That claim is investigated in Sections 6.3–6.4.

6.2 GLOBAL NIHILISM[2]

Dissatisfied with all attempts to say what is wrong with sorites arguments, one may be tempted by the simple thought that nothing is wrong with them: a typical sorites argument is sound, its conclusion strange but true.[3] The consequences of this thought are in flagrant violation of common sense beliefs; no doubt anyone who assents to the former in theory will continue to act on the latter in practice. The violation of common sense may be a good reason to reject the simple thought. Unfortunately, it may also be a good reason to reject every theory of vagueness ever proposed. No logic for a vague language seems wholly congenial to common sense. Indeed, that sorites paradoxes expose inconsistencies in common sense is not a far-fetched view. The truth about vagueness must be strange. We may therefore try to prescind from common sense and ask whether the simple thought, or anything like it, gives a theoretically stable view.

Consider ten thousand men in order of decreasing wealth. By everyday standards the first is very rich, the last very poor, and none significantly wealthier than the next. A typical sorites argument starts from the major premise that, for all i from 1 to 9,999, if the ith man is rich, then the $(i+1)$th man is rich, and the minor premise that the first man is rich; it concludes that the ten-thousandth man is rich. If the simple thought is to be believed, the argument is sound: it is valid and its premises are true; its conclusion is therefore true too. The ten-thousandth man, by everyday standards very poor, is in fact rich.

A moment's reflection shows that the simple thought is not to be believed. An equally typical sorites argument starts from the major premise that, for all i from 9,999 to 1, if the $(i+1)$th man is poor, then the ith man is poor, and the minor premise that the ten-thousandth man is poor; it concludes that the first man is poor. The simple thought endorses the latter argument just as it endorses the former. Yet they cannot both be sound, for the minor premise of each is inconsistent with the conclusion of the other. No one can be both rich and poor (at the same time and in the same respect).

It would not even help to deny that obvious truth, for a third typical sorites argument starts from the major premise that, for all i from 9,999 to 1, if the $(i + 1)$th man is not rich, then the ith man is not rich, and the minor premise that the ten-thousandth man is not rich; it concludes that the first man is not rich. This argument and the first cannot both be sound, for the minor premise of each explicitly contradicts the conclusion of the other.

Indiscriminate trust in sorites arguments leads to no consistent view, however bizarre. That is hardly surprising. The usual reason for believing the minor premise of a sorites paradox is of exactly the same kind as the usual reason for disbelieving its conclusion. They are common sense reasons, yet the simple thought involved a withdrawal from common sense. A typical sorites argument could still be used in an attempted *reductio ad absurdum* of common sense, but not as a positive reason for believing its conclusion. What this shows is that the simple thought does not do justice to its motivation. The original attitude was dissatisfaction with every diagnosis of a general error in sorites arguments.[4] More specifically, it was dissatisfaction with every attempt to assess the *major* premise of a sorites argument as meaningful but less than perfectly true. Every such attempt, it is held, travesties the vagueness of the relevant expression at some point. This attitude is so far quite neutral over the minor premise, whose endorsement by common sense will be treated as at best defeasible evidence in its favour. What the attitude entails is just that the minor premise of a typical sorites argument is no more true than its conclusion.[5] If the first man is rich, then the last man is rich; if the last man is poor, then the first man is poor. So if anyone is rich, everyone is rich; if anyone is poor, everyone is poor. These conditionals appear to be consistent with each other. Contrary to appearances, we live in an egalitarian world. Similar claims would follow about other vague simple predicates. The result may be called the *all-or-nothing view.*[6]

The all-or-nothing view can be explored in a literal-minded way. For the present case, it has four options: (a) ‘rich’ and ‘poor’ apply to everyone; (b) ‘rich’ applies to everyone and ‘poor’ to no one; (c) ‘rich’ applies to no one and ‘poor’ to everyone; (d) ‘rich’ and ‘poor’ apply to no one. The trouble with (a) is that ‘rich’ and ‘poor’ are antonyms. (b) and (c) treat ‘rich’ and ‘poor’ with implausible asymmetry. Although (d) involves an asymmetry between ‘rich’ and ‘not rich’, that is less implausible. ‘Rich’ and ‘poor’ presumably have the same amount of semantic structure, whereas 'not rich' is more complex than 'rich'. (d) generalizes to the view that vague simple predicates are true of nothing; neither (b) nor (c) generalizes smoothly. Only a smoothly generalizable claim is theoretically attractive. The all-or-nothing theorist will endorse (d). Moreover, the sorites arguments for 'rich' and 'poor' can easily be extended to counterfactual circumstances. It is not a contingent fact that we live in an egalitarian world. Thus the all-or-nothing theorist will say that 'rich' and 'poor' could not apply to anyone.[7]

To claim that vague predicates are necessarily not true of anything is not yet to claim that they are semantically defective. The predicate 'is a prime number greater than all other prime numbers' is necessarily not true of anything, but it is not semantically defective, for it occurs in sentences that constitute a sound proof that there is no such number. To stand for a necessarily unpossessed property is not to stand for no property at all. However, the all-or-nothing theorist cannot easily resist the further claim that vague predicates are semantically defective. The word 'rich' is usually explained by examples. The explanation may succeed even if the person pointed at happens to be bankrupt, but it seems to require at least that the learner can envisage possible circumstances in which someone would have the relevant property. Contrast 'prime number greater than all other prime numbers', whose meaning depends on the meanings of its parts. If 'rich' is defined as 'richer than most', the main problem switches to the vague quantifier 'most'. Sorites arguments can be formulated for 'Most *F*s are *G*s' in which the ratio of *F*s that are *G*s to *F*s varies gradually. If nothing of the form 'Most *F*s are *G*s' can be true, then 'most' surely is semantically defective, for how could it be explained? Consequently, if 'rich' is defined as 'richer than most', then 'rich' is semantically defective, because it has a semantically defective constituent. Again, if we could work out from our understanding of 'rich' that it was false of everyone, then it would cease to have any genuine borderline cases, and so could hardly count as vague. The nihilist would be better advised to treat 'rich' as semantically defective.

A semantically defective expression need not be thought of as sheer gibberish. If the name 'Moses' as used in the Pentateuch has no bearer, it is semantically defective, but that does not reduce it to a mere string of letters. It is empty, but there is a hole where its referent should be. It is a failed *name*, not, for example, a failed verb. The all-or-nothing theorist might say something similar of 'rich'. It stands for no property and so is semantically defective, but it is not a mere string of letters; it is a failed *adjective.*

Theoretically the least unattractive option for the all-or-nothing theorist is nihilism, the view that vague expressions are empty. In the long run, however, the choice may not make much difference. Whatever option is taken, vague lexical items will not significantly enhance the referential power of a language. The addition of predicates necessarily true of everything or necessarily true of nothing to a language with a modicum of logical vocabulary does not increase the class of definable properties.[8]The same goes for other syntactic categories, such as the quantitiers. For the all-or-nothing theorist, whatever can be said can be said precisely. How much does this leave us to say? For definiteness, the discussion will focus on nihilism, but similar remarks apply to other all-or-nothing views.

Some nihilists might try to mitigate the implausibility of their doctrine by saying that common sense beliefs are useful, if not true. They help us to get about in the world. That may be so, but the sentence 'Common sense beliefs are useful' is evidently vague. 'Common sense', 'belief' and 'useful' are all soritessusceptible. By nihilist standards, 'Common sense beliefs are useful' is not true. It would not help the nihilist to say that it is useful. Not only is 'useful' sorites-susceptible; if usefulness is the standard by which the debate is to be judged, nihilism loses at once. Only if truth is the standard has it a chance. A nihilist striving to say only what is true should not say that common sense beliefs are useful, or that it is useful to believe that common sense beliefs are useful.

The nihilist may resort to stipulation in order to define precise terms in which to articulate a defence of nihilism. This would work only if the stipulations were themselves made in precise terms. A citadel of precision is needed from which to stipulate outwards.[9] What is that citadel? The language of mathematics and logic will not do, for empirical concepts are not definable in such terms. If anything will do, it is the language of physics. But the nihilist will not find it easy to maintain that the language of physics is perfectly precise. A sorites argument could be constructed for the statement 'There are electrons', starting with the actual situation and moving through a series of counterfactual situations increasingly unlike it. If the counterfactual element is resisted (although it is common in nihilist arguments), sorites arguments could be constructed about the actual world.

For example, ‘contains an electron’ does not seem to be a perfectly precise predicate of space–time regions when small enough variations in their boundaries are considered. The vagueness is compounded if a qualification such as ‘with probability 90 per cent’ is added. There is little for the thorough-going nihilist to say.

We can use our vague language to make our language less vague, by a gradual process of clarification. This does not help the nihilist unless the product is perfectly precise. If it is merely less vague, then the steps of the sorites argument must be correspondingly smaller, but the upshot will be the same as before. When the nihilist is ready to claim perfect precision for some empirical vocabulary, it will be time to examine the claim. Nor is the problem confined to language. Our concepts are as sorites-susceptible as our words. There is no hope for the nihilist in the idea that, although our language is vague, our thoughts are precise.

The nihilist may accept that precision cannot be achieved in practice, while suggesting that we can talk as though it had been. We are to pretend that our words have precise meanings, although there are no precise meanings that we are to pretend that they have. An immediate difficulty is that the word ‘we’ is sorites-susceptible, and therefore empty on nihilist terms, both because it is vague exactly who counts as one of us, and because, for each one of us, it is vague exactly where his or her spatiotemporal boundaries lie. Thus an impersonal formulation may be better: it is to be pretended that words have precise meanings. But the content of this pretence is itself vague. The word ‘word’, for example, is sorites-susceptible, both because it is vague exactly what counts as a word type, and because, for each word type, it is vague exactly what counts as a token of that type, and because, for each token of a word type, it is vague exactly where the spatio-temporal boundaries of that token lie. Thus the pretence is empty on nihilist terms. In any case, it is idle to pretend merely that a word has some precise meaning or other; what is needed is the pretence, of some particular kind of precise meaning, that the word has a meaning of that kind. But for the nihilist the words with which such a kind might be delineated are empty.

The foregoing arguments do not show nihilism to be logically inconsistent. After all, even the extreme nihilist claim that nothing exists is logically consistent. However, not every logically consistent claim need be taken seriously. ‘Nothing exists’ is incapable of being a true sentence in its present sense, for if the sentence is true then it itself is something. The form of nihilism shaped by sorites arguments is somewhat different, but seems equally unlikely to receive true expression, whether in language or thought. Such an expression would be an utterance, inscription, thought, or the like; once again, ‘utterance’, ‘inscription’, ‘thought’ and ‘expression’ itself are sorites-susceptible, and so would be empty if nihilism were true. Philosophy becomes superficial when it omits itself from the world it describes.

The attempt to develop nihilism as a positive doctrine fails on its own terms. At best the nihilist can supply arguments by *reductio ad absurdum* against the alternatives. If all those arguments work, we are in a bad way. It is the task of the next chapters to argue that at least one of them does not work.

6.3 LOCAL NIHILISM: APPEARANCES

Only a restricted nihilism is a tenable position. Nor will it do for the nihilist to endorse a restricted class of sorites arguments if the grounds for endorsing them apply without restriction. For example, it would be arbitrary to use sorites reasoning against ‘There are tables’ but not against ‘There are electrons’, if the grounds for accepting it in the former case are analogous to grounds for accepting it in the latter. Local nihilism requires an inherently restricted reason to regard some vague expressions as incoherent.

One thought begins as follows. For just about any vague expression, an appropriate sorites series can be constructed. In order to use the expression correctly, one is obliged to treat the first member of the series differently from the last, but at no point in the series is one obliged to treat successive members differently. Since what is not obligatory may nevertheless be permissible, two people may draw the line at different more or less arbitrary points and both count as using the expression correctly. But for vague expressions of a restricted kind, a sorites series can be constructed to meet a stronger condition. In order to use the expression correctly, one is obliged to treat the first member of the series differently from the last, but at every point in the series one is obliged *not* to treat successive members differently. Confronted with the sorites series, one cannot use the expression correctly. In that way the expression is semantically incoherent, a point made in the material mode by the corresponding sorites paradox. This incoherence would not extend to vague expressions outside the restricted kind.

Why should correct use of an expression forbid one to treat successive members of a sorites series differently? Well, suppose that there is no *apparent* difference between successive members in a certain respect, and that whether the expression '*F*' applies to something depends solely on its appearance in that respect. It is then tempting to argue that since no difference between successive members is relevant to '*F*', if '*F*' applies to one member of the series then it must also apply to the next. For example, whether the predicate 'looks square' applies to something depends solely on its visual appearance in respect of shape. Whether the object is square is not a matter of its appearance, but whether it looks square is simply a matter of how it looks. A sorites series of rectangles can be constructed, running from a perfect square to a rectangle whose length is many times its height, such that successive members of the series look the same in shape to normal observers in normal conditions of observation. Neighbouring rectangles are not merely indiscriminable in shape by the naked eye; they are seen as positively having the same shape. Does it not follow that if one rectangle in the series looks square, then so does the next? If so, the major premise of the sorites paradox is underwritten, not by the mere vagueness of 'looks square', but by the fact that 'looks square' applies only in virtue of appearances. Contrast a sorites paradox for the vague predicate 'squarish'; since to be squarish is to be (not: look) roughly square, squarishness is not simply a matter of appearance. There is not the same reason to suppose that if two things look the same in shape and one of them is squarish, then so is the other.[10]

In the rest of this chapter, several different attempts to make such a local nihilist argument rigorous will be examined and seen to fail for interrelated reasons. The coherence of the vague terms in question is defensible without any appeal to non-classical logic or semantics. However, the eventual failure of local nihilist arguments will involve the falsity of other assumptions that are by no means confined to local nihilists.

If two things have exactly the same shape, and one of them is square, it follows uncontroversially that the other is square too. The suggestion is that if it *looks* as though the premises of such a logically valid inference are true, then it *looks* as though its conclusion is also true. Thus if it looks as though two things have exactly the same shape, and it looks as though one of them is square, it is supposed to follow that it looks as though the other is square too. By repeated applications of this inference to the sorites series, one can move from the minor premise that it looks as though the first member is square to the conclusion that it looks as though the last member is square. Yet the first member is a perfect square while the last is a rectangle of length many times its height. If the notion of 'how things look' is coherent, the minor premise is true and the conclusion false. The local nihilist concludes that 'how things look' is not a coherent notion.

Our limited powers of discrimination make sorites arguments a special problem for expressions like 'looks square'. If the problem is insoluble, then predicates of appearance are incoherent, and for an inherently restricted reason. The reason would not extend to everything we say about what we perceive, still less to everything vague we say. It is specific to terms whose application is supposed to depend solely on 'how things appear'. If the right moral were that perception does not involve a dimension of pure appearance, that is something with which we could probably learn to live. On most occasions when we now use 'looks square' we could use 'squarish' instead. Of course, the two phrases are not equivalent, otherwise 'squarish' would be no improvement on 'looks square'. In some cases of severe distortion, a squarish thing does not look square or a square-looking thing is not squarish (as we say). But such cases are rare; the loss of 'looks square' might be manageable for practical purposes.

Unfortunately for the local nihilist, the argument above for the incoherence of 'looks square' is not watertight. One may challenge the assumption of the deductive closure of 'looks', that if it *looks* as though the premises of a logically valid inference hold, then it *looks* as though its conclusion also holds. Counterexamples threaten when the inference is too complex to be taken in at a glance, and when the conclusion introduces concepts not found in the premises. If 'looks' is not deductively closed, then even if it looks as though two things have exactly the same shape, and it looks as though one of them is square, how can one infer that it looks as though the other is square too? However, the inference from 'This and that have exactly the same shape' and 'This is square' to 'That is square' is simple enough to be taken in at a glance, and its conclusion does not introduce any concepts not found in the premises. In the present case, the deductive closure of 'looks' will be accepted for the sake of argument, for the nihilist argument commits another fallacy: it equivocates over the context of the looking.

Let x_{i-1}, x_i and x_{i+1} be three successive members of the sorites series. No doubt it looks to the observer when she compares x_{i-1} and x_i as though they have exactly the same shape, and, granted the deductive closure of 'looks', it follows that if it looks to her in that context as though x_{i-1} is square then it also looks to her in that context as though x_i is square. Similarly, no doubt it looks to her when she compares x_i and x_{i+1} as though they have exactly the same shape, and, granted the deductive closure of 'looks', it follows that if it looks to her in the new context as though x_i is square then it also looks to her in the new context as though x_{i+1} is square. However, one cannot put the two conditionals together in the intended way, for the consequent of the former does not imply the antecedent of the latter. Although it may look to the observer when she compares x_{i-1} and x_i as though x_i is square, it in no way follows that it looks to her when she compares x_i and x_{i+1} as though x_i, is square. Whether a rectangle looks square may depend on what it is compared with. A normal observer may be more likely to see it as square when it is in the company of something nearer than it to being square than when it is in the company of something less near.[11]

The nihilist might seek to avoid the problem of contextual change by supposing that each rectangle is seen only once, and by itself. One sees x_i; it looks as though x_i is square. It is then replaced by x_{i+1}; it now looks as though x_{i+1} has exactly the same shape as x_i had. However, even granted the deductive closure of 'looks', it does not follow that it now looks as though x_{i+1} is square, unless it *now* looks as though x_i was square. For looks can change over time; the deductive closure of 'looks' has a chance of being valid only if the premises and conclusion concern looks at the same time. In what sense does it still look as though x_i was square, now that it is no longer seen? One might allow that it looks as though an unseen object had a property when present visual perception provides a basis of some kind for the judgement that the object had the property; to the hunter's experienced eye, examining the tracks, it may look as though the unseen deer was wounded. However, there is no guarantee that in this sense it does still look, once x_i has been replaced by x_{i+1}, as though x_i was square. Perhaps $x_{\mathrm{i}+1}$ looks as though it might well not be square, and the sight of it prompts the observer to doubt, as she had not before, whether even the very similar x_i was square in the first place.

An alternative move for the nihilist is to use a dispositional notion of looking. The way objects dispositionally look does not depend on what is currently seen, but on the way those objects would occurrently look if they were seen, perhaps under normal conditions. To compare the rectangle x_i with $x_{\mathrm{i}+1}$ after x_{i-1} is not to change its dispositional looks. But this move merely shifts the problem, for 'looks' in the dispositional sense cannot be expected to satisfy deductive closure. Suppose that it dispositionally looks as though x_i is square and it dispositionally looks as though xi and x_{i+1} have exactly the same shape. Thus when x_i is seen under normal conditions, it occurrently looks as though x_i is square; when x_i and x_{i+1} are seen together under normal conditions, it occurrently looks as though xi and x_{i+1} have exactly the same shape. It does not follow that when x_{i+1} is seen under normal conditions it occurrently looks as though x_{i+1} is square, even granted the deductive closure of occurrent looking. For the premises say nothing about the way things occurrently look when x_{i+1} is seen under normal conditions in the absence of x_i. Given the contextual effects noted above, in those circumstances it may well not occurrently look as though x_{i+1} is square. Thus it does not follow from the premises that it dispositionally looks as though x_{i+1} is square. Dispositional looking is not deductively closed.

In order to restore deductive closure, the nihilist might abstract even further from specific perceptual contexts. The phrase 'It looks on consideration as though . . .' may be used for the judgements it would be reasonable to make in the light of all the relevant perceptions – for example, after one has seen the whole sorites series. Considered looking in this sense should satisfy deductive closure, for it represents an equilibrium attained by judgement through reflection on perceptual information under the constraint of reason that propositions are to be accepted only if their logical consequences are to be accepted. The difficulty for the nihilist is that it does not look on consideration as though x_i and x_{i+1} have exactly the same shape. Once one has seen the whole series, one can work out that many of its members do not have exactly the same shape as their immediate successors; x_i and x_{i+1} are as likely to be a case in point as any other pair. It would not be reasonable to judge that x_i and x_{i+1} have exactly the same shape.

The nihilist might use the deductive closure of considered looking in a different way. If x_i is square and x_{i+1} is not square, it follows that x_i and x_{i+1} do not have exactly the same shape. By deductive closure, if it looks on consideration as though x_i is square and it looks on consideration as though x_{i+1} is not square, then it looks on consideration as though x_i and x_{i+1} do not have exactly the same shape. But it does not look on consideration positively as though x_i and x_{i+1} do not have exactly the same shape. For all a normal observer can tell by the naked eye, they have exactly the same shape, differences too small to be discerned occurring only elsewhere in the sorites series. Thus if it looks on consideration as though x_i is square, then it does not look on consideration as though x_{i+1} is not square. But that is not what the nihilist needs. To move from one stage of the sorites paradox to the next, the required conditional is that if it looks on consideration as though x_i is square, then it looks on consideration as though x_{i+1} is square.[12] To fill the gap, the nihilist must argue that if it does not look on consideration as though x_{i+1} is not square, then it looks on consideration positively as though x_{i+1} is square. The nihilist would presumably so argue by appeal to the general principle that either it looks on consideration as though things are a certain way or it looks on consideration as though they are not that way. However, that principle is not plausible. Agnosticism is sometimes the reasonable path. When one has seen the whole series, it may be neither reasonable to judge that x_{i+1} is square nor reasonable to judge that it is not.

The nihilist may feel balked by the fact that being square is a mind-independent condition in a sense in which looking square is not. Sometimes one's information is simply too limited to decide for or against the squareness of a thing, obliging one to remain agnostic. The nihilist may therefore seek to rescue the preceding argument by replacing the property of being square in it by the property of looking square. The result might run as follows.

Since x_i and x_{i+1} cannot be distinguished by the naked eye, it would not be reasonable to judge after inspection of the sorites series that it looks on consideration as though x_i is square but does not look on consideration as though x_{i+1} is square. Thus if it looks on consideration as though it looks on consideration as though x_i is square, then it does not look on consideration as though it does not look on consideration as though x_{i+1} is square. That is not yet what the nihilist needs. The plan is to construct a sorites paradox from the apparently true proposition that it looks on consideration as though it looks on consideration as though x_0 is square to the apparently false proposition that it looks on consideration as though it looks on consideration as though x_n is square. To move from one stage of this sorites paradox to the next, the required conditional is that if it looks on consideration as though it looks on consideration as though x_i is square, then it looks on consideration as though it looks on consideration as though x_{i+1} is square. To fill the gap between what has already been established and that conditional, the nihilist must argue that if it does not look on consideration as though it does not look on consideration as though x_{i+1} is square, then it looks on consideration positively as though it looks on consideration as though x_{i+1} is square. However, the nihilist need no longer argue so by appeal to the implausible general principle that either it looks on consideration as though things are a certain way or it looks on consideration as though they are not that way. It is enough to appeal to the more limited principle that either it looks on consideration as though it looks on consideration as though things are a certain way or it looks on consideration as though it does not look on consideration as though things are that way. Call that principle *the transparency of considered looking.*

The idea behind the transparency of considered looking is that, although one cannot decide some questions, one can decide which questions have been decided. That is usually the case in mathematics: a proposition is either known to have been proved, known to have been refuted, or known to have been neither proved nor refuted. Even there, however, the principle is an idealization. The status of a purported proof may be uncertain for months, until enough experts have had a chance to examine it. In empirical matters, the transparency of considered looking is quite implausible. It is often unclear what assertions are warranted by the evidence in physics or history, and the same goes for perceptual warrant. One may be justifiably agnostic as to whether one has enough warrant to make an assertion. In borderline cases, warranted assertibility is not a decidable matter. The mind is not transparent to itself; the conception of warranted assertibility as everywhere decidable is just the self-transparency of mind given a linguistic turn.

It may be assumed, as an instance of the law of excluded middle, that either it looks on consideration as though x_{i+1} is square or it does not look on consideration as though x_{i+1} is square; but the nihilist is not entitled to assume that if it looks on consideration as though x_{i+1} is square then it looks on consideration as though it looks on consideration as though x_{i+1} is square, or that if it does not look on consideration as though x_{i+1} is square then it looks on consideration as though it does not look on consideration as though x_{i+1} is square. Considered looking is not transparent, and the argument for the major premise of the sorites argument fails.

That warranted assertibility is not decidable might be thought to show that our language is not fully governed by rules. This would follow if rules were algorithms telling one in all situations just what, if anything, to say. For any declarative sentence, the rules would effect a threefold classification of all situations into those warranting assertion, those warranting denial, and those not warranting either; any situation in one of those classes would be clearly in it. Such classifications are not to be found. There are always hard cases demanding good judgement. Of course, it may be doubted whether the primitive algorithmic conception of rules is the only one possible.

We cannot always tell by looking whether something looks square, not even whether it occurrently looks square. I may judge that it looks square and later reasonably decide that I had misdescribed it; after all, it merely looked squarish. I may hesitate to judge that it looks square and later reasonably decide that I had been over-cautious; after all, it did *look* square.

Looking is no doubt the best way to tell whether something looks square, but the best way is not an infallible way.

The local nihilist might attempt to recast the argument by speaking of the look of a thing. When xi and x_{i+1} are successive members of the sorites series, the look of x_i and the look of x_{i+1} are identical. By Leibniz's law, identical objects have the same properties. Thus any property of the look of x_i is a property of the look of x_{i+1}. Now the look of an object cannot have the property of literally being square, for the look, unlike the object, is not literally in space. However, it can have the property that the look of a square object has in normal conditions, which may be called squarelikeness. Thus the look of x_0 is squarelike and the look of x_n is not squarelike. But by what has just been said, if the look of x_i is squarelike then the look of x_{i+1} is squarelike. The new sorites paradox threatens 'squarelike', indeed the very concept of the look of a thing.

Why does the local nihilist suppose that the look of x_i and the look of x_{i+1} are identical? The reason is presumably that a normal observer cannot discriminate with the naked eye between x_i and x_{i+1} when they are presented successively. Thus the indiscriminability of x_i and x_{i+1} is supposed to be a sufficient condition for the identity of their looks; it is certainly a necessary condition. But indiscriminability of this sort cannot be a necessary and sufficient condition for the identity of anything, for, unlike identity, it is not a transitive relation. If u is indiscriminable from v, and v from w, it does not follow that u is indiscriminable from w. In contrast, if the look of u is identical with the look of v, and the look of v with the look of w, then it does follow that the look of u is identical with the look of w. Thus the new local nihilist argument relies on an immediately incoherent conception of the 'look' of an object; there is no need to mention 'squarelike'. The repercussions of this result are too limited to satisfy the local nihilist. It does not threaten the coherence of the propositional uses of 'looks' discussed above, nor does it prevent one from developing a notion of the looks of things for the identity of which the indiscriminability of those things is necessary but not sufficient.[13] Nihilism was not intended to be as local as that.

6.4 LOCAL NIHILISM: COLOURS

Predicates of appearance are not as such paradoxical, for not even appearances are just what they appear to be. The gap between squareness and the appearance of squareness generates a corresponding gap between the appearance of squareness and the appearance of the appearance of squareness. However, there are some terms whose meaning might be thought to abolish the gap between reality and appearance from the very beginning. If so, it will not be merely because they describe appearances, for 'looks square' does that. Rather, it will be because they somehow succeed in singling out fixed points of the function that maps reality to appearance; i.e. they will express properties whose reality *is* their appearance. Such terms may be called *observational.* The word 'red' is traditionally supposed to be observational. Of course, in a strange light or to a colour-blind observer non-red things may look red and red things may not. But it is supposed that in normal conditions something will look red to a normal observer just in case it is red.

Although there are notorious problems in specifying what normal conditions and normal observers are without circularity, they need not worry the nihilist, for whom the important point is that there are supposed to be such conditions and observers, however they are specified. For the nihilist can present a normal observer in normal conditions with a sorites series for 'red', whose successive members are so similar as to be indiscriminable in colour by the naked eye. Of each member, the observer is asked 'Is it red?'. If every possible attempt to answer leads to absurdity, then 'red' is incoherent.

There are many ways in which the nihilist might attempt to argue. Some of them appeal just to the idea that 'red' is used to describe appearances; in doing so, they commit variants of the fallacies already noted with respect to 'looks square'. However, a more powerful argument is also available to the local nihilist.

Suppose that conditions and the observer are normal. The observer understands the word 'red'; indeed, it will be convenient to build that into the definition of 'normal'. The observer scrutinizes the sorites series. Let x_i and x_{i+1} be successive members. On the account above, in normal conditions something is red if and only if it is recognizably red to a normal observer. In the circumstances, the following conditional should therefore hold:

(1) If x_i is red, then the observer knows that x_i is red.

Now suppose that the observer does know that x_i is red. A necessary (if not sufficient) condition for this is that the observer's classification of x_i as red should be the outcome of a reliable capacity to classify things correctly as red. Such a capacity need not be infallible, for misclassifications can occur in abnormal conditions. Nevertheless, its reliability does imply that in the present normal conditions it will issue in correct classifications of neighbouring members of the sorites series. Now an observer who is disposed to classify x_i as red and who cannot discriminate between xn and x_{i+1} in colour will be liable to classify x_{i+1} as red, even if the disposition to do so may be less strong than in the case of x_i. Unless x_{i+1} is red, such a classification would be a misclassification. Since x_i was reliably classified as red, it follows that x_{i+1} is red, whether or not the observer classifies it as such. The upshot is the following conditional:

(2) If the observer knows that x_i is red, then x_{i+1} is red.

(1) and (2) entail:[14]

(3) If x_i is red, then x_{i+1} is red.

(3) is just what is needed to move a sorites paradox for 'red' on from one stage to the next. Has the nihilist demonstrated that 'red' is incoherent?

It will be assumed in what follows that (2) is correct. The argument for it certainly needs more discussion, which will be supplied in Chapter 8, where a generalization of (2) is endorsed. The focus will therefore be on (1). In the nihilist's favour, it may be noted that many non-nihilists accept something like (1), on something like the grounds above. The nihilist's appeal to (1) is therefore not question-begging in any crude way. But an objection to (1) is implicit in the defence of (2). Suppose that x_i is red, but indiscriminable in colour by the observer from something non-red. Then the observer will *not* know that x_i is red, for a classification of x_i as red would not be the outcome of a locally reliable capacity to classify things as red. However, objections to (1) are not really what is needed. For if the nihilist is right and ‘red’ is incoherent, one would expect compelling arguments both for and against principles such as (1). To defuse the paradox, one must show that the argument for (1) is not compelling, not that an argument against it is compelling.

The argument for (1) runs roughly as follows. In so far as ‘red’ has boundaries, they are not natural ones. We could easily have used the word with a slightly broader or narrower extension; the term would not thereby have failed to cut nature at its joints in any sense in which it actually succeeds in doing so. The red things do not form a natural kind. Rather, the extension of ‘red’ is fixed just by our dispositions to classify things under ‘red’ or ‘not red’. More exactly, our practice of using the word assigns a canonical status to the dispositions of certain observers in certain conditions; we call them ‘normal’. If something is red, it falls in the extension of ‘red’, and that is because normal observers in normal conditions classify it under ‘red’. Since such classifications are thereby guaranteed to be correct, they constitute knowledge. Thus (1) is true, or so the argument goes.

If the argument for (1) works at all, it should work under the simplifying assumption that something is red if and only if in normal conditions all normal observers always classify it as red. That assumption is certainly too crude to be true, if only because it makes no allowance for everyday mistakes, but it simplifies in a direction that should be helpful to the argument for (1).[15] Suppose that x_i is red. Thus in normal conditions all normal observers always classify x_i as red. Does it follow that in classifying x_i as red in normal conditions, a normal observer knows that x_i is red? By the argument for (2), the observer knows that x_i is red only if that classification is the outcome of a locally reliable capacity to classify things as red. The argument for (2) also assumes that the normal observer in question cannot discriminate in colour between x_i and x_{i+1}. The colour discrimination of a normal observer need not be perfect; it can be that of an ordinary human being with good eyesight, for it is creatures of the latter sort who use ‘red’, and whose use determines its extension. Thus the normal observer who classifies x_i as red in normal conditions is liable sometimes to classify x_{i+1} as red in those conditions too. Such a classification is not automatically correct. By the operative assumption about ‘red’, it is correct to classify x_{i+1} as red only if in normal conditions all normal observers always classify x_{i+1} as red. What has been granted so far is that, in normal conditions, all normal observers always classify xi as red and some normal observers sometimes classify x_{i+1} as red; it by no means follows that all normal observers always classify x_{i+1} as red. Between x_i and x_{i+1} the consensus may begin to crack, and classification as red become a probabilistic matter. If so, it is not true that x_{i+1} is red, and the classification of x_i as red is therefore not the outcome of a locally reliable capacity to classify things correctly as red. In correctly classifying x_i as red in normal conditions, a normal observer need not know that x_i is red.

The argument for (1) breaks down even under the simplest assumption helpful to it about the way in which the extension of ‘red’ is determined. Its chances are, if anything, worse under more realistic assumptions, on which, for example, normal observers in normal conditions sometimes make everyday mistakes as to which things are red. In effect, the objection from (2) to (1) can be made within the terms of the argument for (1), and locates the point at which that argument breaks down. The local nihilist could meet the objection to the argument for (1) only by recasting (1) in terms on which (2) failed; but then the argument for (3) would be lost, and with it the intended sorites paradox.

The plausibility of (2) suggests one way in which the account of colour might be made more realistic. The spontaneous colour judgements of normal observers in normal conditions do not simply tend to be true; they also tend to be knowledgeable. Any account of colour terms should meet this condition. Given (2), the account must therefore include more in the extension of ‘red’ than normal observers in normal conditions tend to call ‘red’, and more in the extension of ‘not red’ than such observers tend to call ‘not red’. Thus no simple equation between the extensions of colour terms and the classifications made by normal observers in normal conditions is to be expected.

The failure of (1) might be thought to threaten the usefulness of ‘red’. Is not the point of the term that we can apply it on the basis of casual observation? (1) fails because some red things cannot be known by observation to be red. However, the red things in question are only the marginally red ones. (2) presents no obstacle to knowing by casual observation of a non-marginally red thing that it is red. Such knowledge is enough to make ‘red’ useful.

Granted (2), how should a normal observer respond to successive members of a sorites series for ‘red’ in normal conditions? Evidently, the last ‘This is red’ should not be immediately followed by a ‘This is not red’, for, by (2), if the former response is known to be correct then the latter is incorrect. Even ‘I don’t know whether this is red’ was seen in Section 1.2 to be a perilous response. However, a strategy of sorts is open to the observer. For suppose that ‘*A*’ is a correct response to a given member of the series; then ‘For all I know, *A*’ (‘I don’t know that not *A*’) is a correct response to the next member.[16] Correspondingly, if one knows ‘*A*’ to be a correct response to a given member, then one is in a position to know ‘For all I know, *A*’ to be a correct response to the next member. Each time one feels the onset of doubt with a given response, one can add another iteration of ‘For all I know’ to it, until one feels sufficient confidence in some new and simpler response, such as ‘This is not red’. For those who cannot bear to be silent, there is always something to say.

In search of a sorites paradox, the local nihilist might try other permutations of the material considered in this chapter, but the futility of the search should by now be evident. Concepts associated with coarsely discriminating recognitional capacities are not as such paradoxical. The limits on our powers of discrimination are intelligible as cognitive phenomena, through principles such as (2). Such principles are at the heart of a provably consistent account of inexact knowledge, to be developed in Chapter 8. Perhaps the impression of paradox stems from a misunderstanding of the insight behind (2); ‘the observer knows that’ is either wrongly omitted from the antecedent or wrongly added to the consequent. Both mistakes result in sorites paradoxes; (2) itself is not paradoxical. This diagnosis can be extended to sorites paradoxes for all vague terms, whether or not they are observational. Before that can be done, however, a positive understanding is needed of the logic and semantics of vague terms. The next chapter begins that task.

Chapter 7

Vagueness as ignorance

7.1 BIVALENCE AND IGNORANCE

No one knows whether I am thin. I am not clearly thin; I am not clearly not thin. The word 'thin' is too vague to enable an utterance of 'TW is thin' to be recognized as true or as false, however accurately my waist is measured and the result compared with vital statistics for the rest of the population. I am a borderline case for 'thin'. If you bet someone that the next person to enter the room will be thin, and I walk through the door, you will not know whether you are entitled to the winnings. Suppose that an utterance of 'TW is thin' is either true or false. Then since we do not know that TW is thin and do not know that TW is not thin, we are ignorant of something. Either 'TW is thin' expresses an unknown truth, or 'TW is not thin' does. We do not even have an idea how to find out whether TW is thin, given my actual measurements and those of the rest of the population. Arguably, we cannot know in the circumstances that TW is thin or that TW is not thin; in that sense, we are necessarily ignorant of something. Most work on vagueness has taken it for granted that these consequences are absurd. It therefore rejects the original supposition that an utterance of 'TW is thin' is either true or false. Borderline cases are held not to involve ignorance, on the grounds that there is no fact of the matter for us to know, hence nothing for us to be ignorant of. On this view, vague utterances in borderline cases are not bivalent.

By abandoning the assumption that vague utterances are bivalent, it is suggested, we free ourselves to understand the phenomena of vagueness. Many attempts have been made to implement the suggestion. Some of the most conspicuous have been examined in previous chapters, and seen to fail. They falsified the phenomena they were supposed to explain. More complex permutations can of course be tried without end, but it seems increasingly unlikely that any formal trick will turn out to unlock the puzzle (the phenomenon of higher-order vagueness is particularly resistant). It begins to look as though abandoning the assumption that vague utterances are bivalent makes vagueness no easier to understand.

If one abandons bivalence for vague utterances, one pays a high price. One can no longer apply classical truth-conditional semantics to them, and probably not even classical logic. Yet classical semantics and logic are vastly superior to the alternatives in simplicity, power, past success, and integration with theories in other domains. It would not be wholly unreasonable to insist on these grounds alone that bivalence *must* somehow apply to vague utterances, attributing any contrary appearances to our lack of insight. Not every anomaly falsifies a theory. That attitude might eventually cease to be tenable, if some non-classical treatment of vagueness were genuinely illuminating. No such treatment has been found.

We have at the very least good reason to re-examine the original objection to bivalence: the supposed absurdity of postulating ignorance in borderline cases. In this chapter, several attempts to find such an absurdity will be examined and refuted. In the next, it will be argued that ignorance in borderline cases is just what one would expect on the basis of independently grounded principles about knowledge. When those principles are misunderstood, the postulation of ignorance seems counter-intuitive. There is no need to insist, unconstructively, that there must be *something* wrong with the objections to bivalence in borderline cases. When articulate, they rest on identifiable fallacies. Properly understood as an epistemic phenomenon, vagueness provides no motive for revising classical semantics or logic, and in particular no motive for denying bivalence.[1]

As a prolegomenon to an epistemic account of vagueness, Section 7.2 provides a motive for not denying bivalence in borderline cases. Section 7.3 argues that it is more difficult than one might suppose to imagine what it would be like for vagueness to be a non-epistemic phenomenon. Later sections elucidate the epistemic view, and reply to objections.

7.2 BIVALENCE AND TRUTH

How is the principle of bivalence to be understood? It does not say that everything is either true or false, for no one supposes that a drop of water is true or false. The principle does not even apply to every meaningful sentence, or use of one to perform a speech act, for there is no need to suppose that a question or command is true or false. Nor does it apply to every well-formed declarative sentence. If a teacher pronounces 'He was there then' as a sample sentence of English, leaving 'he', 'there' and 'then' undetermined in reference, nothing has been said to be the case, truly or falsely. The principle of bivalence claims truth or falsity when, and only when, something has been said to be the case.

To say that something is the case, in the relevant sense, is not always to assert that it is. The notions of truth and falsity apply to suppositions as well as assertions. If something true is said by an utterance of 'Not *P*', it is because something false is said, but not asserted, by the component utterance of '*P*'. The same applies to other truth-functors.

Bivalence is often formulated with respect to the object of the saying, a proposition (statement, . . .). The principle then reads: every proposition is either true or false. However, on this reading it does not bear very directly on problems of vagueness. A philosopher might endorse bivalence for propositions, while treating vagueness as the failure of an utterance to express a unique proposition. On this view, a vague utterance in a borderline case expresses some true propositions and some false ones (a form of supervaluationism might result). There is no commitment to a bivalent classification of utterances, or to the ignorance on our part that such a classification implies. The problem of vagueness is a problem about the classification of utterances. To debate a form of bivalence in which the truth-bearers are propositions is to miss the point of the controversy.

In a relevant form of bivalence, the truth-bearers are (perhaps with a little artificiality) the utterances themselves. The principle is explicitly restricted to occasions when someone uses an utterance to say that something is the case, in brief (if again with a little artificiality), when the utterance says that something is the case. The principle may be formulated as a schema:

(B) If u says that P, then either u is true or u is false.

In (B), '*u*' is to be replaced by a name of an utterance and '*P*' by a declarative sentence whose inscription says that something is the case. Since the utterance named is presumably not a constituent of the relevant instance of (B), it need not be in English. Since the sentence in place of '*P*' is a constituent of that instance, it must be in English. Although the notion of saying in (B) is not perfectly precise, it is precise enough for present purposes.[2]

The notions of truth and falsity in (B) may be elucidated by Aristotle's dictum that 'To say of what is that is is not, or of what is not that it is, is false, while to say of what is that it is, or of what is not that it is not, is true'. These are the notions that Tarski intended his semantic definitions to capture.[3] Aristotle's account can be put in present terminology as follows:

(T) If u says that P, then u is true if and only if P.

(F) If u says that P, then u is false if and only if not P.

The role of '*u*' and '*P*' in (T) and (F) is as in (B). For example, a typical utterance of 'Snow is white' says that snow is white; it is therefore true if and only if snow is white, and false if and only if snow is not white. There may or may not be much more to be said about truth and falsity than (T) and (F) say, but they do at least seem central to our ordinary understanding of those notions.

Now suppose that an utterance u – for example, of 'TW is thin' – is a counterexample to (B). It must therefore verify the antecedent of (B) by saying that something is the case. For some sentence in place of '*P*' – for example, 'TW is thin' – we must have:

(0) u says that P.

The consequent of (B) must also be falsified, giving

(1) Not: either u is true or u is false.

Using (0), we can detach the consequents of (T) and (F):

(2a) u is true if and only if P.

(2b) u is false if and only if not P.

Now (2a) and (2b) allow us to substitute their right-hand sides for their left-hand sides in (1), giving

(3) Not: either *P* or not *P*.

By one of De Morgan's laws, the negation of a disjunction entails the conjunction of the negations of its disjuncts. Thus (3) yields:

(4) Not *P* and not not *P*.

The second conjunct of (4) contradicts the first. There is no need to eliminate the double negation; (4) already contradicts itself. The supposition of a counterexample to bivalence has led straight to a contradiction, through the elucidations of truth and falsity and some trivial logic.[4]

For a sentence of English without context-dependent expressions, instances of (2a) and (2b) may be assumed without explicit appeal to (T) and (F). One can use quotation marks to form designators of the particular inscriptions they enclose, and count those inscriptions as utterances in an extended sense. Then (2a) might say that 'TW is thin' is true if and only if TW is thin, and (2b) that 'TW is thin' is false if and only if TW is not thin. One can then argue as before. Even here, however, (T) and (F) seem to explain (2a) and (2b); it is because it says that TW is thin that 'TW is thin' is true if and only if TW is thin and false if and only if TW is not thin. As predicates of utterances, truth and falsity are disquotational when saying is. Be that as it may, the conclusion is as before: the supposition of a counterexample to bivalence leads to a contradiction.[5]

Some remarks about the logic of the argument are in order. The statements of truth- and falsity-conditions for *u* in (2a) and (2b) are used to equate the denial of bivalence in (1) with the denial of the law of excluded middle in (3).[6] The denial of excluded middle leads to a contradiction by an uncontroversial rule of inference; just about every logic proposed for a vague language allows that the denial of a disjunction entails the conjunction of the denials of its disjuncts. As for the absurdity of (4), no attempt will be made to argue with those who think it acceptable to contradict oneself. Furthermore, the inference from (1), (2a) and (2b) to (3) should not be controversial when the biconditionals in (2a) and (2b) are read as equating their two sides in semantic value. The semantic value of an expression encapsulates those features of it relevant to the truth-value of sentences that result from combining that expression with others by means of logical operators, such as negation and disjunction. If one's semantic theory deals in degree of truth, for example, then (2a) is to be read as implying that '*u* is true' and '*P*' are true to exactly the same degree, and similarly for (2b). If one's semantic theory deals in admissible interpretations, then (2a) is to be read as implying that '*u* is true' and '*P*' are true on exactly the same admissible interpretations, and similarly again for (2b). The substitution of the right-hand sides of (2a) and (2b) for their left-hand sides in (1) to yield (3) should therefore preserve the semantic value of (1) in (3), for the semantic value of the negation of a disjunction depends only on the semantic values of its disjuncts. Thus the whole weight of the argument is thrown onto (2a) and (2b), and through them onto (T) and (F).[7]

The rationale for (T) and (F) is simple. Given that an utterance says that TW is thin, what it takes for it to be true is just for TW to be thin, and what it takes for it to be false is for TW not to be thin. No more and no less is required. To put the condition for truth or falsity any higher or lower would be to misconceive the nature of truth or falsity.

It might be replied that if *u* says that *P* and is neither true nor false, then '*u* is true' is false while '*P*' is neither true nor false, so that the two sides of (2a) do not match in semantic value, and neither (2a) nor (T) is true. A parallel reply might be made to (2b) and (F). The trouble with this objection is that it does nothing to meet the rationale for (T) and (F). It gives no hint, when *u* says that TW is thin, of any way in which *u* could fail to be true, other than by TW failing to be thin, or of any way in which *u* could fail to be false, other than by TW failing to be not thin. The whole point of the original argument is that the supposition of a counterexample to (B) is inconsistent. It is therefore just what one would expect that the supposition leads to inconsistencies with fundamental principles about truth and falsity, such as (T) and (F). Such inconsistencies would undermine the argument only if they exposed limitations in the arguments *for* those principles, but they do no such thing. The arguments against (T) and (F) from the supposition of a counterexample to (B) completely miss the point of the rationale for (T) and (F). The argument against the supposition is therefore not undermined.

In a formal semantics it is consistent to label two properties 't' and 'f' and suppose that some sentences have neither (or both, for that matter). Evidently, such a manœeuvre shows nothing of philosophical interest. No connection has yet been made between the properties labelled 't' and 'f' and the properties of truth and falsity. If one claims that the former properties *are* the latter, one must be prepared to claim that they are governed by the same principles. What principles govern truth and falsity, if not (T) and (F)?

Formal semantic treatments of vague languages – many-valued logics, supervaluations and the like – are characteristically framed in a meta-language that is conceived as precise.[8] Thus one cannot say in the precise meta-language what utterances in the vague object-language say, for to do so one must speak vaguely; one can only make precise remarks about those vague utterances. Since the expressive limitations of such a meta-language render it incapable of giving the meanings of object-language utterances, it can hardly be regarded as adequate for a genuine semantic treatment of the object-language. Certainly, the argument against supposed counterexamples to bivalence cannot be expressed in such a meta-language. For it requires one to say what the vague utterance says, in order to express the relevant instances of (T) and (F); a precise meta-language contains no appropriate substitution for '*P*'. Thus the formality of the semantics conceals the force of the argument. It does so at the cost of giving up the central task of genuine semantics: saying what utterances of the object-language mean. To suppose that such utterances constitute counterexamples to (B) is, after all, to suppose that they do mean something, so there can be no objection in principle to an enlargement of the meta-language that would enable it to say what they mean. Once the enlargement has been made, the argument against supposed counterexamples to (B) can be expressed. To produce a formal semantics in which a simulacrum of bivalence fails is to distract attention from the argument, not to answer it.

Many-valued logics, supervaluations and the like provide no objection to (T) and (F). The possibility of a genuine objection to (T) and (F) is not being dismissed out of hand, however. To be genuine, it would need to make it clear that truth and falsity themselves, not formal simulacra of them, were at issue. The objection might be that no property at all could satisfy (T) (or

(F)), or that none could satisfy both (T) (or (F)) and some other constraint that seemed at least as important to the notion of truth (or falsity). Now on every current semantic treatment of vague languages, both (T) and (F) are satisfiable. The only question is whether what satisfies (T) (or (F)) might thereby fail some other constraint on truth (or falsity). What could such a constraint be? It is hard to find an answer explicit in the writings of those who oppose bivalence. One may, however, suspect the presence of an assumed constraint in the background of much that is written. For (T) and (F) make attributions of truth and falsity as vague as the utterances to which they are attributed. If an utterance of 'TW is thin' is true just in case TW is thin, and it is indeed vague whether TW is thin, then it is also vague whether the utterance is true. Again, if the utterance is false just in case TW is not thin, and it is indeed vague whether TW is not thin, then it is also vague whether the utterance is false.[9] Thus a philosopher who assumed that attributions of truth and falsity must be precise might well object to (T) and (F).[10] But that assumption is just another form taken by the dream of a perfectly precise meta-language. As noted above, that dream prevents us from giving a genuine semantics for a vague language; in any case, previous chapters have found the dream to be unrealizable. We must accept that our attributions of truth and falsity, like just about all our other utterances, have some element of vagueness. The constraint that clashes with (T) and (F) is thus to be rejected on independent grounds. No plausible constraint on the notions of truth and falsity has been found to clash with (T) and (F). Those principles can therefore bear the weight of the argument against supposed counterexamples to bivalence.

The argument may nevertheless evoke some surprise. Can the denial of bivalence – a not uncommon move in philosophy – really be reduced to absurdity as easily as that? Some further elucidation may help.

(a) The argument does not purport to reduce every denial of bivalence to absurdity. It engages only with denials of particular instances of the general principle. It therefore does not touch intuitionism in mathematics, for example. Although intuitionists deny the general principle of bivalence, they are forbidden to give particular counterexamples, just because the inference from (1) to (4) is intuitionistically valid.[11] They sometimes refrain from asserting bivalence in a particular case, but they never deny it.[12] This does not undermine their denial of the general principle, for 'Not every utterance satisfies bivalence' does not intuitionistically entail 'Some utterance does not satisfy bivalence'. Vagueness is a different matter. Vague utterances are supposed to be obviously not bivalent in borderline cases, and the usual way of evoking the supposed sense of obviousness is by vivid descriptions of particular examples. It is obvious that not every vague utterance satisfies bivalence only if for some vague utterance it is obvious that *it* does not satisfy bivalence. Since, for any vague utterance, the supposition that it does not satisfy bivalence leads to a contradiction, it can hardly be obvious that not every vague utterance satisfies bivalence.[13]

(b) The argument does not immediately show that bivalence must be asserted for particular utterances, only that it must not be denied. Would it be consistent to accept the argument while adopting a neutral attitude to bivalence in particular cases? Such an attitude might even be made explicit by means of so-called *weak negation.* The idea is that the weak denial 'Ne *P*' of '*P*' is correct just in case the assertion '*P*' is incorrect. The assertion '*P*' is incorrect if either the ordinary, strong denial 'Not *P*' or a neutral attitude is correct. Bivalence might be weakly denied for particular utterances. In place of (1), one would have 'Ne: either *u* is true or *u* is false'. In place of (4), one would correspondingly have 'Ne *P* and ne not *P*', which is not a contradiction.

The meta-linguistic analogue of weak negation is weak falsity, explained in terms of 'ne' just as (F) explains ordinary falsity in terms of 'not'. If *u* says that *P*, then *u* is weakly false if and only if ne *P*.

Appeals to weak negation neglect higher-order vagueness. If first-order vagueness sometimes makes it incorrect to assert that an utterance is true or false, then second-order vagueness sometimes makes it incorrect to assert that an utterance is true or weakly false. If *u* is such an utterance, then (by the explanation of weak denial) ne: either *u* is true or *u* is weakly false. But this assumption regenerates the argument from (1) to (4) with 'ne' in place of 'not' (by the explanation of weak falsity). The conclusion, 'Ne *P* and ne ne *P*', is a contradiction if weak negation is any kind of negation at all.

Could the new argument be met by the postulation of a third kind of negation, weakly weak negation? That would begin a regress, which would involve as many kinds of negation as there are orders of vagueness.

Moreover, all these kinds of negation would have to be taken as primitive, for the strategy fails when weaker negations are defined in terms of stronger ones. If one tries to explain the truth-conditions of 'Ne *P*' in terms of 'not', one will hold that 'Ne *P*' is true if and only if '*P*' is not true (given, as in the cases at issue, that '*P*' says that something is the case); but then 'Ne *P*' will have the same truth-conditions as 'Not *P*', and its introduction will have achieved nothing. The alternative is a regress of primitive kinds of negation; it has no plausibility.

(c) The argument is not directed against supposed counterexamples to a strengthening of the principle of bivalence that is sometimes discussed under the same name. In present terminology, this is the principle that if *u* says that *P*, then *u* is either *definitely* true or *definitely* false. For then presumably *u* is definitely true if and only if definitely *P*, and *u* is definitely false if and only if definitely not *P*. If one tried to adapt the argument to use it against a supposed counterexample to the strengthened principle, one would reach, in place of (4), the conclusion that not definitely *P* and not definitely not *P*. That does not seem to be absurd.

The difficulty with the new principle is to understand what 'definitely' means. If it were just a rhetorical device for added emphasis, like underlining, then we could certainly understand it, but nothing would have been gained. The new principle would be equivalent to the old, and supposed counterexamples to it would be equally reducible to absurdity; 'not definitely *P* and not definitely not *P*' would after all be a contradiction. 'Definitely' must be intended to add something to the content of sentences in which it appears. It must add enough to prevent definite truth from satisfying the analogue of (T) or definite falsity from satisfying the analogue of (F), otherwise an analogue of the original argument could be used against supposed counterexamples to the new principle. Definite truth is supposed to be more than mere truth, and definite falsity more than mere falsity. But what more could it take for an utterance to be definitely true than just for it to be true? Given that it cannot be neither true nor false, how could it fail to be definitely true other than by failing to be true? Such questions are equally pressing with 'false' in place of 'true'. Again, 'TW is thin' is no doubt definitely true if and only if TW is definitely thin, but what is the difference between being thin and being definitely thin? Is it like the difference between being thin and being very thin? Can 'definitely' be explained in other terms, or are we supposed to grasp it as primitive?

Previous chapters examined the standard attempts to explain 'definitely' in formal terms; all were found to fail. Perhaps one can explain it by examples of borderline cases. One points at TW and says 'He is neither definitely thin nor definitely not thin'. Someone might indeed learn the word in such a way, and go on to say that an utterance of 'TW is thin' is neither definitely true nor definitely false. Now on the face of it, the claim that an utterance *u* is neither definitely true nor definitely false has no more to do with bivalence than has the claim that *u* is neither necessarily true nor necessarily false, or that *u* is neither obviously true nor obviously false.[14] Yet although the new principle is not supposed to collapse into the original principle of bivalence, it is supposed to be assumed by the classical framework within which bivalence seems compelling; why else discuss it? The question is whether, when 'definitely' is explained by examples, it acquires a sense on which the claim that an utterance is neither definitely true nor definitely false has any bearing on semantics. For nothing has been said to rule out the possibility that 'definitely' has acquired an *epistemic* sense, something like 'knowably'. If further stipulations are made in an attempt to rule out that possibility, it is not obvious that 'definitely' retains any coherent sense. Until it is properly explained in such a way as to make clear a conflict between principles of classical semantics and logic and the supposition that an utterance says that something is the case without being either definitely true or definitely false, there is no need to argue against that supposition.

The present point is closely related to that under (b), for the so-called weak negation of '*P*', 'Ne *P*', might be explained as 'Not definitely *P*', the strong negation of 'Definitely *P*'. In both cases, the difficulty is to find a coherent non-epistemic sense for the technical term.

(d) The argument does not purport to show that a vague utterance cannot be neither true nor false, just that it cannot both say that something is the case and be neither true nor false. If a vague utterance fails to say that anything is the case, it is no counterexample to (B).

However, we have good reason not to attribute such a failure to many vague utterances, even in borderline cases.

To suppose that no vague utterance says anything is to adopt a form of global nihilism, whose untenability was noted in Section 6.2. The view that needs to be considered here is the more limited one that vague utterances in borderline cases fail to say anything. On this view, whether someone who utters 'TW is thin' has said something depends on the state of my body at the time. Now suppose that I have a twin, TW2, whose dimensions are the same as mine. It seems to follow that an utterance of the material conditional 'If TW is thin, then TW2 is thin' is true. Moreover, its truth seems to depend, not on what our shared dimensions are, but just on the fact that we share them. In particular, it seems to be true even if we are both borderline cases of thinness. Since it is true, it must have said something. The conditional says something only because its antecedent and consequent also do. They are not assertions, but it has already been noted that not all sayings are assertions. If the antecedent or consequent lacked content, so would the conditional as a whole. Thus 'TW is thin' has content; it says that something is the case.

Even if I am a borderline case, we know what an utterance of 'TW is thin' says because we understand its subject term 'TW', its predicate 'is thin' and the significance of putting them together in that way, and we are aware of any relevant contextual factors (for example, that the set of human beings is an appropriate comparison class). Moreover, we can envisage circumstances in which the utterance would have been clearly true (because TW would have been clearly thin) while saying just what it actually says, and other circumstances in which it would have been clearly false (because TW would have been clearly not thin) while again saying just what it actually says.

Borderline cases may be contrasted with cases of reference failure for singular terms. The phrase 'this dagger' may fail to single anything out when used by someone under a hallucination. Arguably, utterances such as 'This dagger is sharp' in which the phrase is used fail to say anything in this context, and so are neither true nor false. That includes complex utterances; even the biconditional '"This dagger is sharp" is true if and only if this dagger is sharp' fails to say anything, for it uses the phrase 'this dagger' on its right-hand side. Although we can envisage circumstances in which an utterance of 'This dagger is sharp' would have been clearly true or clearly false, they are not ones in which it says what the utterance actually says, for it actually says nothing. In order to understand an utterance of 'This dagger is sharp', one must know which thing the demonstrative singles out in the given context; when reference fails, there is no such fact to be known. There is no parallel obstacle to understanding vague utterances in borderline cases.

(e) The argument does not have untoward consequences when applied to semantic paradoxes such as the Liar. Suppose that an utterance *u* of the sentence 'This utterance is not true' says that *u* is not true. By (T), it follows that *u* is true if and only if *u* is not true. That is impossible, so *u* does not say that *u* is not true. Either *u* does not say anything, or it says something else. Both alternatives can be explored; they may not be as far apart as they seem. For example, we can begin with a certain notion of saying, 'say_0', and use (T) and (F) to define corresponding notions of truth and falsity, 'true_0' and 'false_0'. Perhaps truth_0 and falsity_0 cannot figure in what is said_0, but we can develop an extended notion of saying, 'say_1', such that they can figure in what is said_1. We then use (T) and (F) again to define extended notions of truth and falsity, 'true_1' and 'false_1', and so on. If 'true' in *u* is taken to mean 'true_i', then *u* does not say_i anything, and hence is not true_i. It does say_{i+1} that *u* is not true_i, but all that follows is that *u* is true_{i+1} if and only if *u* is not true*i* – so *u* is true_{i+1}. No contradiction results. More sophisticated approaches are possible too. At each level *i*, a consistent theory results when one subscripts 'true', 'false' and 'say' in (B), (T) and (F) with '*i*'. For vague utterances such as 'TW is thin' in borderline cases, only level 0 is relevant; thus 'true', 'false' and 'say' in the discussion prior to (e) may be read as 'true_0', 'false_0' and 'say_0'. Although these remarks are no substitute for a fully worked out treatment of the semantic paradoxes, they do indicate that the present approach to vagueness is consistent with a number of such approaches.[15]

(f) The argument does not deny ambiguity. If *u* both says that *P* and says that *Q*, then by (T) *u* is both true if and only if *P* and true if and only if *Q*, so *P* if and only if *Q*. Nevertheless, ambiguity occurs. It does so primarily at the level of linguistic types. The string of words ‘Flying planes can be dangerous’ is ambiguous because syntactically different sentences consist of those words. The string of letters ‘bank’ is ambiguous because semantically different words consist of those letters. Such ambiguity has nothing to do with vagueness. An utterance of a sentence is not an utterance of any other sentence, even if the hearer is not certain which sentence has been uttered. An utterance of a word is not an utterance of any other word, even if the hearer is not certain which word has been uttered. Utterances are to be individuated by what they are utterances of. Very occasionally, a speaker may exploit ambiguities in linguistic types to say each of two things simultaneously; in such a case, there are two simultaneous utterances. But if the speaker has simply not decided which of two things to say, and makes noises appropriate to both, then nothing has been said to be the case. In no case does ambiguity threaten (T) or (F). My present utterance of ‘TW is thin’ says that TW is thin; it does not say something else.

The upshot of the argument is that it is incoherent to suppose that vague utterances in borderline cases both say something and fail to be either true or false. It is coherent to suppose them to be neither true nor false only at the cost of treating them as though they said nothing. Formal semantics pays the cost by affecting to use a precise meta-language in which one cannot say what utterances in the vague object-language mean. Since we are inescapably committed to the practice of using vague language, we cannot permanently afford that price. Rapid alternation between perspectives inside and outside the practice can obscure the issue; it does not answer the argument.

7.3 OMNISCIENT SPEAKERS

The idea that borderline cases present counterexamples to the principle of bivalence is a manifestation of the idea that in such cases nothing is hidden. Since the vague utterance is neither obviously true nor obviously false, it is held not to be either true or false. Although individual speakers may happen to be ignorant of various facts, their ignorance is conceived as quite inessential to the nature of the borderline case. If it were essential, vagueness would be an epistemic phenomenon.

On the view that nothing is hidden, it should be harmless to imagine omniscient speakers, ignorant of nothing relevant to the borderline case. Such a hypothesis of itself carries no commitment to bivalence or excluded middle. It is supposed, for example, that if TW is thin then the omniscient speaker knows that TW is thin, and that if TW is not thin then the omniscient speaker knows that TW is not thin. It is no part of the hypothesis that TW is either thin or not thin. Thus it is no part of the hypothesis that either the omniscient speaker knows that TW is thin or the omniscient speaker knows that TW is not thin.[16] On the contrary, it is supposed that if, for example, TW is neither definitely thin nor definitely not thin, then the omniscient speaker knows that TW is neither definitely thin nor definitely not thin. Although reason emerged in Section 7.2 to doubt that 'definitely' can be given an appropriate non-epistemic sense, those doubts may temporarily be bracketed. The hypothesis of an omniscient speaker gives us a different way to appreciate the difficulty of thinking through a non-epistemic view of vagueness.

Accompanied by an omniscient speaker of English, you remove grain after grain from a heap. After each removal you ask 'Is there still a heap?'. The omniscient speaker is not required to answer 'Yes' or 'No'; she can say 'That is indeterminate' or 'To degree 0.917' or 'You are asking the wrong question' if she likes. If there is nothing to say, she will remain silent. She will not say 'I don't know', nor will she hesitate, unsure of the best answer. You can ask her not to mumble. She is determined to be as cooperative and as relevantly informative as she can.

After the first few removals, what is left remains quite clearly a heap. The first few times you ask 'Is there still a heap?', the omniscient speaker answers 'Yes'; she may say other things too, but whatever else she says, on those first few occasions she says 'Yes'. She is not an obscurantist. After sufficiently many of the grains have been removed, she does not say 'Yes'. She is not a liar. For some number *n*, she says 'Yes' after each of the first *n* removals, but not after *n* + 1. You do not know the value of '*n*' in advance. Yet on the view in question, her stopping point cannot represent a hidden line between truth and something less than truth. How can that be?

You repeat the experiment with other omniscient speakers, starting with exactly similar arrangements of grains in exactly similar contexts, then removing the grains in exactly the same way. The trials are independent;

there is no collusion between the omniscient speakers. If they all stop at the same point, it evidently does mark some sort of previously hidden boundary, although it may be a delicate matter to say just what it is a boundary between. The view in question must therefore hold that different omniscient speakers would stop at different points. They are conceived as having some sort of discretion. In some cases it is mandatory to apply the term 'heap'; in others it is permissible but not mandatory. In the latter cases, some but not all omniscient speakers answer 'Yes'.

You can instruct the omniscient speakers how to use their discretion. For example, you can instruct them to use it conservatively, so that they answer 'Yes' to as few questions as is permissible. They will still answer 'Yes' to the first few questions, for the same reason as before: at that stage, it is mandatory to apply the term 'heap'. Now if two omniscient speakers stop answering 'Yes' at different points, both having been instructed to be conservative, the one who stops later has disobeyed your instructions, for the actions of the other show that the former could have used her discretion to answer 'Yes' to fewer questions than she actually did. But the omniscient speakers are cooperative. They will obey your instructions if they can. Since your instructions are capable of being obeyed, they will be obeyed. It is not as though, however many times they said 'Yes', they could have said it fewer times, for the sorites series is finite. They will not disobey your instructions. Thus, if all are instructed to be conservative, all will stop at the same point. You do not know in advance where it will come. It marks some sort of previously hidden boundary, although it may be a delicate matter to say just what it is a boundary between.

The same problem arises if you instruct the omniscient speakers to use their discretion liberally, so that they answer 'Yes' to as many questions as is permissible. They will still cease to answer 'Yes' after sufficiently many questions, for the same reason as before: by that stage, it is forbidden to apply the term 'heap'. Now if two omniscient speakers stop answering 'Yes' at different points, both having been instructed to be liberal, the one who stops sooner has disobeyed your instructions, for the actions of the other show that the former could have used her discretion to answer 'Yes' to more questions than she actually did. But, as before, the omniscient speakers will not disobey your instructions. Thus, if all are instructed to be liberal, then all will stop at the same point. You do not know in advance where it will come. It marks another previously hidden boundary. It may or may not be the boundary between truth and falsity. What matters is that it is of semantic significance, and was hidden from ordinary speakers.

Some may by now feel tempted to repudiate the very possibility of omniscient speakers. To do so is to endorse a strong form of the view that vagueness is an epistemic phenomenon, for it is to treat ignorance as an essential feature of borderline cases.

We have no idea how to conceive borderline cases in such a way that nothing in them lies hidden from ordinary speakers. Once hidden lines are admitted, why should a line between truth and falsity not be one of them? After all, Section 7.2 found the supposition of intermediate cases to be incoherent. Moreover, the failure of non-classical logics to mesh with a satisfactory account of vagueness gives us reason tentatively to return to classical logic. By excluded middle, either TW is thin or TW is not thin. Since an utterance of 'TW is thin' says that TW is thin, it follows by (T), (F) and more classical logic that the utterance is either true or false. Bivalence holds in borderline cases. It is just that we are in no position to find out which truth-value the vague utterance has.

No coherent alternative to the epistemic view has been found. Yet the view has often been assumed to have philosophically intolerable consequences. Sections 7.4–7.6 examine some of its supposed consequences, and show that they are either not intolerable or not consequences. More positive developments will be left to the next chapter.

7.4 THE SUPERVENIENCE OF VAGUENESS

Some accounts of vagueness rule out the epistemic view by definition. A term is said to be vague only if it can have borderline cases, and a case is said to be borderline only if our inability to decide it does not depend on ignorance. The epistemic view is therefore held to imply that vagueness does not really occur.

Tendentious definitions achieve little. It would need to be shown that the terms we ordinarily think of as vague are so in the defined sense of 'vague'. In particular, it would need to be shown that the cases we ordinarily think of as borderline are so in the defined sense of 'borderline'. To make those assumptions without showing them to be correct is merely to beg the question against the epistemic view. For purposes of communication, it is more fruitful to define ‘vague’ and ‘borderline’ by giving examples. We can all agree that ‘thin’ is a vague term, and TW a borderline case for thinness, and then argue on that basis about the underlying nature of the examples. The epistemic view is that ignorance is the real essence of the phenomenon ostensively identified as vagueness.

The epistemic theorist is sometimes asked: of what fact could we be ignorant? The answer is obvious. In the present case, we are ignorant either of the fact that TW is thin or of the fact that TW is not thin. Our ignorance naturally prevents us from knowing which of those is indeed a fact, and therefore from knowing which fact we are ignorant of. That the facts are specified by use of the word ‘thin’ is just what one would expect in the light of Section 7.2. There is no general requirement that vague words be definable in other terms.

A further question is sometimes asked: what kind of fact is the fact that TW is thin? The question is a bad one, for a reason unconnected with vagueness. We do not have a proper taxonomy of facts, not even of precise ones. When a taxonomy is provided, it will be time to say what kind of fact a vague fact is.

Those wholly predictable opening moves against the epistemic view mismanage a deeper objection. It can be made using the idea that vague facts *supervene* on precise ones. If two possible situations are identical in all precisely specified respects, then they are identical in all vaguely specified respects too. For example, if x and y have exactly the same physical measurements, then x is thin if and only if y is thin. Strictly speaking, whether x and y are thin may depend on the physical measurements of the relevant comparison class as well as on those of x and y themselves. For simplicity, the comparison class will be held constant, but nothing in the conclusions to be drawn depends on that. A more general formulation of the supervenience thesis is:

(*) If x has exactly the same physical measurements in a possible situation s as y has in a possible situation t, then x is thin in s if and only if y is thin in t.

The objection to the epistemic view can now be formulated. Let my exact physical measurements be m. According to the epistemic view, I am either thin or not thin. By (*), if I am thin then necessarily everyone with physical measurements m is thin. Similarly, if I am not thin then necessarily no one with physical measurements m is thin. Thus either being thin is a necessary consequence of having physical measurements m, or not being thin is. Suppose that I find out, as I can, what my physical measurements are. I would then seem to be in a position either to deduce that I am thin or to deduce that I am not. But it has already been conceded that no amount of measuring will enable me to decide whether I am thin.

The basis of this objection to the epistemic view is not that one can know all the relevant facts in a case ordinarily classified as borderline but that one can know a set of facts on which all the relevant facts supervene, without being able to decide the case. Unlike the former claim, the latter does not beg the question against the epistemic view. The epistemic theorist has as much reason as everyone else to accept supervenience claims like (*). However, the objection commits a subtler fallacy.

The kind of possibility and necessity at issue in supervenience claims like (*) is metaphysical. There *could not be* two situations differing vaguely but not precisely. Suppose that I am in fact thin. By (*), it is metaphysically necessary that everyone with physical measurements m is thin. If I know that I have physical measurements m, then, in order to know by deduction that I am thin, I must first *know* that everyone with physical measurements m is thin. The plausibility of the objection to the epistemic view thus depends on something like the inference that, since the supervenience generalizations are metaphysically necessary, they can be known *a priori*. The inference from metaphysical necessity to *a priori* knowability may be a tempting one: but, as Kripke has emphasized, it is fallacious. Indeed, metaphysical necessities cannot be assumed knowable in any way at all.

Consider mathematical truths. They are all metaphysically necessary; there is no presumption that they are all knowable. A standard example is

Goldbach's Conjecture, which says that every even number greater than 2 is the sum of two prime numbers. The Conjecture has been neither proved nor refuted; for all we know, no humanly intelligible method of argument can decide it one way or the other. But for any particular even number greater than 2, one can in principle decide within a finite time whether it is the sum of two primes; if it is not, then the Conjecture is refutable. Contrapositively, if the Conjecture is undecidable, then there is no counterexample to it, so it is true. For all we know, Goldbach's Conjecture is a humanly unknowable, metaphysically necessary truth. Vague truths can be in that position too. It is integral to the epistemic view that metaphysically necessary claims like 'Everyone with physical measurements *m* is thin' can be as unknowable as physically contingent ones like 'TW is thin'.

One should not be surprised that the known supervenience of *A*-facts on *B*-facts does not provide a route from knowledge of *B*-facts to knowledge of *A*-facts. A more familiar case is the postulated supervenience of mental facts on physical facts. Suppose, for illustration, that pain is known to supervene on the state of the brain. Then if *b* is a maximally specific brain state (specified in physical terms) of someone in pain, it is metaphysically necessary that everyone in brain state *b* is in pain. However, there is no presumption that one could have found out that someone was in pain simply by measuring his brain state and invoking supervenience. 'Everyone in brain state *b* is in pain' certainly cannot be known *a priori.* Perhaps one can know it *a posteriori*, because one can find out that someone is in pain by asking him, and then combine that knowledge with knowledge of his brain state and of the supervenience of mental states on brain states. 'Everyone with physical measurements *m* is thin' cannot be known *a posteriori* in a parallel way, for no route to independent knowledge of someone with physical measurements *m* that he is thin corresponds to asking someone whether he is in pain.

The epistemic view of vagueness is consistent with the supervenience of vague facts on precise ones. The next section considers a different source of objections to the epistemic view, and makes another application of the concept of supervenience.

7.5 MEANING AND USE

A common complaint against the epistemic view of vagueness is that it severs a necessary connection between meaning and use. Words mean what they do because we use them as we do; to postulate a fact of the matter in borderline cases is (it is charged) to suppose, incoherently, that the meanings of our words draw lines where our use of them does not. The point is perhaps better put at the level of complete speech acts, in terms of sentences rather than single words. If the meaning of a declarative sentence may provisionally be identified with its truth-conditions, and its use with our dispositions to assent to and dissent from it in varying circumstances, then the complaint is that the epistemic view of vagueness sets truth-conditions floating unacceptably free of our dispositions to assent and dissent.

On the epistemic view, it is not very plausible to identify the meaning of a declarative sentence with its truth-conditions. Arguably, if we make stipulations about the future use of a hitherto vague term, we change its meaning, for anyone ignorant of the stipulations fails to understand it in its new sense. For the epistemic theorist, the stipulations might by chance leave its extension intact with respect to all possible worlds. Sentences containing the term would have changed in meaning but not in truth-conditions. However, such further aspects of meaning are not what concern the objector, for their dependence on use is comparatively plain. The complaint concerns truth-conditions.

The bare complaint is too general to be convincing. Environmental factors beyond our dispositions to assent and dissent may play a role in fixing truth-conditions. Suppose that a sub-atomic particle of a new kind is detected in a laboratory; particles of that kind are called ‘X-particles’. We can make sense of the further supposition that there are also particles of a different and much rarer kind, Y-particles, none of which has yet been detected, whose underlying nature is quite different from that of X-particles, and more similar to that of some particles already distinguished from X-particles, these similarities and differences being humanly undetectable (man is not the measure of all things). If scientists did detect a Y-particle, they would incorrectly classify it as an X-particle. Our dispositions to assent to and dissent from the sentence ‘An X-particle is present' do not discriminate between X-particles and Y-particles, but the truth-conditions of the sentence are sensitive to the difference.

What the objector needs to emphasize is that there is no sharp natural division for the truth-conditions of 'He is thin' to follow that might correspond to the sharp natural division between X-particles and Y-particles followed by the truth-conditions of 'An X-particle is present'. The thin things do not form a natural kind. The thought is that, if nature does not draw a line for us, then a line is drawn only if we draw it ourselves, by our use. So (it is held) there is no line, for our use leaves not a line but a smear.

Before one allows oneself to be persuaded by the revised complaint, one should probe its conception of drawing a line. On the face of it, 'drawing' is just a metaphor for 'determining'. To say that use determines meaning is just to say that meaning *supervenes* on use. That is: same use entails same meaning, so no difference in meaning without a difference in use. More formally:

(#) If an expression *e* is used in a possible situation *s* in the same way as an expression *f* is used in a possible situation *t*, then *e* has the same meaning in *s* as *f* has in *t*.

There are various problems with (#), such as its neglect of the environment as a constitutive factor in meaning (as noted above) and its crude notion of two expressions being 'used in the same way'. However, it will be assumed for the sake of argument that some refinement of (#) is correct. For the epistemic view of vagueness is quite consistent with (#) and its refinements. Although the view does not permit simple-minded reductions of meaning to use, it in no way entails the possibility of a difference in meaning without any corresponding difference in use. Had 'TW is thin' had different truth-conditions, our dispositions to assent to and dissent from it would have been different too.

Although meaning may supervene on use, there is no algorithm for calculating the former from the latter. Truth-conditions cannot be reduced to the statistics of assent and dissent. In particular, the line between truth and falsity is not to be equated with the line between unanimous and less than unanimous assent, or with the line between majority assent and its absence. The study of vagueness has regrettably served as the last refuge for the consensus theory of truth, a theory no more tenable for vague utterances than it is for precise ones. We can certainly be wrong about whether someone is thin, for we can be wrong both about the person's shape and size and about normal shapes and sizes in the relevant comparison class. These errors may be systematic; some people may characteristically look thinner or less thin than they really are, and there may be characteristic misconceptions about the prevalence of various shapes and sizes. Appeal might be made to dispositions to assent and dissent in epistemically ideal situations or given perfect information, but that is merely to swamp normal speakers of English with more measurements and statistics than they can handle. Perhaps the dispositions to assent and dissent of an epistemically ideal speaker of English would be an infallible guide to thinness, but then such a speaker might know the truth-value of 'TW is thin'. The ordinary basis for attributions of 'thin' is perceptual; such a basis is inherently fallible.

Every known recipe for extracting meaning from use breaks down even in cases to which vagueness is irrelevant. The inability of the epistemic view of vagueness to provide a successful recipe is an inability it shares with all its rivals. Nor is there any reason to suppose that such a recipe must exist.

It may still be thought that the epistemic view makes special difficulties for an account of the relation between meaning and use. Suppose, for simplicity, that in normal perceptual conditions any competent speaker of English refuses to classify me as thin and refuses to classify me as not thin. How could the truth or falsity of 'TW is thin' possibly supervene on that pattern of use?

Two distinguishable worries may lie behind the question. One is that an assignment of truth or falsity to 'TW is thin' would violate native speakers' intuitions. The other is that since the situation is perfectly symmetrical between assertion and denial at the level of use, any assignment of truth or falsity would involve an implausible breaking of symmetry in an arbitrarily chosen direction. The two worries will be discussed in turn.

Native speakers may well feel that it would be wrong of them to assert or deny 'TW is thin'. Such a feeling might simply be caused by their knowledge that they do not know whether TW is thin, so that they are in no position to make either claim. To rule out this hypothesis, the objector must suppose that native speakers will say something like 'There is no fact of the matter'. In effect, the objector hopes that native speakers will spontaneously come out with the objector's own theory of vagueness. The objector passes the buck to the native speaker. But even if the latter does say what the former wants, that will not show it to be correct. After all, if every native speaker says instead 'There is a fact of the matter, but I don't know what it is', no one would suppose that to establish the correctness of the epistemic view, and the case is the same in reverse. The dispute is a philosophical one, on which the views of native speakers are not authoritative. In fact, it was only with some difficulty that philosophers learnt to distinguish between epistemic and non-epistemic conceptions of vagueness. Why suppose that native speakers have the distinction at their fingertips? Their refusal to assert or deny 'TW is thin' must be taken seriously, but the gloss they put on their refusal is not a manifestation of their knowledge of the language. A bad philosophical theory does not become any better by being put into the mouths of native speakers.

The other worry concerns the apparent symmetry of the situation. However, the concepts of truth and falsity are not symmetrical. The asymmetry is visible in the fundamental principles governing them, for (F) is essentially more complex than (T), by its use of negation. The epistemic theorist can see things this way: if everything is symmetrical at the level of use, then the utterance fails to be true, and is false in virtue of that failure (if it says that something is the case). In that sense, truth is primary. At the level of truth and falsity, there is no symmetry to break.

In some cases we may be able to say something more specific about the connection between meaning and use. Suppose, for example, that we have a (fallible) mechanism for recognizing the property of thinness. Although everything has or lacks the property, the reliability of the mechanism depends on its giving neither a positive nor a negative response in marginal cases. The cost of having the mechanism answer in such cases would be many wrong answers. It is safer to have a mechanism that often gives no answer than one that often gives the wrong answer. Nevertheless, given the mechanism, we can ask 'Which properties does this mechanism best register?'.[17]Perhaps thinness is the conjunction of all those properties that the mechanism detects as well as it detects any property. If the mechanism registers one property better than all others, then it is thinness.

A subject whose primary access to a property is through a recognitional mechanism may not be helped to detect it by extra information of a kind that cannot be processed by that mechanism, even if the new information is in fact a reliable indicator of the presence of the property, for the subject may be unable to read the signs. My exact measurements may in fact be a sufficient condition for thinness, and knowledge of the former still not enable us to derive knowledge of the latter. For all that, thinness may be the property best registered by our perceptual recognitional capacity for thinness, or the conjunction of several such properties.

Such speculations should not mislead one into supposing that a causal theory of reference is essential to an epistemic view of vagueness. They illustrate only one way in which our use of a vague expression might determine a sharp property. Even for 'thin', the truth is no doubt infinitely more subtle. The epistemic theory of vagueness makes the connection between meaning and use no harder to understand than it already is. At worst, there may be no account to be had, beyond a few vague salutary remarks. Meaning may supervene on use in an unsurveyably chaotic way.

7.6 UNDERSTANDING

The charge against the epistemic view of vagueness may be revised. If the view does not force what we mean to transcend what we *do*, then perhaps it forces what we mean to transcend what we *know*. The clearest form of the charge is simply that the view prevents us from knowing what we mean.

The basis for the charge is less clear. Section 7.5 allowed the meaning of an utterance to supervene on the use of expressions. That account can be extended from language to thought.[18] On the extended account, the content of a thought-occurrence, like that of an utterance, supervenes on facts specified without reference to content. An utterance of 'TW is thin' and a thought that TW is thin supervene on the same facts. So if I think that my utterance of 'TW is thin' says that TW is thin, what determines what the utterance says is the same as what determines what I think the utterance says. No gap need open between what one means and what one thinks one means when the same facts determine both. I know that my utterance of 'TW is thin' says that, and is therefore true if and only if, TW is thin.[19]

Our methods for deciding questions of thinness may fail when applied to 'TW is thin', just as our methods for deciding mathematical questions may fail when applied to Goldbach's Conjecture. It does not follow that we do not know the truth-conditions of 'TW is thin', or of Goldbach's Conjecture. Only an extreme verificationist would ground actual knowledge of the truth-conditions of an utterance in possible knowledge of its truth-value, and attempts to realize semantic theories for natural languages on such lines remain hopelessly programmatic.

Some non-verificationists may reply that the epistemic view makes us ignorant of the sense of a vague term, not just of its reference. Of course we do not know where all the thin things are in physical space; the charge is that we should not even know where they all are in conceptual space. We should be using a term that does in fact determine a line in conceptual space without being able to find that line. We should understand the term partially, as one partially understands a word one has heard used once or twice. But in the latter case the word's meaning is backed by other speakers' full understanding, whereas no one is allowed full understanding of the vague term. The objection to the epistemic view is that it attributes partial understanding to the speech community as a whole. It is not entitled to say that we know what we mean, for it is not entitled to assume that we know what we mean by '"TW is thin" says that TW is thin'. It attributes to the community incomplete knowledge of a complete meaning; would it not be more reasonable to attribute complete knowledge of an incomplete meaning?

The objection is based on the Fregean model of the sense of a term as a region in conceptual space: to grasp a sense is to know where its boundary runs. Individual points in the space are located by means of precise descriptions such as 'having exact physical measurements *m*'. Thus the demand that one know which points are in the region marked off by a vague term such as 'thin' is simply the demand that one know truths such as 'Anyone having exact physical measurements *m* is thin' or 'No one having exact physical measurements *m* is thin'. The unreasonableness of that demand was noted in Section 7.4; the metaphysical necessity of such truths does not justify the demand to know them. The metaphor of conceptual space adds no force to the demand. Rather, its illicit function is to collapse distinctions between concepts whose equivalence is metaphysically necessary but not *a priori*, by identifying them with the same region in conceptual space. It is well known that when a proposition is identified with the set of possible worlds at which it is true, a region in the space of possible worlds, cognitively significant distinctions are lost. Exactly the same happens when the objection identifies a sense with a region in conceptual space, conceived as above.

On the epistemic view, our understanding of vague terms is not partial. The measure of full understanding is not possession of a complete set of metaphysically necessary truths but complete induction into a practice. When I have heard a word used only once or twice, my understanding is partial because there is more to the community's use of it than I yet know. I have not got fully inside the practice; I am to some extent still an outsider. Indeed, I probably think of myself as an outsider, knowing that there is more to the practice than I yet know; my use of the term will be correspondingly tentative and deferential. It does not follow that if we had all understood the term in the vague way I do, although without tentativeness or deference, then all our understandings would have been partial. In that counterfactual situation, we should all have been insiders. To know what a word means is to be completely inducted into a practice that does in fact determine a meaning.[20]

To be inducted into a practice, it is not necessary to acquire dispositions that exactly match those of other insiders. Of two people who understand the word 'thin', one may be willing to apply it in a slightly wider range of cases than the other. Rough matching is enough. Perhaps no two speakers of English match exactly in their dispositions to use 'thin'. It does not follow that no two speakers of English mean exactly the same by 'thin'. For what individual speakers mean by a word can be parasitic on its meaning in a public language. The dispositions of all practitioners collectively determine a sense that is available to each.

When an individual is inducted into a practice, other individuals must already be practitioners. The growth of understanding therefore seems to require the existence of a community. Is this a quick argument against the possibility of a private language? Hardly. Although an isolated individual cannot be inducted into a practice, no reason has emerged why one should not initiate a practice, and engage in it, alone. For all that has been said, such a practice might amount to the use of a language. In using the language, the individual would understand it only in the sense of creating it. The question of understanding in any further sense would arise only if other individuals sought to learn the language. But this is not to say that no substantive distinction between sense and nonsense applies to the language of the isolated individual. For to characterize what the individual has as a language is already to imply that it is not nonsense. Although initiating a practice is not like being inducted into it, the achievement is no less substantive. This is not the place to discuss in general what it takes to use a language, or whether in particular an isolated individual can use one. For present purposes, it is enough that the epistemic view of vagueness opens no objectionable gap between using a language and understanding it.

7.7 DECIDABLE CASES

Some objections to an epistemic account of a vague utterance in a borderline case are consistent with the assumption that the utterance is true or false. The objector may deny that any unknowable truth is involved on the grounds that the truth in question is knowable, if unknown. Before such objections are discussed, the claimed unknowability must be examined more closely.

Suppose that persons with exact physical measurements *m* are borderline cases for 'thin'. The epistemic theorist has no special reason to deny that a being with cognitive powers greater than any we can imagine could know of someone with exact physical measurements *m* whether he is thin. Who knows what such a being might know? On the epistemic view, vague utterances in borderline cases are true or false and we humans have no idea how to find out which. It is quite consistent with this view that what is a borderline case for us is not a borderline case for creatures with cognitive powers greater than any we can imagine. Equally, the epistemic theorist has no special reason to assert that such a being could know of someone with exact physical measurements *m* whether he is thin. The cognitive capacities of creatures outside the speech community are simply not to the point.

The epistemic theorist will certainly allow that a case may appear to be borderline without really being so. Trivially, a thin person may disguise himself to look like a borderline case of thinness. Less trivially, an astonishingly plausible case has been made that we can indeed work out the least number of grains needed for a heap: it is four. The suggestion is that 'heap' is after all a natural kind term, and that we can discover the physics that constitutes the real essence of the kind. To simplify: for the presence of a heap it is necessary and sufficient that at least one grain should stably rest on other grains. The least number of grains to make such a configuration possible is four.[21] Were we to accept the argument, we should presumably say that 'heap' was not as vague as had been supposed, because the cases that appeared to be on its borderline were not really so. That would show that 'heap' was not a good example; it would not show that anything was wrong with our account of vagueness. We could use 'large heap' as our example in place of 'heap'.

It has been suggested that, strictly speaking, 'bald' is true only of those people who have no hair on their scalps at all. On this supposition, one can still explain why we use the term when speaking of people with some but very little hair on their scalps. Such loose talk is false but informative; a very specific statement that is false but close to being true is often more useful than its true but very unspecific negation.[22] Remarks like those in the previous case apply, for we could use 'baldish' as our example in place of 'bald'.

In artificially simple cases, the considerations of Section 7.5 may help us to decide what might otherwise appear to be borderline cases. Suppose, for example, that we have a word 'dommal', of which our use turns out to be constrained by just two principles: that every dog is a dommal, and that every dommal is a mammal. Is a cat a dommal? At the level of use, we may refuse with equal firmness to answer 'Yes' or 'No'. A cat may appear to be a borderline case for 'dommal'. But if one accepts the asymmetry between truth and falsity postulated in Section 7.5, one can argue that since our use does not do enough to make an utterance of 'A cat is a dommal' true, it *thereby* does enough to make it false (given that it says something). If the argument is sound, it enables one to know that a cat is not a dommal. It is after all not really a borderline case. One's knowledge that a cat is not a dommal depends on one's knowledge of the symmetry at the level of use. In more realistic examples, knowledge of the latter kind is not to be expected.

The epistemic view comes under threat only if such arguments can be generalized to all borderline cases, for then it could no longer claim to be dealing with the genuine phenomenon of vagueness. There is no prospect of generalizing the arguments above, for the features they exploit are highly specific to the cases at issue. A more systematic strategy would be needed.

One strategy may be quickly dismissed: that of stipulation. We might indeed stipulate that everyone with physical measurements *m* is to count as 'thin'. On just about any view, such a stipulation changes the meaning of 'thin' in those contexts in which it has authority. One needs to know more than was previously necessary to understand the word. There is also likely to be a slight change in its extension. Such a change in meaning can be beneficial, even on the epistemic view, for the exact location of a cut-off point is sometimes less important for the purpose in hand than the fact that we know its location. What the stipulation does not do is to tell us anything about the truth or falsity of utterances already made using 'thin' in its old sense. We cannot stipulate truth-values for them any more than we can stipulate birth-dates for the emperors of Rome. For certain purposes we may choose to treat them as true, or as false, but that does not make them true, or make them false. Stipulations do not answer old questions; they enable us to ask new and sometimes better ones.

A different strategy appeals to the context dependence of vague expressions. The extension of 'thin', for example, is supposed to vary without need of stipulation from context to context, depending on the purpose in hand, the salient comparison classes, previous uses of it in the conversation, and so on. Something constitutes a borderline case if the term applies to it in some contexts and not in others. The suggestion is that, once one fixes the context, one can find out whether the term applies. Sorites paradoxes may be held to exploit the change in context brought about by each concession to the argument. To answer 'Yes' to the question 'Is x_i red?' is to add something to the context of the next question 'Is x_{i+1} red?' that liberalizes the condition for 'red' to apply and thus makes it harder to escape answering 'Yes' to it too.[23]

The objection to the new strategy is that it expects far too much of the context. If one person accuses another of having left a heap of sand on the floor, and is told that it was not a heap on the grounds that there were only a few grains, that context need not enable them to resolve their difference. In a conversation about the social structure of Italy, someone may ask 'How many Italian farmers are very rich?'; the context does not enable one to read the answer to that question off detailed statistics about wealth and income in Italy. Examples could be multiplied without end. If classical logic and semantics are to be applied in such cases, the epistemic account is needed.

Vagueness and context dependence are separate phenomena. The word 'now' is not vague merely because its reference depends on the time of utterance. Conversely, vagueness remains even when the context is fixed. In principle, a vague word might exhibit no context dependence whatsoever. In practice, the lack of natural boundaries for vague words makes context dependence hard to avoid, but that is an empirical correlation, not an *a priori* law.

If one is plausibly to maintain that vague utterances satisfy classical logic and semantics, one must appeal to ignorance. However, little has been said so far to explain *why* we should be ignorant of the truth-values of vague utterances. That is the task of the next chapter. Such ignorance, it will be argued, is just what independently justified epistemic principles would lead one to expect.

Chapter 8

Inexact knowledge

8.1 THE EXPLANATORY TASK

Ignorance is a natural human state. The limits on what creatures of other species can know are not the limits of the universe; our species is unlikely to be very different. Our knowledge stands in more need of explanation than does our ignorance. Self-knowledge is no exception. We do not expect creatures of other species to be capable of knowing everything about themselves; we should not expect to be capable of knowing everything about ourselves. The same applies to our knowledge of our own creations. In some sense we create our language, but it does not follow that it is in every respect open to our gaze. Why should the boundaries of our terms not be invisible to us?

Such general considerations do not lift the whole burden of proof from the epistemic view of ignorance. For most vague terms, there is knowledge to be explained as well as ignorance. Although we cannot know whether the term applies in a borderline case, we can know whether it applies in many cases that are not borderline. The epistemic view may reasonably be expected to explain why the methods successfully used to acquire knowledge in the latter cases fail in the former. This chapter provides such an explanation.

Ignorance in borderline cases will be assimilated to a much wider phenomenon, a kind of ignorance that occurs wherever our knowledge is *inexact*. The notion of inexact knowledge, like that of vagueness, is best introduced by examples. Vision gives knowledge about the height of a tree, hearing about the loudness of a noise, touch about the temperature of a surface, smell about the age of an egg, taste about the constituents of a drink. Memory gives knowledge about the length of a walk, testimony about the physical characteristics of a criminal. The list could of course be continued indefinitely. In each case, the knowledge is inexact. One sees roughly but not exactly how many books a room contains, for example: it is certainly more than two hundred and less than twenty thousand, but one does not know the exact number. Yet there need be no relevant vagueness in the number. The inexactness was in the knowledge, not in the object about which it was acquired.

Section 8.2 examines a case of inexact knowledge in detail, to reveal its underlying structure. Section 8.3 draws the moral that each case of inexact knowledge is governed by a principle requiring knowledge to leave a *margin for error.* Section 8.4 applies the account of inexact knowledge to the problem of vagueness, using margin for error principles to explain our ignorance in borderline cases. It discusses the features that make knowledge involving vague concepts a distinct species of the genus, inexact knowledge. Vague terms are sharp but unstable in meaning; this instability is a distinctive source of inexactness. Section 8.5 characterizes vagueness in a concept as its indiscriminability from other possible concepts, and reconciles this with our knowledge of the meaning of vague terms. Section 8.6 relates inexact knowledge to the non-transitivity of indiscriminability more generally. Section 8.7 extends the notion of inexactness to cognitive attitudes weaker than knowledge, such as reasonable belief.

8.2 THE CROWD

I see a vast crowd in a stadium. I wonder how many people there are. Naturally, I cannot know exactly just by looking. My eyesight and ability to judge numbers are nothing like that good, and a few people may not even be visible from where I stand. Since I have no other source of relevant information at present, I do not know exactly how many people there are. For no number m do I know that there are exactly m people. Nevertheless, by looking I have gained some knowledge. I know that there are not exactly two hundred or two hundred thousand people; I do not know whether there are exactly twenty thousand people. For many numbers m, I do not know that there are not exactly m people. More precisely: for many numbers m, I do not have knowledge whose content is expressed by the result of replacing 'm' in 'There are not exactly m people' by a numeral designating m (e.g. '20,000'). Not under discussion is knowledge whose content is expressed by a sentence in which m is designated by a definite description, such as 'The number of people in the stadium minus one', for I may not know which number fits the description.

The least number principle says that every non-empty set of natural numbers has a least member. It is equivalent to the principle of mathematical induction, that if 0 has a property F, and whenever m has F so has $m + 1$, then every natural number has F. The set of natural numbers m such that I do not know that there are not exactly m people is non-empty, for it contains the number 20,000. By the least number principle, the set contains a least member n. Thus:

(1) I know that there are not exactly $n - 1$ people.

(2) I do not know that there are not exactly n people.

'n' abbreviates the definite description 'the least number m such that I do not know that there are not exactly m people'. '$n - 1$' abbreviates a corresponding description. However, by remarks at the end of the last paragraph, (1) and (2) do not concern knowledge in which numbers are designated by such descriptions; what makes them true or false is knowledge or the lack of it in which n or $n - 1$ is designated by an ordinary numeral such as '20,000' or '19,999'. The descriptions determine which numerals should replace them (in technical terms, they have wide scope).

I do not know exactly how many people there are. I might guess but, even if I guessed right, that would not be *knowing* exactly how many people there are. If I judge that there are exactly m people, and in fact there are, then for all I know there are really $m - 1$ or $m + 1$; I do not know that there are not. *A fortiori*, if I do not judge that there are exactly m people, when in fact there are, I do not know that there are exactly m. Either way, if there are exactly m people, then I do not know that there are not exactly $m - 1$ or $m + 1$. These reflections are quite independent of the value of 'm'. There can be *no* number m such that there are exactly m people and I know that there not, say, exactly $m - 1$, for if there are exactly m and I judge that there are not exactly $m - 1$, I am merely guessing. Anyone who can know by looking from where I am that there are not exactly $m - 1$, when in fact there are exactly m, has much better eyesight than I have (including X-ray eyes to see through various obstacles) and a much greater ability to judge numbers. For every number m, if there are exactly m people, then I do not know that there are not exactly $m - 1$.[1]

The foregoing reflections are mine. *I* know that my eyesight and ability to judge numbers are not good enough for me to know by looking that there are not $m - 1$ people, when in fact there are m. I know that, for every number m, if there are exactly m people, then I do not know that there are not exactly $m - 1$. Indeed, I may be assumed to have pedantically instantiated this knowledge for each relevant number. For each such m, I know that if there are exactly m people, then I do not know that there are not exactly $m - 1$. For example, I know that if there are exactly 20,000 people, then I do not know that there are not exactly 19,999. Since the number n in (1) and (2) is one of the many for which I have instantiated my knowledge, it follows too that

(3) I know that if there are exactly n people, then I do not know that there are not exactly $n - 1$ people.

But now a contradiction threatens. By (1), I *do* know that there are not exactly $n - 1$ people. If I combine this knowledge with my knowledge in (3), I can apparently deduce, and thereby come to know, that there are not exactly n people, which contradicts (2). What has gone wrong?

The argument can be made more explicit. There is a limit to how much I can add to my knowledge of the number of people by reflecting on the limitations of my eyesight and ability to judge numbers, and making deductions from what I thereby know. It is indeed plausible, although this point will not be pressed, that such processes add *nothing* to the knowledge of the number that I have already gained. Be that as it may, the supposition can legitimately be built into the example that I go to the trouble of deducing in a logically competent way all propositions of the form 'There are not exactly m people' that can be deduced from what I know, for every relevant value of 'm'. After some time I have completed the task (the number of people in the stadium may be assumed not to change over this period). 'n' in (1)–(3) can be defined with respect to a time t at which this reflective equilibrium has been attained. Since the defining description for 'n' does not figure in the content of the knowledge under discussion, this procedure does not involve circularity, or a vicious regress. t can also be taken as the time of the knowing at issue in (1)–(3). Thus (3) says that I know at t that if there are exactly n people, then I do not know at t that there are not exactly $n - 1$ people. There is nothing viciously self-referential or ungrounded about this present knowledge of present (rather than past) ignorance, for it is based on general considerations about my eyesight and ability to judge numbers, not on a futile attempt to survey all the propositions I do not presently know. Now if it follows from what I know at t that there are not exactly n people, then by hypothesis I have already deduced that by t, and so come to know it. Thus at t:

(4) If I know some propositions, and from those propositions it logically follows that there are not exactly n people, then I know that there are not exactly n people.

Note that (4) is not a general principle about knowledge; it is merely a description of my state in a particular situation (the same goes for (1)–(3), of course). The paradox is that each of (1)–(4) appears to be true with respect to the envisaged mundanely possible situation, yet they appear to be mutually inconsistent.

One thought is that I might know my premises and competently deduce my conclusion without thereby coming to know it because, although the premises are probable enough to count as known, the conclusion is not. For a logical consequence of two propositions may be less probable than each of them. Thus (4) might fail. For several reasons, this is an unpromising diagnosis.

(a) It is counter-intuitive to suppose that making competent deductions from what one knows is not in general a way of extending one's knowledge.
(b) The connection assumed between knowledge and probability has not been made out. No degree of probability less than 1 by itself makes knowledge out of true belief; the pessimistic owner of a ticket in a fair lottery who rightly believes that it will not win does not know that it will not, however many tickets have been sold.
(c) What sort of probability is in question? Degrees of belief seem too dependent on the subject, and propensities in the world not dependent enough, to form a standard for knowledge. If one considers

probabilities conditional on what I know, then what I know has probability 1 and the objection lapses.

(d) Nothing in the example disallows the assumption that there is a probability in the relevant sense of 1 that my eyesight and ability to judge numbers are not good enough for me to know that there are not $n - 1$ people, given that there are in fact n. Then the probability of the conditional in (3) will be 1, as will be the probability of its consequent conditional on its antecedent. But then the objection to (4) still lapses, for the required drop in probability from premises to conclusion cannot occur. If two premises entail a conclusion, and one premise has probability 1, then the conclusion is at least as probable as the other premise.

(e) One can always introduce a notion of knowledge#, where to know# a proposition is to know either it or propositions from which one has competently deduced it. I do not know# exactly how many people there are. One can run through the argument with 'know&##8217; in place of 'know'; (4) is then certainly correct.[2]

In the light of (a)–(e), the paradox should not be blamed on the failure of competent deduction to extend knowledge.

Another thought is that vagueness itself is somehow to blame. If it were, the case could not be used to cast independent light on the problem of vagueness. However, it will be argued without appeal to the epistemic theory of vagueness that vagueness is not to be blamed in the present case. Some examples of inexact knowledge depend on vagueness, but not all do.

The least number principle is sometimes held not to be valid in the presence of vagueness; what is the least number of grains to make a heap? The charge is that 'n' in (1)–(4) fails to refer. It was defined as 'the least number m such that I do not know that there are not exactly m people'. The only word here that might be relevantly vague is 'know'. It is indeed vague to some extent; there are borderline cases of knowledge. The question is whether this vagueness is the source of the paradox. If it is, the paradox would vanish if 'know' were made precise by arbitrary stipulations. We could not expect in practice to eliminate all the vagueness in 'know', but we do not need to. All we need are stipulations that resolve each borderline case of the form 'I know that there are not exactly m people' in the present context, for then we could apply the least number principle to define 'n', making (1) and (2) true. For example, let us adopt the rule that all relevant borderline cases are to be resolved conservatively with respect to 'know', thereby raising the standard for what it takes to 'know'. A corresponding rule may be adopted to resolve any relevant borderline cases for 'borderline case', and so on. Such stipulations do nothing to improve my eyesight or ability to judge numbers, of course. Now the conservative stipulations will, if anything, reinforce the truth of conditionals of the form 'If there are exactly m people, then I do not know that there are not exactly $m - 1$ people', for they make it harder to know that there are not exactly $m - 1$ people in the new sense of 'know'. Moreover, my attitude to the conditionals was already a clear case of knowledge, not a borderline case, so I should know them in the new sense too. So (3) remains true on the new reading, as does (4). Thus the apparently inconsistent propositions (1)–(4) remain apparently true when the stipulations are made. Since the paradox would not vanish if 'know' were made relevantly precise, and no other term is relevantly vague, vagueness is not the source of the difficulty. It is therefore legitimate in what follows to treat 'know' as though it were precise.

The paradoxical reasoning needs to be examined in more detail. When (4) is applied, what are the propositions from which it logically follows that there are not exactly n people? The major premise is the conditional 'If there are exactly n people, then I do not know that there are not exactly $n - 1$ people'. The form of the inference is from 'If A, then not B' and 'B' to 'Not A'. Thus 'A' is 'There are exactly n people' and 'B' is 'I know that there are not exactly $n - 1$ people'. If I am to use the inference to gain knowledge of its conclusion, I must know its premises, as (4) requires. (3) says that I know the major premise. I must also know the minor, 'B', i.e. (1). That is, the paradoxical reasoning assumes:

(5) I know that I know that there are not exactly $n - 1$ people.

The mutually inconsistent propositions are (2)–(5), not (1)–(4).

One might think for a moment that since I know by definition of 'n' that (1) is true, (5) is true. However, that would be to mistake the role of the definite description abbreviated by '$n - 1$', 'the predecessor of the least number m such that I do not know that there are not exactly m people', in (5). Certainly I know that, for every number k, if k fits the description then I know that there are not exactly k people. But that is not what (5) says.

The whole argument concerns knowledge in which numbers are designated by numerals, not by definite descriptions such as that above; only on this understanding is (3) plausible. What (5) says is that if a number fits the description abbreviated by '$n - 1$', then I know that I know that there are not exactly that number (as designated by a numeral) people. In technical terms, the definite description must be given the widest possible scope. If (5) were read otherwise, the reasoning would be trivially fallacious. Since I do not know what number fits the definite description abbreviated by '$n - 1$', my reflection on the description does nothing to make (5) true.

The apparently true propositions are (1)–(4). The inconsistent ones are (2)–(5). The right thing to do is the obvious one: accept (1)–(4) and reject (5). This is to reject the 'KK' principle that if I know something, then I know that I know it. I know that there are not exactly $n - 1$ people, but I do not know that I know that there are not exactly $n - 1$ people.[3]

Before we adopt the proposal, we should check that it really does meet the difficulty. We can do this by giving a consistency proof for (1)–(4) and the negation of (5). This will be done by means of a simple model of some situations in one of which (1)–(4) are true and (5) is false. The model is not intended to be realistic; what it shows is the structure of the envisaged solution to the paradox. Although the argument could be carried out in purely formal terms, the construction is more revealing when made informally.

For each natural number m, let s_m be a situation in which there are exactly m people in the stadium. Thus for any number k (the result of substituting the numeral for k for 'k' in) 'There are exactly k people' is true in s_m if and only if $k = m$. Truth-values can now be assigned recursively to compound sentences in different situations. 'Not A' is true in s_m if and only if 'A' is not true in s_m. 'If A, then B' is true in s_m if and only if either 'A' is not true in s_m or 'B' is true in s_m. 'I know that A' is true in s_m if and only if 'A' is true in each of s_{m-1} (if it exists), s_m and s_{m+1}. I know the proposition just in case it is true in the situation I am in and every situation like it except for a difference of one in the number of people in the stadium.

It is easy to check that, for any positive m, the conditional 'If there are exactly m people, then I do not know that there are not exactly $m - 1$ people' is true in every situation. Consequently, 'I know that if there are exactly m people, then I do not know that there are not exactly $m - 1$ people' is also true in every situation. In particular, (3) is true in every situation. Moreover, in this simple model I know all the logical consequences of what I know, for if every premise of a logically valid argument is true in each of s_{m-1}, s_m and s_{m+1}, then so too is its conclusion. Thus (4) is also true in every situation. Since 'There are not exactly $n - 1$ people' is true in every situation but s_{n-1}, 'I know that there are not exactly $n - 1$ people', i.e. (1), is true in every situation but s_{n-2}, s_{n-1} and s_n. Hence 'I know that I know that there are not exactly $n - 1$ people', i.e. (5), is true in every situation but s_{n-3}, s_{n-2}, s_{n-1}, s_n and s_{n+1}. Similarly, 'I know that there are not exactly n people' is true in every situation but s_{n-1}, s_n and s_{n+1}, so its negation (2) is true in just those three situations. Thus (1)–(4) are all true in s_{n+1}, while (5) is false.[4]

The model shows that the proposed solution to the paradox is consistent. In order to restore consistency, it was enough to deny the KK principle. Indeed, the model proves that one can consistently combine (1) and (2) with unqualified versions of the generalizations from which (3) and (4) were derived:

(3+) For every number m, I know that if there are exactly m people, then I do not know that there are not exactly $m - 1$ people.

(4+) For every number m, if I know some propositions, and from those propositions it logically follows that there are not exactly m people, then I know that there are not exactly m people.

To emphasize: (3+) and (4+) purport to describe my state at time t, not to lay down general principles about knowledge. In a realistic example, an upper bound would be imposed on m, if I am not to know infinitely many propositions.

The paradoxical reasoning is generated by the combination of (3+) and (4+) with the KK principle. It can be reformulated as appealing directly to the principle of mathematical induction, rather than to the least number principle. I certainly know that there are not exactly 0 people. But if I know that there are not exactly $m - 1$ people, then by the KK principle I know that I know that there are not exactly $m - 1$ people; so by (3+) I know the premises of a valid deductive argument whose conclusion is that there are not exactly m people; so by (4+) I know that there are not exactly m people. By mathematical induction, for every natural number m I know that there are not exactly m people. That is impossible, for the number of people in the stadium is finite. As the case was described, I do not know that there are not exactly 20,000 people. This upper bound means that, strictly speaking, the appeal to mathematical induction can be dispensed with, for one can reach absurdity arguing step by step without use of special principles of mathematics. The appeal to the least number principle was equally unnecessary. However, the principles conveniently label the difference in flavour between the two versions, and speed up both. What neither argument can dispense with is the KK principle. It is the culprit.

The failure of the KK principle is not news. However, the standard counterexamples involve knowing subjects who lack the concept of knowledge or have not reflected on their knowledge and therefore do not know that they know. The present case is quite different. It concerns a subject who has the concept of knowledge and has reached reflective equilibrium with respect to the propositions at issue. Still I know without knowing that I know.

Our knowledge is riddled with failures of the KK principle, for it is riddled with inexactness. The problem about the stadium could be duplicated with respect to almost any case of sense perception, for almost any such case gives inexact knowledge about numbers or quantities of some kind. In many cases, the quantities will lie on a continuous scale, as when I see a tree and have inexact knowledge of its height. That presents no obstacle to the argument, for continuous scales are divided into units; I have inexact knowledge of the height of the tree in inches to the nearest inch. The point generalizes to knowledge from sources beyond present perception, such as memory and testimony. This is partly because they pass on inexact knowledge derived from past perception, partly because they add further inexactness of their own. In each case the possible answers to a question lie so close together that if a given answer is in fact correct, then one does not know that its neighbouring answers are not correct, and one can know that one's powers of discrimination have that limit. The argument proceeds as before.

The aim of Section 8.3 is to describe a model of inexact knowledge on which its failure to provide the KK principle looks utterly natural.

8.3 MARGINS FOR ERROR

Suppose that, in the situation described in Section 8.2, there are exactly *i* people and I have the true belief that there are not exactly *j* people. If the difference between *i* and *j* is too small, then even if there had been exactly *j* people, I might easily still have believed that there were not exactly *j*. It is not the case that if there had been exactly *j* people I should not have believed that there were not exactly *j*. My eyesight and ability to judge numbers are limited. But then my true belief that there are not exactly *j* is not reliably true; it is too risky to constitute knowledge. On the other hand, if the difference between *i* and *j* is large enough, and my belief is formed in a normal way, then it may well be the case that if there had been exactly *j* people I should not have believed that there were not exactly *j*. My true belief may then be reliably true, and constitute knowledge.[5] Other things being equal, I know that there are not exactly *j* people if and only if the difference between *i* (the actual number) and *j* is large enough. How large is large enough? That depends on the circumstances: my eyesight and ability to judge numbers, the obstacles occluding part of the crowd, the quality of the light.

Where our knowledge is inexact, our beliefs are reliable only if we leave a margin for error. The belief that a general condition obtains in a particular case has a margin for error if the condition also obtains in all similar cases. The degree and kind of the required similarity depend on the circumstances. For given cognitive capacities, reliability increases with the width of the margin. The more accurate the cognitive capacities, the narrower is the margin needed to achieve a given level of reliability. Since a belief constitutes knowledge only if it is reliable enough, the belief that a general condition obtains in a particular case constitutes knowledge only if the condition obtains in all cases similar enough in the relevant respects to achieve the required level of reliability. Knowledge that the condition obtains is available only if it does obtain (whether knowably or not) in all sufficiently similar cases. If it obtains in a case sufficiently similar to a case in which it does not obtain, then knowledge that it obtains is unavailable in both cases. It cannot be known to obtain within its margin for error.

A *margin for error principle* is a principle of the form: ‘*A*’ is true in all cases similar to cases in which ‘It is known that *A*’ is true. Which margin for error principles obtain depends on the circumstances; one cannot specify *a priori* the required degree and kind of similarity. One can, however, state a margin for error meta-principle: that where knowledge is inexact, some margin for error principle holds. The meta-principle is necessarily unspecific, but it is not trivial. In particular, it does not hold merely by definition of ‘inexact knowledge’, for the phrase was defined by examples, not by reference to margin for error principles. Rather, inexact knowledge is a widespread and easily recognized cognitive phenomenon, whose underlying nature turns out to be characterized by the holding of margin for error principles.

The required similarity need not be specified in the same terms as those used to express the knowledge in question. For example, I can know by looking that there are exactly five people in a room. That belief would be false in any case differing from the actual one in the number of people. Yet it constitutes knowledge, and its source is of the same general kind as my inexact knowledge of the number of people in the stadium. It is indeed governed by a margin for error principle. If I know by looking that the room is in a certain condition (such as that of containing five people), then the room is in that condition in any case differing from the present one only by rearrangements of a few molecules. I can know that there are exactly five people because there would still be exactly five in any case within the relevant margin for error.

A special case of inexact knowledge is that in which the proposition ‘*A*’ is itself of the form ‘It is known that *B*’. Just as we are not perfectly accurate judges of the number in a crowd, so we are not perfectly accurate judges of the reliability of a belief. A margin for error principle for 'It is known that *B*' in place of '*A*' says that 'It is known that *B*' is true in all cases similar to cases in which 'It is known that it is known that *B*' is true. As usual, the required degree and kind of similarity depend on the circumstances, for example on one's ability to judge reliability; 'It is known that *B*' and '*B*' may need margins for error of different widths. If 'It is known that *B*' is true but there are sufficiently similar cases in which it is false, then it is not available to be known. It cannot be known within its margin for error. Thus the failure of the KK principle is a natural consequence of the inexactness of our knowledge of our knowledge. By the margin for error meta-principle, our knowledge of our knowledge is governed by a margin for error principle, from which it follows that the KK principle is false.

Suppose that 'It is known that it is known that *B*' is true in a given case. By a margin for error principle for 'It is known that *B*', the latter proposition is true in all cases similar to the given case. But then by a margin for error principle for '*B*', '*B*' is true in all cases similar to cases similar to the given case. In effect, knowledge that one knows requires two margins for error. More generally, every iteration of knowledge widens the required margin. Any number of iterations of knowledge is possible in principle, but is available in a narrower range of cases than any lower number of iterations.

Believing is often compared to shooting at a target, the truth. The comparison is not quite apt, for the truth is a single point (the actual case), like a bullet, while the proposition believed covers an area (a set of possible cases), like a target. Instead, the believer's task may be conceived as drawing a boundary on a wall at which a machine is to fire a bullet. The belief is true if the bullet hits the bounded area, false otherwise. If truth is a hit, knowledge is a safe hit.[6] That is, the point of impact is within the bounded area and not so near its boundary that the bullet could very easily have landed outside (had a light breeze blown). For example, a hit might be safe just in case every point on the wall less than an inch from the point of impact is within the bounded area. The one-inch margin inside the boundary corresponds to the cases in which '*B*' is true but unknown; when this margin is removed from the bounded area, the remaining area corresponds to the cases in which '*B*' is known. When another margin is removed, the result corresponds to the cases in which '*B*' is known to be known. The iteration of knowledge operators is a process of gradual erosion.[7]

An area lacks a margin only if no point in it is less than an inch from a point on the wall not in the area. Since any point on the wall can be reached from any other via a sequence of points each less than an inch from the next, if an area lacks a margin then either no points are in it or all points are in it. It is either the 'null area', corresponding to a contradiction, or the whole wall, corresponding to a tautology.[8] On this model, a proposition is available to be known whenever it is true only if it is either logically true or logically false. A contingent proposition corresponds to a non-null area less than the whole wall and has a margin; such a proposition is true in some cases where it is not available to be known.

Not only can any point on the wall be reached from any other via a sequence of points each less than an inch from the next: there is a finite bound to the number of points needed, since diagonally opposite corners of the wall are as far apart as any two points on it. Thus any area less than the whole wall is reduced to nothing by a finite number of removals of one-inch margins. This corresponds to the claim that for any proposition other than a logical truth there is a finite bound to the number of iterations of knowledge one can have of it – a mildly sceptical result.[9]

The remarks in the last two paragraphs should be treated with more than usual caution, for they depend on specific features of the very simple layout imagined. If the wall were divided down the middle by a partition at right angles to it, so that there was no danger of a bullet fired on one side of the partition landing on the other, the half of the wall on one side of the partition would have no margin, and would therefore correspond to a contingent proposition available to be known whenever true. If the wall were infinitely long, with no partition, any two points on it would still be a finite distance apart, but the half of the wall on one side of an arbitrary vertical line would not be reduced to nothing by any finite number of removals of one-inch margins; it would correspond to a contingent proposition of which one could have any finite number of iterations of knowledge. The constant width of the margin is another simplification. If it were draughtier on the left than on the right, making the bullet's flight less predictable on that side, a wider margin would be needed there. Our cognitive capacities are more accurate in some areas than in others.

Such qualifications could be multiplied. However, the point of a model is to omit all but a few of the original's complications. What has been sketched is a picture of inexact knowledge in which the systematic failure of the KK principle is utterly natural. It is to be expected not just because we are not perfectly reflective, but because, however reflective we are, our cognitive capacities are not perfectly accurate. The next task is to apply these ideas to knowledge whose inexactness stems from the vagueness of its content.[10]

8.4 CONCEPTUAL SOURCES OF INEXACTNESS

The usual sources of inexactness infect vague judgements just as much as they do precise ones. Perceptual knowledge of someone's girth in inches is inexact; so is perceptual knowledge of his thinness. In borderline cases, however, our ignorance goes deeper than that.

On the epistemic view of vagueness, there are values of '*m*' for which anyone with exact physical measurements *m* is unknowably thin. Now if I have physical measurements *m*, I do not have them essentially; the sentence 'TW is thin' expresses a contingent truth. But since thinness supervenes on exact physical measurements, the generalization 'Everyone with physical measurements *m* is thin' expresses a necessary truth (see Section 7.4). Since I can be known to have physical measurements *m*, the ignorance postulated by the epistemic view involves ignorance of such necessary truths.[11] Yet they seem trivially to satisfy the necessary condition for being known imposed by a margin for error principle. A necessary truth is true in all cases; *a fortiori*, it is true in all cases similar to the case in which it is a candidate for being known. How then can a margin for error principle explain our ignorance of a necessary truth?

Someone who asserts 'Everyone with physical measurements *m* is thin' is asserting a necessary truth, but he is still lucky to be speaking the truth. He does not know the truth of what he says. Although he could not have asserted the proposition he actually asserted without speaking truly, he could very easily have asserted a different and necessarily false proposition with the same words.[12]The extension of the word 'thin' could very easily have been slightly different, so that it would have excluded everyone with physical measurements *m*. What distinguishes vagueness as a source of inexactness is that the margin for error principles to which it gives rise advert to small differences in meaning, not to small differences in the objects under discussion.

Consider again the supervenience of meaning on use, at least for a fixed contribution from the environment (as in Section 7.5). For any difference in meaning, there is a difference in use. The converse does not always hold. The meaning of a word may be stabilized by natural divisions, so that a small difference in use would make no difference in meaning. A slightly increased propensity to mistake fool's gold for gold would not change the meaning or extension of the word 'gold'. But the meaning of a vague word is not stabilized by natural divisions in this way. A slight shift along one axis of measurement in all our dispositions to use 'thin' would slightly shift the meaning and extension of 'thin'. On the epistemic view, the boundary of 'thin' is sharp but unstable.

Suppose that I am on the 'thin' side of the boundary, but only just. If our use of 'thin' had been very slightly different, as it easily could have been, then I should have been on the 'not thin' side. The sentence 'TW is thin' is true, but could very easily have been false without any change in my physical measurements or those of the relevant comparison class. Moreover, someone who utters the sentence assertively could very easily have done so falsely, for the decision to utter it was not sensitive to all the slight shifts in the use of 'thin' that would make the utterance false.

The point is not confined to public language. Even idiolects are vague. You may have no settled disposition to assent to or dissent from 'TW is thin'. If you were forced to go one way or the other, which way you went would depend on your circumstances and mood. If you assented, that would not automatically make the utterance true in your idiolect; if you dissented, that would not automatically make it false. What you mean by 'thin' does not depend solely on what you would say in your present circumstances and mood. You have no way of making each part of your use perfectly sensitive to the whole, for you have no way of surveying the whole. To imagine away this sprawling quality of your use is to imagine away its vagueness.

Similar points apply to concepts. You have no way of making your use of a concept on a particular occasion perfectly sensitive to your overall pattern of use, for you have no way of surveying that pattern in all its details. Since the content of the concept depends on the overall pattern, you have no way of making your use of a concept on a particular occasion perfectly sensitive to its content. Even if you did know all the details of the pattern (which you could not), you would still be ignorant of the manner in which they determined the content of the concept.

The plausibility of the claim that vagueness gives rise to margin for error principles does not depend on the epistemic view of vagueness. Consider the term 'heap', used in such a way that it is very vague.[13] Someone who asserts 'n grains make a heap' might very easily have made an assertion with that sentence even if our overall use had been slightly different in such a way as to assign the sentence the semantic status presently possessed by '$n - 1$ grains make a heap'. A small shift in the distribution of uses would not carry every individual use along with it. The actual assertion is the outcome of a disposition to be reliably right only if the counterfactual assertion would have been right. Thus the actual assertion expresses knowledge only if the counterfactual assertion would have expressed a truth. By hypothesis, the semantic status of 'n grains make a heap' in the counterfactual situation is the same as that of '$n - 1$ grains make a heap' in the actual situation; if the former expresses a truth, so does the latter. Hence, in the present situation, 'n grains make a heap' expresses knowledge only if '$n - 1$ grains make a heap' expresses a truth. In other words, a margin for error principle holds:

(!) If we know that n grains make a heap, then $n - 1$ grains make a heap.

The argument for (!) does not appeal to the epistemic view of vagueness at any point. Someone who rejected the view would not be forced to reject (!), and might well wish to accept it.

(!) might be thought to generate a sorites paradox, not for 'heap' but for 'known heap'. If we know (!) (by philosophical reflection on our vague use of 'heap'), and deduce the relevant logical consequences of what we know, does it not follow that if we know that n grains make a heap, then we know that $n - 1$ grains make a heap? Since we know that 10,000 grains make a heap, it would follow by 10,000 applications of *modus ponens* that 0 grains make a heap, which they do not. However, this is just a variant of the fallacious argument about the crowd. In order to know that $n - 1$ grains make a heap, we should have to know the premises from which we deduced '$n - 1$ grains make a heap'. They are the relevant instance of (!) and its antecedent, 'We know that n grains make a heap'. Thus we should have to know that we know that n grains make a heap. But the previous stage of the argument showed only that we know that n grains make a heap. With the KK principle, a sorites paradox would indeed be forthcoming. Without it, one iteration of knowledge is lost at each stage of the argument. For any reasonable number of iterations at the start, the argument runs out of steam before reaching a false conclusion. Indeed, it must do so, for the model used in Section 8.2 to show the consistency of (1)–(4) can equally be used to show that it is consistent to assume that we have several iterations of knowledge of each instance of (!), and of both the proposition that 10,000 grains make a heap and the proposition that 0 grains do not.

Given (!), we cannot know a conjunction of the form 'n grains make a heap and $n - 1$ grains do not make a heap'. To know the conjunction, we should have to know its first conjunct; but then by (!) its second conjunct would be false, making the whole conjunction false and therefore unknown. (!) encapsulates our ignorance of the cut-off point for 'heap'. A similar account can be given for other vague terms. A well-constructed sorites series makes an analogue of (!) true, because it proceeds by steps smaller than the relevant margin for error, so that if the term is known to apply to one member, then it does apply to the next.

What (!) conceals is the source of the inexactness. The small differences it displays are in the number of grains, but the underlying explanation appealed to small differences in the use of 'heap'. The argument for (!) could translate the latter into the former because a shift in the whole pattern of use of 'heap' by one step along the sorites series would be suitably small. One can construct artificial cases in which such a shift would be noticeable, so that the analogue of (!) might fail. Suppose, for the sake of argument, that 'several' is used in such a way that any plurality of more than three things is clearly several things and any plurality of less than three things is clearly not several things. Pluralities of three are the only borderline cases. A shift in the whole pattern of use of 'several' by one, so that pluralities of four became the only borderline cases, might well be noticeable. One could not maintain the analogues of (!) for both 'several' and 'not several', for that would involve maintaining both that if we know that four are several then three are several and that if we know that two are not several then three are not several. For we do know both that four are several and that two are not, so three would have to be both several and not several. If no one ever classified pluralities of three as 'several' or as 'not several', then one might after all be able to know whether three were several by the kind of argument mentioned in Section 7.7. If, however, the use of 'several' in this case is messy in the way characteristic of vague terms in natural language, that will not be possible. Rather, individuals will sometimes classify three as 'several' or 'not several', and might well have used the same words even if the frequency of such uses had been slightly different. A margin for error principle still governs such uses: if 'Three are several' expresses knowledge, then it would still have expressed a truth in those counterfactual situations. But even if 'Three are several' does express a truth, it would fail to do so in some of the counterfactual situations. By the margin for error principle, three are not known to be several. Thus our ignorance can still be explained by appeal to a margin for error principle in the form that most closely reflects the conceptual source of the inexactness.

Without appeal to the epistemic account of vagueness, one can argue that if vague terms have sharp boundaries, then we shall not be able to find those boundaries. Once one has seen this point, one can hardly regard our inability to find them as evidence that they do not exist. But if one has not seen the point, one might naturally suppose that if they exist then we should be able to find them, and so regard our inability as evidence of their non-existence. Thus margin for error principles explain both the ignorance postulated by the epistemic view and the apparent intuitions that run counter to that view. They do not commit one to the view, but they undermine some popular reasons for resisting it.

8.5 RECOGNITION OF VAGUE CONCEPTS

Vagueness is a source of inexactness, Section 8.3 argued, because individual uses of a vague term are not fully sensitive to small differences in the overall pattern on which small differences in meaning supervene. The argument seems to appeal to indiscriminable differences in meaning.

It is tempting to conclude that, contrary to Section 7.6, speakers of a vague language do not know exactly what their utterances mean. The argument of this section is that the tempting conclusion does not follow.

On the epistemic view, an utterance of a vague sentence such as '*n* grains make a heap' may express a necessary truth in a borderline case. A speaker who made such an assertion would not be expressing knowledge that *n* grains make a heap, for he might easily have used those words even if their overall use had been slightly different, so that they expressed a necessary falsehood. His utterance *u* does not manifest a disposition to be reliably right. As things actually are, *u* says that *P*. In the counterfactual situation, *u* says that *P**. The speaker would not recognize the difference. He does not seem to know in the actual situation that *u* does not say that *P**.

The epistemic theorist may concede that speakers do not know that *u* does not say that *P**. The question is whether it follows that they do not know what *u* says, in other words, that they do not know that *u* says that *P*. To put it the other way round, if they know that *u* says that *P*, does it follow that they know that *u* does not say that *P**? Does it even follow that they *can* know that u does not say that *P**?

If *u* says that *P*, then it does not say that *P**. Now if speakers of the language know that conditional, and also know that *u* says that *P*, then they can combine those pieces of knowledge and deduce that *u* does not say that *P**. But although the conditional is true, it does not follow that speakers of the language know it to be true. If they cannot discriminate what *u* actually says from what it counterfactually says, they cannot be expected to know the conditional. On the epistemic view, perhaps speakers know that *u* says that *P*, but cannot know that *u* does not say that *P**, and so cannot know that if *u* says that *P*, then *u* does not say that *P**.

The epistemic theorist is not alone in supposing that our ability to recognize the meaning of an utterance does not require us to discriminate it from all other possible meanings. On almost any view, the meaning of a vague utterance lies on something like a continuum. Even fuzzy boundaries lie in a continuum of possible fuzzy boundaries, varying in location and fuzziness. The sentence could have expressed a very slightly different meaning, and would have done so if its use had been very slightly different. One cannot expect speakers of the language to be able to discriminate between all such possible meanings. Several indiscriminable semantic differences can add up to a discriminable semantic difference. In being forced to acknowledge this fact, the epistemic theory is no worse off than its rivals.

One might react to the phenomenon of indiscriminable semantic differences by concluding that speakers only roughly know what their utterances mean; they cannot uniquely identify their meanings. If this reaction is open to anyone, it is open to the epistemic theorist. However, a less sceptical line of thought deserves to be explored.

Consider our ability to recognize faces. We often know exactly who someone is by seeing her face. Nevertheless, there could easily have been (and perhaps is) someone else facially indiscriminable from the known person, whom we should have misidentified on confrontation as the person we know. Our ability to recognize our friends and relations is not undermined by the mere possibility of look-alikes, although it might be undermined by their actual presence in the neighbourhood. Similarly, why should our ability to recognize the meaning of utterances in our language be undermined by the mere possibility of indiscriminably different meanings? It is not as though such meanings need be actually present in the language. In particular, slight differences in use between speakers do not generate indiscriminably different meanings, for linguistic meaning is socially determined (Section 7.6). Of course, just as there are genuine look-alikes, so indiscriminable semantic variation may genuinely occur, and where it does so our knowledge of meaning is rather uncontroversially undermined. The point is that the epistemic theorist has no more reason than anyone else to suppose that such variation is actually universal. There is a sense in which we often know exactly what an utterance means.

One can think of actual people as located on a continuum of possible people: but it does not follow that to recognize a person one must locate her on that continuum. It is enough to know which of the actual people she is. Similarly, one can think of actual meanings as located on a continuum of possible meanings: but it does not follow that to recognize a meaning one must locate it on that continuum. It is enough to know which of the actual meanings it is. To do that, it is enough to use the term within the appropriate practice, as discussed in Section 7.6.

The view just sketched is quite consistent with the relevant margin for error principles. If 'heap' had meant something slightly different, speakers would have recognized that slightly different meaning. They would not have misidentified it as the present meaning. Whatever the exact details of their dispositions to assent and dissent, they would then have been participants in the practice of using 'heap' as it would then have been. The identification even of a vague meaning can manifest a disposition to be reliably right.

The vagueness of an expression consists in the semantic differences between it and other possible expressions that would be indiscriminable by those who understood them. Similarly, the vagueness of a concept consists in the differences between it and other possible concepts that would be indiscriminable by those who grasped them. The greater the indiscriminable differences, the greater the vagueness. Nevertheless, vague expressions can be understood, and vague concepts grasped, for the indiscriminable differences need not actually arise.

8.6 INDISCRIMINABLE DIFFERENCES

Vagueness issues from our limited powers of conceptual discrimination. It is often associated with the expression in logic of such limits: the non-transitivity of indiscriminability. If a sample x is indiscriminable in colour from a sample y, for example, and y is indiscriminable in colour from a sample z, it does not follow that x is indiscriminable in colour from z. Someone who can discriminate in colour between x and z may count x as 'red' and z as 'not red'; y seems destined to be a borderline case. This section investigates the connection between vagueness and the non-transitivity of indiscriminability. More generally, it investigates the connection between inexactness and the latter. For although inexactness has perceptual as well as conceptual sources, the resulting limitations on our knowledge share a common structure; they all give rise to margin for error principles. Since the structure is easier to discern when the source of the inexactness is perceptual, such examples will be used.[14]

The non-transitivity of indiscriminability is sometimes held to characterize only discrimination by direct comparison. On this view, a transitive form of indiscriminability is restored once indirect comparisons are permitted. For example, one can discriminate in colour between *x* and *y* indirectly, by noticing that one can directly discriminate the former but not the latter in colour from *z*. Call two things indirectly indiscriminable in a certain respect just in case they are directly indiscriminable in that respect from exactly the same things. Indirect indiscriminability is by definition a transitive relation. Admittedly, it can be hard to know that two things are indirectly indiscriminable, for all the things that might be directly indiscriminable from one but not the other must somehow be surveyed. In our everyday use of language, we rarely bother with indirect discriminations, and therefore lapse into vagueness. If we started to rely on such discriminations, we could no longer base our judgements on casual observation. Our use would lose its convenient vagueness. Nevertheless, it is suggested, indirect discrimination is available in principle as a standard, if we care to be more precise.[15]

The appeal to indirect discrimination presupposes that it is indeed a form of discrimination. To discriminate between *x* and *y* is to know that they are different.[16] Unless indirect discrimination involves such knowledge, it cannot be used as a standard for precise use, a reason for treating *x* and *y* differently. Now even if one has found a *z* from which one can in fact directly discriminate *x* but not *y*, that alone does not enable one to know that *x* and *y* are different; one must *know* that one can directly discriminate *x* but not y from *z*. Such knowledge may not be forthcoming in cases of inexactness. The point is closely related to the failure of the KK principle. If one has directly discriminated *x* from *z*, one knows that they are different; but if one is in no position to know that one knows that they are different, how can one know that one has discriminated between them? Equally, one might be unable to discriminate directly between *y* and *z*, but not be sure that one had not done so; how could one then know that one could not directly discriminate between them?

The point can best be substantiated by an example. I have been passing a certain tree on most days for several years, often giving it a casual glance, never attempting to measure it. My present knowledge of its height on each of the past 5,000 days is certainly inexact. My eyesight, my memory and my ability to judge heights are all imperfect. I do know that the height of the tree now (on day 5,000) is greater than it was at the beginning (on day 0), so:

(6) I know that the height on day 0 is not the same as the height on day 5,000.

I also know that, if the tree grew by less than a millimetre between day *i* and day *j*, then, for all I know, it did not grow at all in that period. To detect growth on that scale, one would need a much better eyesight, memory and ability to judge heights than I actually have. For all *i* and *j* between 0 and 5,000:

(7) I know that if the height on day *i* and the height on day *j* differ by less than a millimetre, then I do not know that the height on day *i* is not the same as the height on day *j*.

It will be convenient to read (7) and similar formulas as concerning only knowledge in which the numbers of the days are designated by numerals, as explained in Section 8.2. Like (3), (7) says that I know the contraposed form of a margin for error principle. I also know, on the testimony of a good botanist, that the tree cannot grow by as much as a millimetre in a single day. For all *i* between 0 and 5,000:

(8) I know that the height on day *i* and the height on day $i + 1$ differ by less than a millimetre.

By a rough estimate of the growth in height over the period, I can in fact estimate that the average growth per day is very much less than a millimetre.

I may be said to discriminate directly between day *i* and day *j* in the height of the tree just in case I know on the basis of my memory of those days that the height of the tree on day *i* was not the same as its height on day *j*. Days are directly indiscriminable in the height of the tree just in case I cannot discriminate directly between them in that respect (given my present circumstances). That relation is clearly not transitive, for each day has it to the next by (7) and (8), but by (6) the first day does not have it to the last.

Can I use indirect comparisons to discriminate more finely, perhaps even to falsify (7)? Day *i* may be said to be indirectly indiscriminable from day *j* just in case, for each day *k*, day *i* and day *k* are directly indiscriminable in the height of the tree if and only if day *j* and day *k* are directly indiscriminable in the height of the tree. Indirect indiscriminability is by definition an equivalence relation: it is reflexive, symmetric and transitive. Days are indirectly discriminable just in case they are not indirectly indiscriminable.

Since day 0 but not day 5,000 is directly discriminable from day 5,000, at least one day is indirectly discriminable from day 0. Let day m be the first such day. Thus day $m - 1$ is indirectly indiscriminable from day 0. If day m were indirectly indiscriminable from day $m - 1$, it would be indirectly indiscriminable from day 0 (by transitivity); since it is not, day m is indirectly discriminable from day $m - 1$. Unfortunately, I do not know which day is the first to be indirectly discriminable from day 0; I do not know which day 'm' designates. For any particular day, it is quite consistent with everything I know that the tree did not grow at all on that day, but grew steadily on every other day. In particular, I am in no position to know, even by inference, that the height of the tree on day $m - 1$ was not the same as on day m, when m and $m - 1$ are designated in my knowledge by numerals, not by definite descriptions such as that used to define 'm' ((6)–(8) concern only knowledge in which the numbers of days are so designated). Yet the two days count as 'indirectly discriminable' in the height of the tree.

Indirect discrimination is not a genuinely cognitive form of discrimination at all. If day i and day j are indirectly discriminable in the height of the tree, it does not follow that I am in a position to know inferentially that the height of the tree on day i was not the same as on day j. The problem is that I do not know exactly which direct discriminations I can make. Thus indirect discrimination is no threat to (7).

(6)–(8) remain a plausible description of my state when the term 'know' is used for inferential as well as non-inferential knowledge. Indeed, I may be assumed to have gained all the relevant knowledge available to me. I have deduced, and thereby come to know, all the relevant logical consequences of what I know; call this assumption 'Closure'. My ignorance of the height of the tree does not result from ignorance of logic.

I am ignorant about my direct discriminations. Either I know that the height of the tree on day i was not the same as on day j, but do not know that I know, or I do not know that the height on day i was not the same as on day j, but do not know that I do not know. The latter case is perhaps easier to imagine; what can now be shown is that the example involves cases of the former kind. As in other examples of inexact knowledge, the KK principle fails. The strategy of the argument is to show the KK principle to imply, for each i, that if I know that the height of the tree on day 0 is not the same as on day i, then I know that its height on day 0 is not the same as on day $i - 1$. Thus if I know that its height on day 0 is not the same as on day 5,000 (as I do, by (6)), then I know that the height on day 0 is not the same as on day 4,999, so I know that it is not the same as on day 4,998, so . . . . By 5,000 steps of the argument, I know that its height on day 0 is not the same as on day 0. Thus the KK principle will have been reduced to absurdity.

Suppose that I know that the height on day 0 is not the same as the height on day i. What must be deduced is that I know that the height on day 0 is not the same as the height on day $i - 1$. By the KK principle, I know that I know that the height on day 0 is not the same as the height on day i. By an instance of (7), I know that if the height on day 0 differs by less than a millimetre from the height on day i, then I do not know that the height on day 0 is not the same as the height on day i. By Closure, I know that the height on day 0 differs by not less than a millimetre from the height on day i. By (8), I know that the height on day $i - 1$ differs by less than a millimetre from the height on day i. By Closure again, I know that the height on day 0 is not the same as the height on day $i - 1$. QED

Once the KK principle is dropped, (6)–(8) and Closure form a consistent set. This can be shown by the construction of a simple model. For each subset X of the set of natural numbers from 1 to 5,000, let w_X be a 'possible world' in which, for each i from 1 to 5,000, the tree grows half a millimetre from day $i - 1$ to day i if i is a member of X, and otherwise does not grow at all in that time. The height of the tree on day 0 is the same in each world. For all worlds w_X and w_Y, say that w_Y is accessible from w_X just in case, for each day, the height of the tree in w_X on that day differs by less than a millimetre from its height in w_Y on that day. The worlds accessible from w_X are to be conceived as those in which everything known in w_X is true; if I am in w_X, and w_Y is accessible from w_X, then for all I know I am in w_Y; w_Y is epistemically possible relative to w_X. I know a proposition in w_X just in case that proposition is true in every world accessible from w_X; this assumption guarantees that I know all the logical consequences of what I know, and therefore validates Closure. It is not difficult to show that (7) and (8) are true in all worlds in the model, and that (6) is true in any world in which the tree grows on at least two days. Thus (6)–(8) and Closure are all true in worlds of the latter kind.

The failure of the KK principle in the model reflects the non-transitivity of the accessibility relation. For suppose that w_Z is accessible from w_Y, and w_Y from w_X, but that w_Z is not accessible from w_Y. Let '*A*' be a proposition true at just those worlds accessible from w_X. Thus 'I know that *A*' is true at w_X. However, it is not true at w_Y, for '*A*' is not true at w_Z, which is accessible from w_Y. Since w_Y is accessible from w_X, 'I know that I know that *A*' is not true at w_X. The KK principle breaks down at w_X. One can think of accessibility in the model above as a relation of indiscriminability between worlds. The KK principle fails because the indiscriminability of worlds is non-transitive.

The example began with the non-transitive indiscriminability of days in the height of the tree, and moved on to a similar phenomenon for worlds. It seems that this can always be done. Whatever *x*, *y* and *z* are, if *x* is indiscriminable from *y*, and *y* from *z*, but *x* is discriminable from *z*, then one can construct miniature worlds w_x, w_y and w_z in which the subject is presented with *x*, *y* and *z* respectively, everything else being relevantly similar. The indiscriminability of the objects is equivalent to the indiscriminability of the corresponding worlds, and therefore to their accessibility. The latter is therefore a non-transitive relation too. The proposition 'This is not *z*' is true in every world accessible from w_x, but not in w_z. As before, 'I know that this is not *z*' will be true in w_x, but 'I know that I know that this is not *z*' will be false. Thus the KK principle fails in w_x.

There is a complication. Discrimination is intentional, for it is a kind of knowledge. Even the example above involved discriminating between days in the height of the tree, i.e. between heights as presented to me by the tree on various days. Discriminability, in one sense of the term, can depend on the way in which objects are presented. Suppose, for example, that in the morning I count the number of birds in a cage and find that there are six; at noon I glance at the cage but do not notice how many birds it contains; in the evening I count the birds again, and find that one has gone. I do not know whether it escaped before or after noon. In that opaque sense, I can discriminate retrospectively the number of birds at the first count from the number at the second, but cannot discriminate the number at the first count from the number at noon, or the number at noon from the number at the second count. Only two numbers, five and six, were presented to me. One number was presented only once, by a count, the other twice, once by a count and once by a glance. I can discriminate the former number from the latter as presented by the count, but not as presented by the glance. In this case, the non-transitivity of indiscriminability in the opaque sense gives rise to no failure of the KK principle. Only two possibilities are relevant: either there were five birds at noon or there were six; I have no idea which. To model the situation, just two worlds are needed, indiscriminable from each other; this accessibility relation is transitive.

The example reveals a more radical way in which indirect discriminability in the opaque sense is not a genuine form of discriminability. I can discriminate the number of birds at the first count, but not the number at noon, from the number at the second count. It obviously does not follow that the number at the first count and the number at noon are distinct, let alone that I know that the number at the first count and the number at noon are distinct. For if the number at the first count is the number at noon, and I know that the number at the first count is not the same as the number at the second count, it does not follow that I know that the number at noon is not the same as the number at the second count. Even if the descriptions 'the number at the first count' and 'the number at noon' in fact denote the same number, the substitution of one for the other in the opaque context 'I know that . . .' can result in a change of truth-value in the sentence as a whole. Indirect discriminability in the opaque sense does not imply distinctness.

In the model used to prove the consistency of (6)–(8) and Closure, accessibility between worlds depends only on the growth of the tree in those worlds. Even there, however, indirect discriminability in the opaque sense does not imply distinctness. Let w_X be a world in which the tree does not grow at all from day 0 to day 1, and grows half a millimetre a day thereafter. Thus in w_X the height on day 0 is the same as the height on day 1. Moreover, I know that the height on day 0 is not the same as the height on day 3, for since the tree has grown a millimetre by day 3, in no world accessible from w_X is the height of the tree on day 0 the same as the height on day 3, for in each such world it has grown at least half a millimetre in that period. However, I do not know in w_X that the height on day 1 is not the same as the height on day 3. For from w_X a world w_Y is accessible in which the tree grows half a millimetre from day 0 to day 1, does not grow at all from day 1 to day 3, and grows half a millimetre a day thereafter; on each day the height in w_X differs from the height in w_Y by at most half a millimetre. In w_X, the height on day 0 and the height on day 1 are indirectly discriminable in the opaque sense, but they are not distinct. Such cases would be multiplied if the model took account of the varying degrees of attention I paid to the tree on various days, and other factors independent of its height.

Analogues of the phenomena discussed in this section will occur in the epistemology of vagueness, although they may be harder to discern. The failure of the KK principle will be manifested as higher-order vagueness. The failure of indirect discriminability to be a genuine form of discriminability will be another obstacle to attempts to make natural languages more precise.

8.7 INEXACT BELIEFS

Knowledge is not the only cognitive relation one cannot have to cut-off points for vague terms. One cannot know that n grains of sand make a heap and $n - 1$ do not; one also cannot reasonably believe it. The epistemic account has more to explain than the absence of knowledge.

The discussion so far has concerned knowledge. Some of its claims are equally plausible when 'reasonable to believe' is substituted for 'known'. Indeed, the discussion in Section 8.2 does not require what I believe to exceed what I know. Corresponding to (3), for example, is the plausible claim that it is reasonable to believe that if there are exactly n people in the stadium, then it is not reasonable to believe that there are not exactly $n - 1$. Margin for error principles, however, seem specific to knowledge. If one Φs a proposition in a situation s, one leaves a margin for error only if that proposition is true in all cases similar enough to s. Since s is certainly similar enough to itself, the proposition must be true in s. Thus if Φing requires a margin for error, one Φs only true propositions. Knowledge is such an attitude; reasonable belief is not. A subject with misleading evidence may reasonably believe false propositions. Since reasonable belief does not satisfy a margin for error principle, how can the epistemic theorist explain the unattainability of reasonable belief in borderline cases?

Parallel questions arise for many other cognitive attitudes, although not for all. An irrational person may certainly *believe* that n grains of sand make a heap and $n - 1$ do not. Parallel questions arise only for attitudes firmly based on evidence. The epistemic theorist can apply the account of inexact knowledge to such attitudes by working with a hypothesis: that one's evidence is simply what one knows. That the grass was wet, if one knows it, can be part of one's evidence for other beliefs. That the grass was wet, if one falsely believes it, cannot be part of one's evidence for other beliefs; only the evidence for the false belief can be part of the evidence for those other beliefs. In this sense, the hypothesis postulates an *externalist* concept of evidence. The status of a proposition as evidence does not depend solely on its place in the internal workings of the subject's head. Since only true propositions are known, only true propositions are evidence. Even so, there can be good evidence for a false proposition: when an innocent person is framed, for example. The restriction of one's evidence to what one knows is just what makes it plausible that, if Φing must be firmly based on evidence, then we cannot Φ that n grains make a heap and $n - 1$ do not.

A belief may be said to be reasonable just in case its probability conditional on the subject's evidence is high. On the present view, this is to say that the belief is reasonable just in case its probability conditional on what the subject knows is high. Now suppose, for simplicity, that the subject knows just the propositions that leave a margin for error δ. In a situation s, this amounts to knowing that one's situation is within δ of s, and knowing no more than that. Thus the probability of a belief conditional on what one knows may be conceived as the proportion of situations within δ of s in which the belief is true. On such simplifying assumptions, a belief is reasonable in a situation s just in case it is true in most worlds within δ of s.[17]

One can now explain, schematically, why for each number n it is not reasonable to believe that n grains make a heap and $n - 1$ do not. Indeed, one can explain a stronger principle:

(!!) If it is reasonable to believe that n grains make a heap, then it is not reasonable to believe that $n - 1$ grains do not make a heap.

Someone who, *per impossibile*, did reasonably believe that n grains make a heap and $n - 1$ do not would also violate (!!), by reasonably believing both that n grains make a heap and that $n - 1$ grains do not; it is reasonable to believe a conjunction only if it is reasonable to believe its conjuncts. Thus if one can explain (!!), one can also explain why it is not reasonable to believe that n grains make a heap and $n - 1$ do not.

In the present case, the relevant epistemically possible situations are those in which the cut-off point for 'heap' varies; no finer distinctions are relevant. Let s_k be the situation in which k is the least number of grains to make a heap. Thus n grains make a heap in s_k just in case n is at least k. Suppose that the situations within the appropriate margin for error of s_k are s_{k-2}, s_{k-1}, s_k, s_{k+1} and s_{k+2}. Suppose also that four out of five count as 'most', but that three out of five do not. Thus it is reasonable to believe a proposition in s_k if and only if it is true in at least four of s_{k-2}, s_{k-1}, s_k, s_{k+1} and s_{k+2}. Now the proposition that n grains make a heap is true in at least four of those situations if and only if n is at least $k + 1$; the proposition that $n - 1$ grains do not make a heap is true in at least four of the situations if and only if n is at most $k - 1$. Since no number is both at least $k + 1$ and at most $k - 1$, the proposition that n grains make a heap is true in most of the situations only if it is not the case that the proposition that $n - 1$ grains do not make a heap is true in most of them. Thus it is reasonable to believe the former proposition only if it is not reasonable to believe the latter. This explains (!!). Although the explanation uses highly simplified assumptions, more complex versions can be developed to cope with more complex assumptions. The underlying idea is the same.

The plausibility of (!!) depends on the assumption that a reasonable belief must have a high probability on the evidence. If a probability barely more than 50 per cent sufficed, the short step from $n - 1$ to n might be enough to tip the balance, and falsify (!!). The propositions that $n - 1$ grains do not make a heap and that n grains do could simultaneously be slightly more probable than not. The foregoing explanation therefore required more than a bare majority of the relevant situations for reasonableness.

The logic of the operator 'It is reasonable to believe that . . .' is like the logic of the operator 'It is highly probable that . . .', on the approach just outlined. Although 'It is reasonable to believe that *A*' does not entail '*A*', it does entail 'It is not reasonable to believe that not *A*'. If '*A*' entails '*B*', then 'It is reasonable to believe that *A*' entails 'It is reasonable to believe that *B*'. However, if several premises 'A_1', 'A_2', . . . , 'A_i' jointly entail '*B*', it does not follow that the premises 'It is reasonable to believe that A_1', 'It is reasonable to believe that A_2', . . . , 'It is reasonable to believe that A_i' jointly entail 'It is reasonable to believe that *B*', for a logical consequence of several propositions may be less probable than each of them individually. In particular, if A_1 and A_2 are independent propositions, then 'It is reasonable to believe that A_1' and 'It is reasonable to believe that A_2' do not jointly entail 'It is reasonable to believe that A_1 and A_2'.

The analogue of the KK principle for what it is reasonable to believe breaks down. Consider again the model used to explain (!!). If '*A*' is true just in s_{k-2}, s_{k-1}, $sk+1$ and s_{k+2}, then 'It is reasonable to believe that *A*' is true just in s_k, and 'It is reasonable to believe that it is reasonable to believe that *A*' in no situation at all. Thus the inference from 'It is reasonable to believe that *A*' to 'It is reasonable to believe that it is reasonable to believe that *A*' has a true premise and a false conclusion in s_k. In probabilistic terms: 'It is highly probable that *A*' does not entail 'It is highly probable that it is highly probable that *A*'. Inferences of the converse form can also be shown to fail. The reason is that the evidence on which the probabilities are conditional is what is known, but may not be known to be known. Thus a non-zero probability may be assigned to a possible situation in which the propositions constituting the evidence in the actual situation do not count as evidence, although they are still true.[18]

The remarks above attempt no more than a sketch. Nevertheless, they show that the epistemic view of vagueness can be extended to a variety of cognitive attitudes. It is not only our knowledge that is inexact. For vague terms, the inexactness has a conceptual source; we cannot even form reasonable beliefs as to the location of their cut-off points, and fall under the illusion that such points do not exist.

Chapter 9

Vagueness in the world

9.1 SUPERVENIENCE AND VAGUE FACTS

Words are objects; since there are incontestably vague words, there are incontestably vague objects. Yet vagueness is often said to be a feature, not of objects themselves, but of the words with which we describe them. The intended thought is Russell's: objects are vague only in their capacity as representations. Again, it is often said that the facts themselves are not vague; only our representations of them are vague. This chapter investigates the meaning and truth of such claims in the light of the understanding of vagueness so far developed.

Section 7.4 postulated the supervenience of vaguely described facts on precisely describable facts. If two possible situations are alike as precisely described in terms of physical measurements, for example, then they are alike as vaguely described with words like 'thin'. It may therefore be concluded that the facts themselves are not vague, for all the facts supervene on precisely describable facts.

Both the validity of the argument and the truth of its premise are questionable. According to Section 6.2, not even our descriptions of physical measurements can be perfectly precise. They are more precise than descriptions with words like 'thin', but the process of increasing precision will never attain its limit. We are not entitled to assume a world of precisely describable facts on which all the facts supervene. Even if we do assume such a world, moreover, we are not entitled to conclude without further argument that all the facts can be precisely described. For to say that all the facts supervene on precisely describable facts is not to say that there are only the precisely describable facts. The definition of supervenience does not require the supervenience base to be logically or metaphysically prior to what supervenes on it. Perhaps vaguely described facts cannot be precisely described, yet supervene on facts that can.

These remarks hint at reasons for which the idea of vagueness in things themselves has attracted some and repelled others. The idea attracts, because it promises to allow a rather direct relation between our vague ordinary words and the facts we use them to describe, for example, between an utterance of 'Blood is red' and the fact that the substance blood has the property of being red. The idea repels, because it promises to forbid a complete description of all the facts in precise scientific words. Opposed metaphysical proclivities underlie the ensuing debate.

The issues are evidently obscure. Section 9.2 sketches some of their ramifications as they have developed in a tradition confined by the assumption that vagueness is a kind of indeterminacy. The question 'Might reality itself be vague?' has therefore been assimilated to the question 'Might reality itself be indeterminate?'. However, previous chapters have seen the assumption not to support a cogent account of vagueness. Section 9.3 returns to the epistemic view of vagueness, and uses it to develop a modest epistemic sense in which things themselves may be vague.

9.2 DETERMINACY IN THE WORLD

The claim that only representations can be vague may be substantiated in different ways.

(a) The facts may be held, as a matter of logic or metaphysics, to be perfectly precise, so that vagueness can only be a feature of representations of the facts.

(b) To apply the concept of vagueness to anything other than a representation may be treated as a category mistake, on the grounds that vagueness is simply a mode of representing.

As they stand, (a) and (b) are incompatible positions. If it is a category mistake to apply the concept of vagueness to something other than a representation, then it must be just as much a category mistake to apply the contrasting concept of precision to that thing: but 'the facts' are supposed to be something other than representations. Thus if (b) is true, (a) commits a category mistake. Granted that facts are not representations, (a) makes ‘The facts are precise’ necessarily true and ‘The facts are vague’ necessarily false; (b) makes both of them equally meaningless. (a) and (b) agree that if there were no representations, there would be no vagueness; they disagree on whether there would still be precision.

The claim that only representations can be vague may be compared to the claim that only representations can be false. The latter claim too may be substantiated in different ways, parallel to (a) and (b). On one view, it is logically or metaphysically necessary that all facts are true facts; on the other, it makes no more sense to speak of true facts than it does to speak of false ones. The two views agree that if there were no representations, there would be no falsity; they disagree on whether there would still be truth.

Can the attitudes expressed in (a) and (b) be reconciled by reinterpretation? If one starts with a narrow concept of precision, applicable only to representations, as in (b), one can use it to define a broader concept, applicable to non-representations too. For one can define something to be precise in the broad sense if and only if all (vague or precise) descriptions of it supervene on descriptions of it that are precise in the narrow sense. Correspondingly, something is vague in a broad sense if and only if it is not precise in the broad sense. It may then be held, perhaps as a matter of logical or metaphysical necessity, that all non-representations are precise in the broad sense. Indeed, all representations as well as all non-representations may be held to be precise in the broad sense, on the grounds that a vague representation can be precisely described. According to the generalized thesis, whatever can be described at all can in principle be described with perfect precision.

Michael Dummett once expressed a view similar to (b): ‘the notion that things might actually *be* vague, as well as being vaguely described, is not properly intelligible’.[1] However, he subsequently withdrew the remark, as lacking visible means of support, and suggested that the generalized thesis might express no more than a deep-rooted prejudice.[2] Certainly, the definition of a broad concept of vagueness on the basis of (b) carries no commitment to the generalized thesis; it merely renders that thesis statable.

Section 9.1 casts doubt on attempts to capture the metaphysical issues with a broad concept of vagueness defined as above. Suppose that no representation can be perfectly precise. We can still replace vague representations by less vague ones. Consider someone who is tall because he is 6 ft 1 in. in height. It does not seem completely senseless to ask whether the vaguer representation 'He is tall' is made true only by the very fact that makes the less vague representation 'He is 6 ft 1 in. in height' true, or whether they are made true by different facts. Even if representations can be perfectly precise, or at least precise to within any required degree short of perfection, the same question can be asked. Perhaps the notion of a fact is suspect, but similar questions can be asked about objects, properties and relations. One can ask whether the vaguer representation represents (albeit vaguely) the very objects, properties and relations represented by the less vague representation, or whether they represent different objects, properties and relations. The underlying question is whether there is an important sense in which what vague representations represent is itself vague; it is not answered by the possibility or impossibility of other, more precise representations.

The question takes at least one traditional form. Suppose that the world contains properties and relations that are just as independent of our representations of them by concepts as are the objects which have those properties and relations. Call such properties and relations 'universals', and call having such a property or relation 'participating' in that universal. Must participating in a universal be an all-or-nothing matter, or can it be a matter of degree? In arguing that only representations can be vague, Russell gave a positive answer to the former question: 'Nothing is more or less what it is, or to a certain extent possessed of the properties which it possesses'.[3] The remark was perhaps directed against a line of neo-Hegelian thinkers for whom truth and reality came essentially in degrees, although those thinkers did not single out vagueness for special attention. They were concerned with the incompleteness of all our descriptions, which they (no doubt wrongly) took to imply that the descriptions could not be completely true.[4] From a quite different starting point, Arthur Burks argued in 1946 that 'an empiricist cannot consistently hold to the view that all universals embodied in the real world are precise universals'.[5] He took this to be a consideration against the view that all universals embodied in the real world are precise, rather than against empiricism. Burks explicitly raised the possibility that our vague concepts might represent vague universals, participation in which is a matter of degree.

One might replace the locution '*x* has property *P*' by the locution '*x* has property *P* to degree *d*'. However, one has gained little if what the latter expresses is a matter of degree too. If not, such vagueness in the world is of a mild kind: there would be no corresponding higher-order vagueness in the world. It would be tempting to attribute even the first-order vagueness to mere inexplicitness in our ordinary forms of speech. In contrast, if possession of a given property to a given degree is itself a matter of degree, and so *ad infinitum*, then the vagueness seems to run deeper.

In Sections 4.11 and 5.5, vagueness turned out to be much less closely related to matters of degree than is often supposed. However, one may ask related questions. Can it be indeterminate whether a given object has a given property? If so, can it be indeterminate whether it is indeterminate whether the object has the property?

Participation in one universal of particular importance has been held to be indeterminate through vagueness: the identity relation.[6] Derek Parfit, for example, took it to be uncontroversial that a pair of nations, or of machines, might be neither determinately identical nor determinately distinct. There is a nation *x*; various historical changes occur, at the end of which there is a nation *y*. Our criteria of identity for nations may give no decision as to whether *x* is the same nation as *y*. According to Parfit, the identity question 'Is *x y*?' has no right answer in such a case. Similarly, our criteria of identity for persons may give no decision as to whether I shall be the person who results from some sufficiently bizarre permutation of my brain and limbs with those of others. Parfit argued that the identity question 'Is the person who will result me?' is equally without a right answer in such a case, even though we find the conclusion far more disturbing in our own case than in that of nations or of machines. If I care very much whether the person who will result is me, it does not follow that it is determinate whether the person who will result is me.[7] In a similar vein, Saul Kripke wrote of cases where 'the identity relation is vague': for example, where parts of a table are replaced.[8] On such views, one may single out a particular object x, and a particular object y, yet it may be just vague whether x is y.

The coherence of such speculations was called into question by Gareth Evans's brief note of 1978, 'Can there be vague objects?'. Not long afterwards, Nathan Salmon independently advanced a similar argument.[9] Although different versions vary in detail, the main idea is simple. Suppose that it is indeterminate whether $x = y$. Thus it is not determinate that $x = y$. But it is determinate that $x = x$. So x has a property that y lacks: the property one has in being such that it is determinate that x = one. Now identity is governed by Leibniz's law: if $x = y$ then every property of x is a property of y. Thus it is not the case that $x = y$. This seems to undercut the initial supposition that it is indeterminate whether $x = y$. We began by assuming the question 'Is x y?' to have no right answer, but this assumption seems to yield an answer to the question: 'No'. How then can it be indeterminate whether $x = y$?

The Evans–Salmon argument has provoked a large body of discussion. Evans began his note by mentioning the view that 'the world might itself *be* vague'. He did not claim it to be a consequence of this view that matters of identity in particular are vague; someone might hold that the world is vague in some respects, but not in respect of matters of identity. However, Evans's choice of title does suggest the claim that x is a vague object if and only if there is an object y such that it is indeterminate whether $x = y$. Moreover, those who think that the world is vague may do so *because* they think that some matters of identity are vague. Thus discussion of whether the world may be vague has come to centre on the Evans–Salmon argument. The specific proposition that for some objects x and y it is indeterminate whether $x = y$ is rather more tractable than the general proposition that the world is vague, and the formal structure of the argument is something definite to work on.

One point soon became clear. The argument does not show, and was not intended to show, that identity *statements* cannot be indeterminate in truth-value. For it leaves open the possibility of indeterminacy in whether two names refer to the same object. Arguably, such a situation can be set up if any statement at all is indeterminate in truth-value. For if it is indeterminate whether A, we can stipulate that the name 'Pardon' refers to London unless A, in which case it refers to Paris. But then the identity statement 'Paris =

Pardon' will be indeterminate in truth-value, for it is equivalent to '*A*'. What happens if one tries to use the Evans–Salmon argument against this claim? Granted, it is not determinate that Paris = Pardon and it is determinate that Paris = Paris. The next step would be to infer that Paris has a property that Pardon lacks: the property of being an object *z* such that it is determinate that Paris = *z*. This is to argue from 'It is not determinate that Paris = Pardon' to 'Pardon is not an object *z* such that it is determinate that Paris = *z*'. However, this inference is invalidated by the indeterminacy in reference of 'Pardon'. For if Paris = Pardon, Pardon *is* an object *z* such that it is determinate that Paris = *z*.[10] Thus if 'Paris = Pardon' is indeterminate in truth-value, 'Pardon is an object *z* such that it is determinate that Paris = *z*' is not false. Hence 'It is not determinate that Paris = Pardon' might be true while 'Pardon is not an object *z* such that it is determinate that Paris = *z*' was indeterminate in truth-value. At any rate, Evans says nothing to block such an objection. Rather, the argument should be advanced only on the assumption that the singular terms flanking '=' are determinate in reference. Most defenders of vague objects would grant that assumption; their idea is indeed that vagueness in the objects referred to sometimes explains an indeterminacy in truth-value of an identity statement better than does indeterminacy in reference in its singular terms.

A defender of vague objects may in desperation suggest that we cannot single them out determinately, and so cannot name them. The argument is therefore best formulated in terms of variables rather than names. Salmon did this, and it was done above. For any pair of objects, vague or precise, unspecifiable or specifiable, there is an assignment of the first to the variable 'x' and the second to the variable 'y', which is just what the argument needs.

Evans and Salmon aimed their arguments specifically against indeterminate identity, taken as implicit in vague identity. However, neither of them specified an overall framework for the handling of indeterminacy. A supervaluational treatment is as congenial to the argument as any.[11] Their critics have tended to use something like a three-valued logic, with 'true', 'false' and 'indeterminate' as the three values. For example, suppose that 'It is determinate that *A*' is true if '*A*' is true, and false if '*A*' is indeterminate or false. Then if '$x = y$' is indeterminate, 'It is determinate that $x = y$' will be false and 'It is determinate that $x = x$' true. Thus '$x = y$' in combination with a true premise leads by Leibniz's law to a false conclusion. On classical assumptions, it follows that '$x = y$' is false: not so in some non-classical systems. What leads in combination with true premises by truth-preserving reasoning to a false conclusion need not be false in a three-valued logic; it may just be indeterminate. Alternatively, more or less *ad hoc* restrictions may be imposed on Leibniz's law in the presence of indeterminacy. There is no doubt that a variety of non-standard systems can be specified in which the Evans–Salmon argument is formally incorrect. The same goes for any argument, valid or invalid. What matters is whether any such system provides us with a correct set of standards for evaluating the argument on its intended interpretation. Here the objections have been found wanting. If they supply an overall semantic framework at all, it is that of many-valued logic: but, as reviewed in Sections 4.11–4.14, there are well-known and powerful reasons to reject any such framework. Those reasons are quite independent of issues about identity; the objectors have done nothing to rebut them. Nor, in the case at hand, have they said enough about what indeterminacy is to dispel the suspicion that for '$x = y$' to be 'indeterminate' is just another way for it to be false.

If the Evans–Salmon argument is correct, it does not follow that the world itself is in no respect vague. Identity is one relation amongst many. Evans writes of 'the idea that the world might contain certain objects about which it is a *fact* that they have fuzzy boundaries'. Yet fuzzy boundaries do not in any obvious way require vague identity. Objects are identical only if their boundaries have exactly the same fuzziness. If Europe has fuzzy boundaries, some points are neither determinately in Europe nor determinately not in it. We might draw a precise closed curve, including all points determinately in Europe and excluding all those determinately not in it, and treat it as the boundary of a sharp object, Europe*. It does not follow that 'Europe = Europe*' is a case of vague identity. One might hold that $x = y$ only if every point is determinately in x just in case it is determinately in y, and determinately not in x just in case it is determinately not in y. Consider a point neither determinately in Europe nor determinately not in it. Since the boundary of Europe* is not fuzzy, the point is either determinately in Europe* or determinately not in it. Thus a necessary condition for Europe to be Europe* fails, and may do so determinately; 'Europe = Europe*' is simply false. On this view, 'Europe' determinately refers to a vague region, rather than being indeterminate in reference between a number of precise regions. Such a treatment is not obligatory, but the point is that it is quite consistent with the Evans–Salmon argument. Moreover, it has the convenience of allowing 'Europe' to be treated as straightforwardly referring to a particular object, Europe.[12]

If objects can have fuzzy spatial boundaries, surely they can have fuzzy temporal, modal or mereological boundaries too. When did Europe begin to exist? Under what counterfactual circumstances would it still have existed? Which cities are part of Europe? Each of these questions might be taken to concern a particular object, Europe, yet to have no determinate answer. In no case does this view entail the possibility of a pair of objects concerning which it is indeterminate whether they are identical.

Suppose that it is vague, concerning Moscow and Europe, whether the former is part of the latter. One might then conclude that parthood is a vague relation, even if identity is not. This reopens the general question: when is a universal vague? One might try saying that a property P is vague just in case there is (or could be) an object x such that it is vague whether x has P, and that a binary relation R is vague just in case there are (or could be) objects x and y such that it is vague whether x has R to y. However, there is a difficulty with such definitions. For example, concerning Europe and Europe*, it is vague whether the former has a larger surface area than the latter. Yet this does not seem to justify calling the relation of having a larger surface area vague; for present purposes it seems quite a precise relation. Intuitively, the vagueness in whether one object has the relation to the other is in this case to be blamed on vagueness in one of the objects, not on vagueness in the relation. Similarly, suppose that the surface area of Europe* is exactly h hectares. Then, concerning Europe, it is vague whether it occupies more than h hectares. Yet this does not seem to justify calling the property of occupying more than h hectares vague; for present purposes it seems quite a precise property. Intuitively, the vagueness in whether the object has the property is in this case to be blamed on vagueness in the object, not on vagueness in the property.[13]

One response to the problem has been to say that a property P is vague just in case there is (or could be) a precisely bounded object x such that it is vague whether x has P, and that a binary relation R is vague just in case there are (or could be) precisely bounded objects x and y such that it is vague whether *x* has *R* to *y*. That would deal with the problem in the last paragraph, for Europe is not a precisely bounded object. However, whereas the previous definition is too permissive, the present one is too restrictive. Consider the property *P* of being a very large object with very vague boundaries. If *x* is a precisely bounded object, then it is not vague whether *x* has *P*; it is clearly determinate that *x* lacks *P*. Thus, by the present definition, *P* is not a vague property: yet, intuitively, it is very vague. A different account is needed of what it is for a universal to be vague. Such an account will be provided by appeal to the epistemic view of vagueness.

9.3 UNCLARITY *DE RE*

The epistemic view of vagueness might be supposed to answer the question 'Can things themselves be vague?' with an immediate negative. There are worse and better reasons for that supposition.

A bad reason is this. Someone who takes it for granted that vagueness is a form of indeterminacy may take the epistemic view to deny vagueness, simply because it denies the relevant indeterminacy. If vagueness does not occur at all, then it does not occur in things themselves. That line of thought rests on a misinterpretation already corrected. The epistemic view allows that vagueness occurs, merely denying that it is a kind of indeterminacy.

There is a better reason. To attribute vagueness to 'things themselves' just is to say that they have it irrespective of whether or how they are represented. But if vagueness is a matter of ignorance, it depends on whether or how things are represented to the supposedly unknowing subject. No representation: no distinction between vague and precise. Thus the epistemic view cannot consistently attribute vagueness to things themselves.

The better reason can be elaborated. According to Section 8.4, vagueness manifests a specific kind of ignorance: that which stems from the indiscriminability of distinct concepts. Indiscriminability depends on the cognitive standpoint of the would-be discriminator; the relevant standpoint here is that of the person who grasps the concepts in question. The notion of grasping is specific to concepts; in that sense there is no such thing as grasping a non-concept. Since a concept is vague to the extent that it would be indiscriminable from another concept by those who grasped them, nothing but a concept can be vague in the same sense. Similarly, a linguistic expression is vague to the extent that it would be indiscriminable in meaning from a non-synonymous linguistic expression by those who understood them; the relevant notion of understanding is specific to representations. Nothing but a representation could be vague in that sense.

Strictly understood, the distinction between vagueness and precision applies only to representations. But that does not rule out the possibility that it reflects a corresponding distinction in what is represented. Might vagueness and precision correspond to features not confined to representations, vagueness* and precision*, what a vague representation represents being vague* and what a precise representation represents precise*? The correspondence will not be exact, if the same thing can be represented both vaguely and precisely, for then it will count as both vague* and precise*, whereas a representation is vague only to the extent that it is not precise. For example, the vague description 'the greatest natural number much less than 100' and the precise description 'the prime number between 72 and 78' might both refer to 73 in a given context. Thus 73 would be both vague* and precise*.[14] This example suggests that, on the epistemic view, the possibility of representing something vaguely says little about the nature of that thing; whatever can be represented can be represented vaguely. Just about everything is vague*. Perhaps precision* is a more interesting feature. It is less plausible that whatever can be represented can be represented precisely; the possibility of representing something precisely might say much about the nature of that thing.

If something is precise* only if it can be represented with *perfect* precision, then one may suspect that almost nothing is precise*. However, even if something cannot be represented with perfect precision, it may still be capable of being represented with *unlimited* precision, in the sense that, for every degree of precision less than perfection, it can be represented more precisely than that. We need not rely on this idea of degrees of precision, for which we have no measure, since we can put the point in terms of clarity. Even if it is impossible to remove all unclarities in the representation of *x*, it may still be possible to remove any particular unclarity.[15] One may express such an unclarity by saying that it is unclear whether *a* is *F*, where '*a*' refers to *x*; the point is that it may be clear whether *b* is *F*, where '*b*' also refers to *x*. This formulation has the advantage of not using the cloudy term 'representation'.

We can use the operator 'it is unclear whether' to focus the issues. Suppose that vagueness in '*a*' makes it unclear whether *a* is *F*; the problem is whether this reflects something about what '*a*' refers to, *x*, or only something about '*a*' (given that '*F*' is relevantly precise). We can pose it by asking a question in which '*a*' is not used: is it unclear, of *x*, whether it is *F*? In traditional terminology, we need to use a construction *de re*, as in that question, rather than a construction de *dicto*, as in 'Is it unclear whether *a* is *F*?'. The distinction between constructions *de re* and *de dicto* will be used to analyse the notion of vagueness in things themselves. We may begin by rehearsing the general distinction (and ignoring many subtleties).

Syntactically, the distinction between constructions *de re* and *de dicto* may be drawn for any sentence functor. In constructions *de dicto*, a term occurs within the scope of the functor, as the term 'this man' occurs within the scope of the functor 'It is clear that . . .' in the sentence 'It is clear that this man is tall'. In constructions *de re,* the term occurs outside the scope of the functor but binds a pronoun or the like within that scope, as 'this man' binds 'he' in 'It is clear of this man that he is tall'.[16] The semantic significance of this syntactic distinction may be illustrated in the case of the context 'He knows that . . .'. Imagine someone who does not know that Constantinople fell in 1453; he knows only that it fell after a great siege some time in the fifteenth century. Thus he does not know that the year Constantinople fell was before 1460. However, he does know that 1453 was before 1460. Indeed, he knows of 1453 that it was before 1460. Since 1453 is the year Constantinople fell, he knows of the year Constantinople fell that it was before 1460. Thus 'He knows of the year Constantinople fell that it was before 1460' (*de re*) does not entail 'He knows that the year Constantinople fell was before 1460' (*de dicto*). The converse entailment can also fail. He knows that there was a great siege in the year Constantinople fell without knowing that there was a great siege in 1453, indeed, without knowing of 1453 that there was a great siege in it. Since

1453 is the year Constantinople fell, he does not know of the year Constantinople fell that there was a great siege in it. Thus 'He knows that there was a great siege in the year Constantinople fell' (*de dicto*) does not entail 'He knows of the year Constantinople fell that there was a great siege in it' (*de re*).[17]

Several principles were tacitly at work in the previous paragraph. It was assumed that, since '1453' and 'the year Constantinople fell' refer to the same year, the sentences *de re* 'He knows of 1453 that it was before 1460' and 'He knows of the year Constantinople fell that it was before 1460' have the same truth-value. In contrast, the sentences *de dicto* 'He knows that 1453 was before 1460' and 'He knows that the year Constantinople fell was before 1460' have different truth-values. The terms in the constructions *de dicto* occur within the subordinate clauses, and specify how someone must denote the year to have the attributed knowledge; thus the substitution of one term for another with the same denotation can alter the truth-value of the sentence. The singular terms in the constructions *de re* occur outside the subordinate clause, and specify only which year someone must denote to have the attributed knowledge; thus the substitution of one term for another with the same denotation cannot alter the truth-value of the sentence. A second principle was also at work. It was assumed that, in the circumstances, 'He knows that 1453 was before 1460' (*de dicto*) is sufficient for 'He knows of 1453 that it was before 1460' (*de re*), but that 'He knows that there was a great siege in the year Constantinople fell' (*de dicto*) is not sufficient for 'He knows of the year Constantinople fell that there was a great siege in it' (*de re*). There is an asymmetry between the terms '1453' and 'the year Constantinople fell'. If one understands '1453', one knows which year it denotes. If one understands 'the year Constantinople fell' (as the man in the example does), one need not know which year it denotes. To move from the report *de dicto* to the report *de re*, one must assume that the subject of the knowledge knows which thing is in question.[18]The date is associated with a way of thinking of the year that enables one to have thoughts *de re*; the description is not.

The *de re*/*de dicto* distinction may be transferred from reports of attitudes to the attitudes themselves. An attitude is *de re* if it is correctly reportable *de re*; it is *de dicto* if it is correctly reportable *de dicto*. Thus the very same attitude may be both *de dicto* and *de re*. For example, our man's knowledge that 1453 was before 1460 amounts to his knowledge of 1453 that it was before 1460.[19]

How does the *de re/de dicto* distinction fit the construction relevant to vagueness, 'It is unclear whether . . .'? Three disanalogies between it and more frequently discussed constructions such as 'He knows that . . .' may be noticed.

(a) 'Whether' is used in place of 'that'. However, it is still followed by a subordinate clause. Moreover, 'It is unclear whether *A*' may be analysed as 'It is unclear that *A* and it is unclear that not *A*' (where 'It is unclear that . . .' does not entail ' . . .'). Thus if *the de re/de dicto* distinction applies to 'It is unclear that . . .', then it also applies to 'It is unclear whether . . .'.

(b) The impersonal 'It is unclear' is used in place of the personal 'He knows'. However, it is still a cognitive notion, just as the impersonal 'It is known' is. 'It is unclear' abstracts from the particular circumstances of particular thinkers, but it adverts to what can or cannot be unclear to some thinkers in some circumstances. The *de re/de dicto* distinction can be drawn for such attitudes. We dig up the fossilized remains of what was in fact the last brontosaurus, but we are in no position to know that it was the last; it is known of the last brontosaurus that it died here, but it is not known that the last brontosaurus died here.

(c) 'It is unclear' looks like the negation of 'It is clear', whereas 'He knows' is not the negation of anything. This may be significant; if a propositional attitude is something like knowledge or belief, hope or fear, there is a sense in which the negation of a propositional attitude is not itself a propositional attitude. To know, believe, hope, or fear that *A*, one must grasp the thought that *A*. In contrast, to fail to know, believe, hope, or fear that *A*, one need not grasp the thought that *A*. Indeed, one may fail to take those attitudes just because one fails to grasp the thought (one may lack the relevant concepts). However, it is not evident that 'It is unclear' is the negation of 'It is clear'. Consider first the personal case. If someone has never heard of hydrogen, the statement 'It is clear to him whether water contains hydrogen' is obviously false, but the statement 'It is unclear to him

> whether water contains hydrogen' is not obviously true. There is a sense in which the matter becomes unclear to him only once he is in a position to consider it. The case *de re* is similar: there is a sense in which it is unclear to him of x that it is *F* only if he has a way of thinking *de re* of *x*.

Does the distinction between unclarity and lack of clarity survive the abstraction from the personal case to the impersonal one relevant to vagueness, or does it depend on the contingent cognitive limitations of particular people in particular circumstances? The distinction *de dicto* between 'It is unclear whether *A*' and 'It is not clear whether *A*' may disappear, for it depended on the person's failure to grasp the thought that *A*. If it makes sense to apply the operator 'It is unclear whether' to '*A*', the thought that *A* can be articulated, and so grasped by someone. However, an impersonal distinction between unclarity *de re* and lack of clarity *de re* still makes sense, if there could be things incapable of being singled out in thought.

Space–time points may constitute an example. If such points constitute a dense array, perhaps the smallest part of space–time that can be individually thought of is a small region rather than a single point. Of course, we can think about such points in general, and have excellent evidence that they exist; our best scientific theories may tell us that they form a dense array. What might be impossible for any creature is to refer to just one of those points and think about it in particular. On this view, no point can be named; something is wrong with the Fregean doctrine that an object is a potential bearer of a proper name. Let *x* be a point ('*x*' is of course a variable, not a name). Since it is impossible to think about *x*, it is not clear of *x* whether it is on Earth. Yet there is a sense in which it is also not unclear of *x* whether it is on Earth. For example, it could never occur to anyone to doubt of *x* whether it is on Earth. In this sense, impersonal unclarity *de re* is not simply the absence of impersonal clarity *de re*.

The sense of 'It is unclear of *x*' in which it is not the negation of 'It is clear of *x*' will be adopted in what follows. The expressive power of the terminology is thereby enlarged, for the negation of 'It is clear of *x*' can always be expressed as such. The more expressive terminology provides a more nuanced description of vagueness *de re*. It is clear of *x* whether it is *F* if and only if, for some way of thinking *de re* of *x* expressed by '*a*', it is clear whether *a* is *F*. It is unclear of *x* whether it is *F* if and only if, for some way of thinking *de re* of *x* expressed by '*a*', it is not clear whether *a* is *F*. The same principles hold with 'that' in place of 'whether'. The usual equivalences between the two forms also hold: it is clear of *x* whether it is *F* if and only if either it is clear of *x* that it is *F* or it is clear of *x* that it is not *F*; it is unclear of *x* whether it is *F* if and only if it is unclear of *x* that it is *F* and it is unclear of *x* that it is not *F*.[20] As already noted, something can be neither clear nor unclear: if there is no way of thinking of *x*, it is neither clear of *x* whether it is *F* nor unclear of *x* whether it is *F*. Equally, it has not been ruled out that something may be both clear and unclear. If '*a*' and '*b*' express different ways of thinking *de re* of *x*, perhaps it is clear whether *a* is *F* and not clear whether *b* is *F*. Then it is both clear of *x* whether it is *F* and unclear of *x* whether it is *F*.[21]

The apparatus is now ready to be applied to an example. Suppose that it is unclear whether the river Thames is more than 209 miles long, because it is unclear where exactly it meets the sea. The epistemic view of vagueness is consistent with the commonsense view that many people have seen and heard about the Thames, know which river it is, and sometimes think *de re* about it. From the supposition and those views it follows that it is unclear of the Thames whether it is more than 209 miles long. There is unclarity *de re*. Since the phrase 'more than 209 miles long' has no relevant vagueness, in a modest sense the Thames itself is vague. This view is to be contrasted with one on which the expression 'the Thames' is indeterminate in reference between a number of precisely defined stretches of water. The epistemic view can preserve the straightforward reference of ordinary singular terms.

What has not been shown is that, on the epistemic view, it is not clear of the Thames whether it is more than 209 miles long. That would involve showing that for *no* way of thinking *de re* of the Thames expressed by '*a*' is it clear that *a* is more than 209 miles long. What might be such a way? One naturally thinks of a precise specification of spatial boundaries. However, more would need to be specified, for more than one thing has those spatial boundaries. At this moment the Thames is constituted by a particular body of water, with the same boundaries in space as it: but soon part of that body of water will be part of the sea and not part of the Thames any more. A river is not just a body of water. Nor is it enough to specify boundaries in time as well as space, for the event consisting of everything that ever happened or will happen in the Thames (its 'life') has the same boundaries as the river in space and time; but whereas the Thames could have had very different boundaries, if it had taken a different course or dried up years ago, that event could not have had those very different boundaries (although some other event could have). A river is not just an event. One would need to specify counterfactual as well as actual boundaries. It is hard to see how that could be done except by something like the specification that the boundaries in question are those of a *river*. Thus '*a*' must mean something like: that river with such-and-such boundaries. But how does one know that *any* river has those boundaries? If no river has them, then '*a*' does not refer. There are two cases to consider.

(i) It is clear that *a* exists. Now if '*a*' is defined in anything like the way envisaged, it is clear that if *a* exists then *a* is a river. Indeed, it is clear that if *a* exists then *a* is the river Thames; no other river has anything like those boundaries. Since clarity *de dicto* is preserved by logical deduction, it is clear that a is the Thames. By hypothesis, it is either clear that a is more than 209 miles long or clear that a is not more than 209 miles long. But if a is the Thames, then it logically follows that a is more than 209 miles long if and only if the Thames is more than 209 miles long. Again because clarity *de dicto* is preserved by logical deduction, it is either clear that the Thames is more than 209 miles long or clear that the Thames is not more than 209 miles long. But the original hypothesis was just that it is not clear whether the Thames is more than 209 miles long. Thus if the vagueness is genuinely irreducible, case (i) will not occur.

(ii) It is not clear that *a* exists. Thus '*a*' means something like: the river (if there is one) with such-and-such boundaries. But then it is not plausible that '*a*' expresses a way of thinking *de re* of the Thames, any more than 'the year Constantinople fell (if it did fall)' expresses a way of thinking *de re* of 1453. Such an '*a*' does not make it clear of the Thames whether it is more than 209 miles long.

In the light of (i) and (ii), it seems to be not clear, as well as unclear, of the Thames whether it is more than 209 miles long. If so, the Thames itself is vague in a less modest sense than that already acknowledged. The argument should be treated with caution, for it is hard to survey all possible definitions of '*a*'. What has become evident, however, is that the epistemic view of vagueness is at the very least compatible with the occurrence of unclarity and lack of clarity *de re*.

Examples could easily be multiplied. Statements of identity and non-identity may be singled out for discussion. Strictly speaking, notions *de re* have not yet been explained for the case of more than one term. The natural generalization is as follows. It is clear of *x* and *y* whether the former is *R* to the latter if and only if, for some way of thinking *de re* of *x* expressed by '*a*' and some way of thinking *de re* of *y* expressed by '*b*', it is clear whether *a* is *R* to *b*. It is unclear of *x* and *y* whether the former is *R* to the latter if and only if, for some way of thinking *de re* of *x* expressed by '*a*' and some way of thinking *de re* of *y* expressed by '*b*', it is not clear whether *a* is *R* to *b*. The same principles hold with 'that' in place of 'whether'. The generalization to more than two terms is obvious.

Suppose that rivers 1 and 2 join; downstream there is river 3; it is unclear whether river 2 is a tributary of river 1. It is correspondingly unclear whether river 1 = river 3. The terms 'river 1', 'river 2' and 'river 3' may be taken to express ways of thinking *de re* of the rivers, perhaps given by perception. It is therefore unclear of river 1 and river 3 whether the former = the latter.

What of the Evans–Salmon argument? Surely it remains open that river 1 = river 3; but how can that identity be unclear? Well, suppose that river 1 = river 3. It is clear that river 1 = river 1. We can take both '*a*' and '*b*' in the above account to be 'river 1', *x* to be river 1 and *y* to be river 3, for since 'river 1' expresses a way of thinking *de re* of river 1, which is river 3, it also expresses a way of thinking *de re* of river 3. Thus it is clear of river 1 and river 3 that the former = the latter; this is a valid form of the Evans–Salmon argument. By the assertion at the end of the previous paragraph, it is unclear of river 1 and river 3 that the former = the latter. However, this is no contradiction, for clarity *de re* and unclarity *de re* have already been seen not to be incompatible. The identity is clear under one pair of ways of thinking of the river, not under another.

Careful notation defuses another apparent contradiction. Since it is unclear of river 1 and river 3 that the former = the latter, it is unclear of river 1 and river 1 that the former = the latter. Yet it cannot be unclear of river 1 that it = it. For, by the explanation of unclarity de re, that would imply that for some way of thinking de re of river 1, expressed by 'a', it is not clear that a = a; no such 'a' exists. There is no contradiction, however, for 'it is unclear of river 1 and river 1 that the former = the latter' is not just a pedantic reformulation of 'it is unclear of river 1 that it = it'. The former permits river 1 to be thought of in distinct ways in the thought of identity; the latter does not.[22]

Thought of identity can be unclear de re; in that modest sense, identity can be vague. Can thought of identity also fail to be clear de re? It was noted above that if river 1 = river 3, then it is clear of river 1 and river 3 that the former = the latter; no clarity de re is lacking there. The substantive question is: can false thought of identity fail to be clear de re?

Suppose that river 1 ≠ river 3. Must there be a way of thinking de re of river 1, expressed by 'a', and a way of thinking de re of river 3, expressed by 'b', for which it is clear that a ≠ b? That will depend on the facts. If river 1 is a tributary of river 2 rather than vice versa, then river 2 = river 3, and we can take 'a' to be 'river 1' and 'b' to be 'river 2', for it is clear that river 1 ≠ river 2. Then it is clear of river 1 and river 3 whether the former = the latter. On the other hand, it could be that rivers 1, 2 and 3 are three distinct rivers: two rivers flow together into a third. In that case, no such 'a' and 'b' may be possible. Then it is not clear of river 1 and river 3 whether the former = the latter. In a less modest sense than that already acknowledged, identity can be vague. One could argue for this conclusion very much as it was argued that it may not be clear of the Thames whether it is more than 209 miles long. Such an argument would be as tentative as before. Nevertheless, it has become evident that the epistemic view of vagueness is at the very least compatible with the occurrence of unclarity and lack of clarity de re in identity thoughts. Similar issues will arise for other relations: 'part of', 'located at', and so on.

It has been crucial to the interest of the examples discussed so far that the constituents with respect to which they are *de re* are those to which the relevant vagueness is confined. Consider an example in which it occurs in the predicate rather than the subject. If it is unclear whether Stac Polly is a mountain (not just a rocky hill), it may well be unclear of Stac Polly whether it is a mountain, but it does not follow that reality itself is vague, for the relevantly vague term is 'mountain', which remains firmly within the scope of the functor 'it is unclear whether . . .'. Nevertheless, it might be suggested that this too is a case of vagueness *de re*, because it is unclear of the mountains whether Stac Polly is one of them, or because it is unclear of the property of mountainhood whether Stac Polly has it. It might also be suggested that it is unclear of the state of affairs of Stac Polly being a mountain whether it obtains. The remainder of this chapter explores these attempted extensions of the idea of unclarity *de re*. Not surprisingly, they are more dubious than the cases considered so far.

It is unclear of the mountains whether Stac Polly is one of them just in case for some way of thinking *de rebus* of the mountains expressed by '*F*s', it is not clear that Stac Polly is one of the *F*s. What is required for thinking *de rebus* of the mountains? On one view, all one needs is a term true of all and only the mountains. Since 'mountain' is such a term, and it is unclear whether Stac Polly is a mountain, it will indeed be unclear of the mountains whether Stac Polly is one of them. However, it will also be clear of the mountains whether Stac Polly is one of them, since for many other terms '*F*' true of all and only the mountains it is clear whether Stac Polly is one of the *F*s. For example, if Stac Polly is a mountain, '*F*' can be 'thing which is either a mountain or Stac Polly'. If it is not a mountain, '*F*' can be 'thing which is both a mountain and not Stac Polly'. However, it is not obvious that every term true of all and only the mountains does give one a way of thinking *de rebus* of the mountains. We understand the term 'mountain', but do we know which the mountains are, if we cannot decide whether Stac Polly is one of them? If we must know which the mountains are in order to think of them, there may be no room for unclarity of Stac Polly whether it is a mountain. On both views of what it takes to think *de rebus* of the mountains, the danger is that the formulations *de rebus* are too indiscriminate to be useful: the former view makes thinking *de rebus* of the mountains too easy, while the latter makes it too difficult. The argument is quite general: it did not turn on any special feature of the example. Similar remarks apply to the idea that it is unclear of the set of mountains whether Stac Polly is a member of it.

One alternative is to treat predicates as referring to properties and relations rather than to pluralities or sets. If Stac Polly is not a mountain, then the mountains other than Stac Polly and the mountains are the same things, but being a mountain other than Stac Polly and being a mountain are not the same property, since Stac Polly could have had the latter without having the former. The argument that threatened to trivialize the first view of *de re* thinking about pluralities or sets does not apply to properties and relations. The property of being *F* is the property of being *G* only if it is necessary (not just true) that everything is *F* if and only if *G*.

What is it to think *de re* of a property? On one view, in order to think *de re* of a property, it suffices to understand a predicate that, necessarily, is true of all and only things with that property. Since we understand the predicate 'is a mountain', we think *de re* of the property of mountainhood. It follows that if it is unclear whether Stac Polly is a mountain, then it is unclear of the property of mountainhood whether Stac Polly has it. However, the above view of what is to think *de re* of a property arguably also makes it also clear of the property of mountainhood whether Stac Polly has it.[23] The danger is again triviality: perhaps *de re* thought about properties has been made too easy. The argument is not decisive; perhaps a better theory of *de re* thought of universals can be developed.

With the idea of unclarity *de re* concerning properties and relations stands or falls the idea of unclarity *de re* concerning facts or states of affairs. An object having a property may be held to constitute a state of affairs, such as the state of affairs of the object Stac Polly having the property of mountainhood. On this view, states of affairs are no more representations than objects and properties are. It is unclear of the state of affairs whether it obtains if and only if it is unclear of the object and the property whether the former has the latter. In such a formulation, every substantial constituent has been evacuated from the subordinate clause, leaving only a trace behind. Any unclarity is then manifestly independent of the way in which the state of affairs was represented in the sentence in question, although not of the ways in which it could be represented in some sentence or other.

The nominalist suspects that properties, relations and states of affairs are mere projections onto the world of our forms of speech. One source of the suspicion is a sense that we could just as well have classified things differently. Vagueness is indeed one manifestation of the fact that our classifications are not fixed by natural boundaries. The vagueness of singular terms suggests that if the nominalist conclusion did follow, it could not exempt the category of objects. The boundaries of a particular mountain reflect our language no less than do those of the property of mountainhood. However, it would be an obvious mistake to conclude that our language created the mountain; it is a less obvious mistake to conclude that it created the property. If a subtler form of dependence is at stake, it is not obvious what it is.

The metaphysical issues remain to be resolved. Even so, one can use the rough and variable notion of thought *de re* to point the moral that, on the epistemic view of vagueness, our contact with the world is as direct in vague thought as it is in any thought. The cause of our ignorance is conceptual; its object is the world.

Appendix
The logic of clarity

The phenomena of vagueness can be described by means of the sentence functor 'clearly' ('it is clearly the case that'). It will be symbolized as C and added to a propositional language with sentence variables p, p', p'', . . . negation (~) and conjunction ($\wedge$), in terms of which the material conditional ($\supset$) and biconditional ($\equiv$) are defined. The predicate logic of clarity will not be discussed; see Section 9.3 for a sketch of some complex issues raised by its semantics. The formal semantics of the propositional language will use the notion of a margin for error developed in Sections 8.2–8.6. It will not attempt to capture all aspects of that discussion.

A *fixed margin model* is a quadruple $<W, d, \alpha, []>$, where W is a set, d a metric on W, α a non-negative real number, and [] a mapping of formulas to subsets of W such that for all formulas A, B:

$$[\sim A] \quad = \quad W - [A]$$

$$[A \wedge B] \quad = \quad [A] \cap [B]$$

$$[\mathrm{C}A] \quad = \quad \{w \in \mathrm{W}\colon \forall x \in W\, d(w, x) \leqslant \alpha \;\Rightarrow x \in [A]\}$$

(d is a metric on W if and only if for all $x, y, z \in W$, $d(x, y)$ is a real number; $d(x, y) \geqslant 0$; $d(x, y) = 0$ just in case $x = y$; $d(x, y) = d(y, x)$; $d(x, z) \leqslant d(x, y) + d(y, z)$). Intuitively, W is a set of possible worlds, d a measure of their similarity, α a margin for error and $[A]$ is the set of worlds at which A is true. CA is true at a world w if A is true at every world within the margin for error α of w. A formula A is *valid* in fixed margin models if and only if for every such model $<W, d, \alpha, []>$, $[A] = W$, i.e. A is true at every world in every fixed margin model.

One can regard a fixed margin model as a standard possible worlds model with C in place of   , the world x being accessible from the world w just in case $d(w, x) \leqslant \alpha$. The definition of a metric requires accessibility to be reflexive and symmetric. These are the only constraints on accessibility, in the sense that, for any reflexive symmetric relation R on a set W, there is a metric d on W such that for all $x, y \in W$, $x R y$ if and only if $d(x, y) \leqslant 1$. One can define a suitable metric d by

$d(x, x) \quad = 0$

$d(x,y) \quad = 1$ if $x R y$ but $x \neq y$

$d(x, y) \quad = 2$ if not $x R y$

Thus the formulas valid in fixed margin models are precisely the modal formulas valid in reflexive symmetric models (when C is identified with). The latter are known to be precisely the theorems of the Brouwersche modal logic KTB.[1] One can thus axiomatize the formulas valid in fixed margin models by the following standard axiomatization of KTB (A, B any formulas):

(Taut) $\vdash A$ if A is a truth-functional tautology

(K) $\vdash \mathrm{C}(A \supset B) \supset (\mathrm{C}A \supset \mathrm{C}B)$

(T) $\vdash \mathrm{C}A \supset A$

(B) $\vdash {\sim}A \supset \mathrm{C} \sim \mathrm{C}A$

(MP) If $\vdash A \supset B$ and A then $\vdash B$

(RN) If $\vdash A$ then $\vdash \mathrm{C}A$

(Taut) and (MP) generate classical propositional logic, as justified for a vague language by the epistemic view of vagueness. (K) and (RN) are justified on the grounds that if a (possibly empty) set of premises are all clear, then so are their logical consequences. (T) is justified on the grounds that what is clearly the case is the case; it corresponds to the condition that every world is within the margin for error of itself. (B) corresponds to the symmetry of the relation $d(w, x) \leqslant \alpha$.

The formula $\mathrm{C}p \supset \mathrm{CC}p$ is not valid in fixed margin models, for $d(w, x) \leqslant \alpha$ and $d(x, y) \leqslant \alpha$ *do* not always imply $d(w, y) \leqslant \alpha$ when $\alpha > 0$. This result can be generalized. For any formula A, the formula $A \supset \mathrm{C}A$ is valid in fixed margin models if and only if either A is valid in fixed margin models or ${\sim}A$ is.[2] Intuitively, any formula permits a margin in which it is true but not clearly true, unless it takes up all or no conceptual space.

One principle about C in the above axiomatization of KTB is unobvious: (B). In fact, a slightly different kind of model invalidates (B). Rather than taking a margin of α to be sufficient, one takes any margin greater than α to be sufficient. α itself is not quite wide enough. Formally, a *variable margin model* is a quadruple $< W, d, \alpha, [] >$ satisfying the same constraints as a fixed margin model, except that the clause for [C*A*] is

$$[CA] \quad = \{w \in W: \exists\delta > a\forall x \in W\, d(w, x) \leqslant \delta \Rightarrow x \in [A]\}$$

A formula *A* is valid in variable margin models if and only if for every such model $< W, d, \alpha, [] >$, $[A] = W$.

Since metrics are by definition symmetric ($d(x, y) = d(y, x)$), the failure of the (B) axiom in variable margin models may seem surprising. To understand it, consider an example. Let *W* be the set of real numbers, $d(x, y) = |x - y|$, $\alpha = 1$, and $[p] = \{w: w \neq 0\}$.Then $[Cp] = \{w: \exists\delta > 1\ \forall x\, |w - x| \leqslant \delta \Rightarrow x \neq 0\} = \{w: w < -1 \text{ or } 1 < w\}$. Hence $[C \sim Cp] = \{w: \exists\delta > 1\ \forall x\, |w - x| \leqslant \delta \Rightarrow -1 \leqslant x \leqslant 1\} = \{\}$.Thus $0 \notin [\sim p \supset C \sim Cp]$, so (B) is invalid.

If $\alpha = 0$ in a variable margin model $<W, d, \alpha, [] >$, then $[Cp \supset CCp] = W$. However, that formula is not valid on variable margin models. In the previous example, $[Cp] = \{w: w < -1 \text{ or } 1 < w\}$ but $[CCp] = \{w: w < -2 \text{ or } 2 < w\}$. As before, the result can be generalized. For any formula *A*, the formula $A \supset CA$ is valid in variable margin models if and only if either *A* is valid in variable margin models or ~*A* is.[3] Any formula permits a margin in which it is true but not clearly true, unless it takes up all or no conceptual space.

Variable margin models are harder than fixed margin models to compare with standard possible world models, for the former do not use a single accessibility relation. Rather, a variable margin model has a family of accessibility relations, one for each $\delta > \alpha$. Nevertheless, the formulas valid in variable margin models correspond precisely to the theorems of the modal system KT. They can be axiomatized simply by the omission of (B) from the axiomatization of KTB above.

Theorem *A* is valid in variable margin models if and only if $\vdash_{KT} A$.

Sketch of proof (I) Soundness. Let $< W, d, \alpha, [] >$ be a variable margin model. We need to show that if *A* is an instance of (Taut), (K) or (T) then $[A] = W$ and that (MP) and (RN) preserve this property. Of these, only (K) is worth looking at. Suppose that $w \in [C(A \supset B)] \cap [CA]$. Then there are δ_1, $\delta_2 > \alpha$ such that $d(w, x) \leqslant \delta_1$ implies $x \in [A \supset B]$ and $d(w, x) \leqslant \delta_2$ implies $x \in [A]$. Let $\delta = \min\{\delta_1, \delta_2\} > \alpha$. Then $d(w, x) \leqslant \delta$ implies $x \in [A \supset B]$ and $x \in [A]$, so it implies $x \in [B]$. Hence $w \in [CB]$. The rest is routine. Thus every theorem of KT is valid in every variable margin model.

(II) Completeness. Suppose that not $\vdash_{KT} A_0$. Now KT is complete for reflexive models.[4] Thus there is a reflexive relation R on a set W and a mapping $|\,|$ from formulas to subsets of W such that $|A_0| \neq W$ and for all formulas A and B:

$$|\sim A| = W - |A|$$
$$|A \wedge B| = |A| \cap |B|$$
$$\mathrm{C}A = \{w \in W\colon \forall x \in W\ w\,R\,x \Rightarrow x \in |A|\}.$$

We use W, R and $|\,|$ to construct a variable margin model $<W^*, d, \alpha, []>$ such that $[A_0] \neq W$. Put:

$$W^* = \{x, y, i\colon x \in W, y \in W, y\,R\,x, i \in \mathrm{N}\}$$
$$\begin{aligned} d(<w, x, i, y, z, j>) &= 0 \text{ if } w = y, x = z \text{ and } i = j;\\ &= 1 \text{ if } w = z \text{ and } x = y;\\ &= (i+2)/(i+1) \text{ if } w \neq z \text{ but } x = y;\\ &= (j+2)/(j+1) \text{ if } w = z \text{ but } x \neq y;\\ &= 2 \text{ if } w \neq z \text{ and } x \neq y. \end{aligned}$$

Define [] by setting $[A] = \{<w, x, i> \in W^*\colon w \in |A|\}$ for each atomic formula A and letting the recursive clauses in the definition of a variable margin model do the rest, with $\alpha = 1$. d is easily seen to be a metric. Thus $<W^*, d, 1, []>$ is a variable margin model.

Lemma $[A] = \{<w, x, i> \in W^*\colon w \in |A|\}$ for every formula A.

Proof By induction on the complexity of A.
Basis By definition of [].
Induction step Suppose $[A] = \{<w, x, i> \in W^*\colon w \in |A|\}$. We must show $[\mathrm{C}A] = \{<w, x, i> \in W^*\colon w \in |\mathrm{C}A|\}$ (the cases of $\wedge$ and $\sim$ are trivial).

Suppose $<w, x, i> \in [\mathrm{C}A]$. We must show $w \in |\mathrm{C}A|$. Suppose $w\,R\,y$. We must show $y \in |A|$. Now for some $\delta > 1$,

$$d(<w, x, i>, <y, w, j>) \leqslant \delta \text{ implies } <y, w, j> \in [A] \text{ for } j \in \mathrm{N}.$$

By definition of d,

$$d(<w, x, i>, <y, w, j>) \leqslant (j+2)/(j+1) \text{ for } j \in .$$

For large enough j, $(j+2)/(j+1) \leqslant \delta$, so

$$d(<w, x, i, y, w, j>) \leqslant \delta.$$

Thus for some j, $<y, w, j> \in [A]$. By induction hypothesis, $y \in |A|$, as required. Thus $w \in |CA|$.

Conversely, suppose $w \in [CA]$. We must show $<w, x, i> \in [CA]$. The value $(i+3)/(i+2)$ for δ will do. Suppose

$$d(<w, x, i>, <y, z, j>) \leqslant (i+3)/(i+2).$$

We must show $<y, z, j> \in [A]$. Since $(i+3)/(i+2) < (i+2)/(i+1) \leqslant 2$, the inequality implies, given the definition of d, that either $w = y$, $x = z$ and $i = j$ or $w = z$. In the former case, note that $w \in |CA|$ (since $w \in |A|$ and R is reflexive), so by induction hypothesis $<y, z, j> = <w, x, i> \in [A]$. In the latter case ($w = z$), since $y, z, j \in W^*$, $z R y$, i.e. $w R y$. Since $w \in |CA|$, $y \in |A|$, so by induction hypothesis $y, z, j \in [A]$. Thus $<w, x, i> \in [CA]$. This completes the proof of the lemma.

To complete the proof of the theorem, we need to show $[A_0] \neq W^*$. Now for some $w \in W$, $w \notin |A_0|$. But $w R w$, since R is reflexive, so $<w, w, 0> \in W^*$; by the lemma $w, w, 0 \notin [A_0]$, so $[A_0] \neq W^*$. ■

It can now be argued that KT is the logic of clarity. The argument does not require variable margin models to be the only intended models of the logic of clarity. Indeed, that would be implausible, for Section 8.3 noted that the necessary margin might vary from proposition to proposition; the use of a single α for all formulas is an idealization. What may legitimately be assumed is that the variable margin models are amongst the intended models. This is quite plausible; no variable margin model should be ruled out on logical grounds alone as giving the structure of clarity for at least some propositions.

The argument is as follows. Evidently, the axioms and inference rules of KT are informally valid in the logic of clarity. Thus all theorems of KT are informally valid. Conversely, suppose that A is informally valid. Then no intended model of the logic of clarity invalidates A. Since all variable margin models are intended models, A is valid in variable margin models.[5] By the completeness theorem above, A is a theorem of KT. Thus the informally valid formulas of the logic of clarity are precisely the theorems of KT.

One further property of KT is worth noting. For any formulas A, B, if CA $\equiv$CB and C ~ A $\equiv$C ~ B are theorems of KT, then so is $A \equiv B$.[6] Thus A and B differ logically in the conditions in which they are true only if they differ logically in the conditions in which they are clearly true or in the conditions in which they are clearly false. Truth-conditions supervene on conditions for clear truth and for clear falsity. The significance of the result should not be exaggerated, for it concerns logical rather than metaphysical possibility, and assumes that atomic propositions are logically independent. Nevertheless, it indicates a sense in which the epistemic view does not force what is true to be unconstrained by what is clear.

Notes

1 THE EARLY HISTORY OF SORITES PARADOXES

1 In this chapter I am greatly indebted to Barnes 1982 and Burnyeat 1982. Barnes gives priority to Abraham, who presented the Lord with a sorites series in his defence of Sodom (Genesis 18: 23–33). Lorenzo Valla mentions the same passage in his discussion of sorites arguments (*Dialecticae disputationes* III xii), commenting that Abraham dared not go below ten for the number of righteous Sodomites because the differences between the numbers were becoming too large. Masson-Oursel 1912 discusses related Indian and Chinese arguments, but they seem to be sorites only in the harmless sense discussed at the end of Section 1.4, not intended to be paradoxical.

2 Diogenes Laertius, *Lives of the Philosophers* 2.108; Kneale and Kneale 1962: 114; Sedley 1977: 102; Barnes 1982: 36–7.

3 Barnes 1982: 32–3; Burnyeat 1982: 317–18.

4 Aristotle, *Physics* 250a 19–22; Moline 1969: 393 n.3; Sedley 1977: 112 n.85, Barnes 1982: 36–7; Sillitti 1984; Mignucci 1993: 231–2.

5 Beth 1959: 21–3; Moline 1969; Barnes 1982: 37–41.

6 Sedley 1977: 78.

7 Sedley 1977: 92–3.

8 Long and Sedley 1987 I: 51.

9 Barnes 1982: 42–3.

10 Barnes 1982: 43–4; Frede 1983.

11 Kneale and Kneale 1962: 116.

12 Barnes 1982: 41–2.

13 Long and Sedley 1987 I: 104, translating Cicero, *On Fate* 21.

14 Barnes 1982: 30.

15 Cicero, *Academica* 2.94; Sedley 1977: 93; Barnes 1982: 52–3; Burnyeat 1982: 335–6. The evidence mentioned in the text suggests that Chrysippus did not blame the sorites on ambiguity as to which proposition was expressed (after all, some of the sorites-susceptible terms were drawn from Stoic theory); for Chrysippus on ambiguity, Long and Sedley 1987 I: 227–9. Mignucci 1993

hints that Chrysippus used a notion of degree of truth in handling the para dox, but does not reconcile such an interpretation with Chrysippus' commitment to bivalence.

16 Frede 1983 for a general account of Stoic epistemology; Long and Sedley 1987 I: 236–66 for the sources.

17 Cicero, *Academica* 2.85.

18 Diogenes Laertius, *Lives of the Philosophers*, 7.177.

19 For the non-omniscience of the Stoic wise man, Kerferd 1978 and Burnyeat 1982: 335. The boundary between wisdom and unwisdom is another sharp Stoic cut-off point.

20 Cicero, *Academica* 2.93; Barnes 1982: 49–50. In taking silence to be accompanied by suspension of judgement, I follow Barnes against Burnyeat 1982: 334, who treats silence as a rhetorical device. Burnyeat's treatment does not explain what pattern of beliefs about the sorites series the Stoic should form. Moreover, for a sorites series for 'cognitive impression', Chrysippan quiescence is contrasted with assent (Sextus Empiricus, *Against the Professors* 7.416); since assent in Stoic epistemology is not a merely verbal matter, nor is quiescence. What could the Stoic hope to gain, even rhetorically, by not asserting out loud a proposition to which he had mentally assented?

21 To argue from (2) to (3a), suppose that *i* are not clearly few. On Stoic principles, *i* are either clearly not few or not clearly not few. If *i* are clearly not few, they are clearly not clearly few, for since what is clear (i.e. the content of a cognitive impression) is true, if *i* were clearly few then *i* would be few, which they clearly are not. On the other hand, if *i* are not clearly not few, by supposition they are also not clearly few, so by (2) they are clearly neither clearly few nor clearly not few, so *i* are clearly not clearly few. Thus (2) implies (3a). (2) implies (3b) by a parallel argument. Conversely, (3a) and (3b) together imply (2), for if two propositions are clear, so is their conjunction. The notion of clarity validates the required assumptions that what is clear is true and that what follows from what is clear is itself clear.

22 For a similar objection to the S5 principle for knowledge see Hintikka 1962: 106; Lenzen 1978: 79; Humberstone 1988: 187.

23 Barnes 1982: 55; I assume that 'Is a_i F?' in his n.78 should read 'Is ai clearly F?'.

24 Barnes's argument also assumes the S4 principle for clarity: what is clear is clearly clear. An argument that sorites series provide counterexamples to the S4 principle for clarity can be extracted from Section 8.4. It is worth noting that the Stoics themselves had the basis for a counterexample to the S4 principle for scientific knowledge. One might suppose that the wise man knows that he is wise, but a later variant on Eubulides' puzzle of the Elusive Man suggests otherwise. 'You say that the wise man is ignorant of nothing, and that nothing eludes him. How, then, can he ever have become wise? For at the instant at which a man becomes wise there is a fact of which he is ignorant,

and which eludes him, namely the fact that he is now wise.' Knowledge of the fact is subsequent to the fact. The case greatly troubled the Stoics, and Chrysippus wrote a book about it. It was held that it is possible to be wise and yet not know that one is (Sedley 1977: 94). The example can obviously be generalized. If one learns that *i* are few, there is an instant at which one knows that *i* are few but does not yet know that one knows that *i* are few. If the reason for the delay is that it takes time for a fact to become clear, there is an instant at which *i* are clearly few but not clearly clearly few. If what takes time is only assent, the case is not itself a counterexample to the S4 principle for clarity, but it is still very similar to the counterexample in Section 8.4, with temporal boundaries in place of conceptual ones. It cannot be taken for granted that Chrysippus would have accepted the S4 principle for clarity.

25 *Against the Professors* 7.416; translation from Long and Sedley 1987 I: 223–4. Annas 1990 argues that the Stoics vacillated as to whether the cognitiveness of an impression must be available to its subject.

26 Barnes 1982: 52; Burnyeat 1982: 335.

27 Barnes 1982: 27.

28 Barnes 1982: 47–9.

29 Long and Sedley 1987 I: 218.

30 For the bearing of Stoic views of conditionals on sorites reasoning, Sedley 1977: 91, 1982: 255, 1984; Barnes 1982: 28–9; Burnyeat 1982: 321; Long and Sedley 1987 I: 229–30.

31 Long and Sedley 1987 I: 222, translating Diogenes Laertius 7.82.

32 Burnyeat 1982: 321.

33 Long and Sedley 1987 I: 306, translating Plutarch, *On Common Conceptions*, 1084 c–d.

34 Long and Sedley 1987 I: 391, translating Alexander, *On Fate* 207.5–21; contrast Mignucci 1993:245 (an interpretation on which the Stoics should have regarded such arguments as invalid).

35 Barnes 1982: 45.

36 Long and Sedley 1987 I: 225, translating Cicero, *Academica* 2.93–4.

37 Long and Sedley 1987 I: 463, translating Sextus Empiricus *Against the Professors* 9.182–4.

38 Barnes 1982: 46; Burnyeat 1982: 326–33.

39 Long and Sedley 1987 I: 223, printing *On Medical Experience* 16.1–17.3.

40 Barnes 1982: 24–7, 57–9.

41 Barnes 1982: 60–3. Chrysippus may have rejected such a treatment of the sorites (Long and Sedley 1987 I: 229). It would hardly have yielded congenial results when applied to the sorites for 'cognitive impression'.

42 Burnyeat 1982: 319–20; contrast Mignucci 1993: 237–8.

43 Cicero, *Academica* 2.92.

44 Moline 1969; Barnes 1982: 60–1.

45 Horace, *Epistles* II i 36–49; Persius, *Satires* VI 75–80; Barnes 1982: 36.

46 *Expositione in Rhetor. Cic.* II 27; Prantl 1855–70 I: 663.

47 Kneale and Kneale 1962: 187–8.

48 Jardine 1977: 161. Relevant works such as Cicero's *Academica* were scarcely known in the Middle Ages (Schmitt 1983: 227). Augustine's reply to Ciceronian scepticism, *Contra Academicos*, was better known but mentions the sorites only in passing (II v 11).

49 Valla's discussion of the sorites is in the work now generally known as *Dialecticae disputationes* (III xii, Valla 1982 I: 306–12). It has been argued that his intended moral was the sceptical one that the status of such arguments must remain problematic for us (Jardine 1977: 161–2, 1983: 272–5, 1988: 180–1). The text follows Monfasani 1990: 193–6 in taking Valla's treatment to be non-sceptical.

50 The reversion was not immediate; there are non-syllogistic sorites in the textbooks of Caesarius (1467–1550) and Melanchthon (1497–1560) (Jardine 1982: 806–7).

51 These topics are discussed in Hamilton 1860 I: 366–85, 464–6.

52 Schmitt 1983.

53 Hart 1991–2 argues with surprising plausibility that heaps *do* form a natural kind; even so, large heaps do not form a natural kind.

54 *Nouveaux Essais* III v 9 and III vi 27; Wiggins 1980: 124.

55 Bain 1870: 160–1; compare 433.

56 Hegel 1892 I: 462–4, 1975: 159.

57 Hegel 1975: 172.

2 THE IDEAL OF PRECISION

1 On whether 'vague' is vague, Aldrich 1937: 94; Austin 1962: 125–31; Alston 1964: 85; Sorensen 1985, 1988b: 227–8; Deas 1989. None of them denies that 'vagueness' in ordinary English extends beyond the blurring of boundaries.

2 The *Begriffsschrift* system is in fact a higher-order predicate calculus, in which one can generalize about concepts as well as about objects.

3 In practice, mathematicians do not always adhere to their own definitions, as Frege complained. The account in the text is of what it is to treat a definition as stipulative.

4 Frege 1960:159 (*Grundgesetze der Arithmetik*, vol. 2, 1903, Section 56). Frege hints at a distinction between lack of sharpness and lack of completeness when he compares the former with a fuzzy line and the latter with a broken one, but he makes nothing of the distinction.

5 For a general survey of Frege's treatment of vagueness see van Heijenoort 1985.

6 Such a function is a first-level concept; higher-level concepts, mapping concepts to truth-values, will be ignored.

7 A worse mistake would be to suppose that one can just stipulate that an expression has a referent, without stipulating what referent it is to have.

8 Recursive function theory needs partial functions for *mathematical* reasons.
9 Frege 1960: 159.
10 Frege 1960: 168 (*Grundgesetze* Section 64).
11 Frege 1960: 159; 1979: 155 ('The argument for my stricter canons of definition', 1897–8 or shortly afterwards); 1979: 195 ('Introduction to logic', 1906); 1980: 114 (letter to Peano, 1896).
12 In Frege 1979: 179 ('On Schoenflies: *Die logischen Paradoxien der Mengenlehre*', 1906), the sharpness requirement is associated with the principle of non-contradiction as well as that of excluded middle.
13 Aminor modification is needed to take account of logical principles specified in terms of a distinguished predicate (e.g. $\forall x\, x = x$).
14 The theorem is 'If x has a property F that is hereditary in the f-sequence, and if y follows x in the f-sequence, then y has property F' (van Heijenoort 1967: 62; *Begriffsschrift* Section 27).
15 Ziff 1974, 1984: 141–2; Weiss 1976; Smith 1984; criticized by Dummett 1975 and Sorensen 1988b: 219–24.
16 Frege 1980: 114; compare 1979: 155.
17 Dummett 1981: 32–3, 34.
18 Frege 1960: 170–1 (*Grundgesetze* Section 66).
19 What makes a coherent account of vagueness impossible on Fregean terms is higher-order vagueness; see Section 2.4 (3). Perhaps the nearest Frege comes to recognizing the phenomenon is the passage mentioned in n.4.
20 Frege 1980: 115. Frege says that for practical purposes it is enough for your thought to be approximately the same as mine.
21 Frege 1979: 155; 1980: 115.
22 Frege 1979: 122 ('Comments on sense and meaning', 1892–5).
23 See McDowell 1977; Evans 1982: 22–30; Dummett 1981: 129–38; Salmon 1990: 235–46. Frege undoubtedly thought that there could be a sense without a referent; the question is whether he was right, given his accounts of these notions.
24 Frege 1960: 159. Burge 1990 suggests that Frege took reference failure to be a comparatively rare phenomenon, and that much of his talk of unclarity is epistemic: one may fail fully to grasp a sense which one deploys in thought (and not because other members of the community do fully grasp it). Although some of Frege's remarks do point towards such a view, it undermines his cognitive criterion for the identity of senses. Burge acknowledges at n.16 that this criterion limits the scope for ignorance of the senses one deploys, but it threatens to leave no scope at all. If a thought which equates the referents of a clearly grasped sense and an unclearly grasped sense is informative, as it appears to be, then the senses are distinct by the cognitive criterion. If it is replied that a clear grasp of the thought would reveal it to be not really informative, what prevents a similar reply to paradigm applications of the cognitive criterion? A Fregean would be better advised to emphasize ignorance of referent rather than of sense, although Frege never treated

vagueness in such terms, and regularly associated vagueness with reference failure. (*Note:* The text ignores indirect contexts.)

25 A conception of a logically perfect language similar to Frege's is found in Wittgenstein's *Tractatus Logico-Philosophicus*, with the difference that 'All the propositions of our everyday language, just as they stand, are in perfect logical order' (5.5563). The nearest the *Tractatus* comes to mentioning vagueness is 3.23: 'The requirement that simple signs be possible is the requirement that sense be determinate'. 'Determinate' here excludes more than vagueness; a proposition unanalysable in terms of simple signs counts as indeterminate, but it need not be vague (see also 3.24). However, vagueness is a form of indeterminacy too. The *Notebooks* for June 1915 confront vagueness as an obstacle to determinacy and full analysis. 'I only want to justify the vagueness of ordinary sentences, for it *can* be justified' (1961b: 70). What a proposition says 'must be susceptible of SHARP definition' (60). Sharpness is clearly distinguished from specificity. A proposition with a sharp sense may be made true by more than one possible arrangement of simple objects, and thus be an incomplete description of any one of them, but there must be a sharp line between those which make it true and those which make it false. Yet a proposition such as 'The book is on the table' does not seem to be perfectly sharp in this way; it can be hard to classify as true or false (67). Wittgenstein is attracted by the view that, although a sentence in isolation does not resolve borderline cases, what one means by it on a particular occasion does (68, 70). He has in mind situations where he sees the book on the table in front of him, supposing that its quite specific spatial relation to the table will give the proposition a sharp sense. This solution is hopeless. It does not apply when the proposition is evidently false, nor when it is true but the speaker cannot perceive the book on the table. Even if the book is on the table and I am looking at it, I do not mean by 'on' an absolutely specific spatial relation; if the book is a fraction of an inch from where it looks to me to be, it can still be firmly on the table in the sense I mean. Even if I did mean an absolutely specific relation by 'on', it would be of no avail when I used the sentence to tell you where the book was, for you would not know which absolutely specific relation I meant. Vagueness is reduced but not eliminated by context and intention. See Carruthers 1990: 47–9 and 67–8 for more discussion.

26 5.498 and 5.508 (all such references are to Peirce 1931–56 by volume and paragraph). For Peirce on vagueness in the light of his overall views, Thompson 1953: 213–27; Brock 1979; Nadin 1980; Hookway 1985: 231–3, 237–8, 1990; Engel-Tiercelin 1986. For his anticipation of many-valued logic, Section 4.3.

27 Baldwin 1901–22: 748.

28 Rather dubiously, Peirce's definition implies that a case in which speakers are determinately disposed to shrug their shoulders and say 'It's a borderline case' does not constitute vagueness.

29 5.505. The distinction between the two senses of ‘vague’ is not, as sometimes alleged, the distinction between vagueness in the predicate and vagueness in the subject. The subject can have blurred boundaries too. Note that a sentence can be vague with respect to one component and general with respect to another, e.g. ‘Someone loves everyone’.

30 Peirce held that, strictly speaking, even a proper name of an individual yields some indeterminacy, because an individual changes from moment to moment (3.93, 5.448 n.1). Certainly an implicit ‘sometimes’ or ‘always’ could make ‘Philip was tactful’ indeterminate in his sense, but no such qualification is implicit in ‘Philip lived 46 years’. No acceptable view of personal identity sustains Peirce’s apparent suggestion that genuinely singular reference has not been achieved until a particular moment of Philip’s life has been specified.

31 5.447, 5.505.

32 Peirce’s account is supposed to apply to signs in thought as well as speech (5.447). Even sensations and images can be vague (3.93). He writes ‘our own thinking is carried on as a dialogue, and though mostly in a lesser degree, is subject to almost every imperfection of language’ (5.506), but does not explain details. If the distinction between utterer and interpreter in a dialogue with oneself is just that between the earlier and the later self, all further determination would be left to the interpreter, the later self, so that our own thinking would be general and not subject to the imperfection of vagueness. If last month I thought vaguely ‘A great event will happen this month’, I must be able to act now as the producer of the thought; if I thought generally ‘Man is mortal’, I must be able to act as its consumer.

33 For semantic games in general, Hintikka 1973; Saarinen 1979. For their connection with Peirce, Brock 1980; Hilpinen 1982. The account in the text ignores many problems, e.g. the existence of unidentifiable objects.

34 Contrast Haack 1974: 109.

35 5.450.

36 5.506.

37 5.447.

38 Note also the phrase ‘a slightly different mode of application’ at 5.449.

39 5.447.

40 5.448.

41 Vagueness is discussed in the 1913 manuscript *Theory of Knowledge* as a property of memory images (Russell 1983–, vol. 7: 174–6), a theme continued in Russell 1921: 180–4. Russell 1923 was adumbrated in the 1918 ‘Lectures on Logical Atomism’ (Russell 1983–, vol. 8: 161) and read at Oxford in 1922. F.C.S. Schiller, the local pragmatist, delivered a sceptical reply (Russell 1983–, vol. 9: 145–6; Schiller 1934). Russell also discusses vagueness in Russell 1983–, vol. 8: 139–42; 1927: 220–4, 280; 1948: 276–7. He is criticized by Black 1937; Kohl 1969; Rolf 1982.

42 Russell 1923: 147–8 (as reprinted in Russell 1983–).

43 Russell 1923: 154.
44 Russell 1923: 148.
45 Rolf 1982: 76–8.
46 The point in the text holds whether vague sentences are those which can have borderline cases or those which contain a vague constituent.
47 If a mode of combination can be vague, count it as a constituent too.
48 Logical words in natural languages may in fact be vague, e.g. it may be vague whether 'if' is truth-functional, but that would not vindicate Russell's attempt to derive their vagueness from that of 'true' and 'false'.
49 Russell 1923: 149.
50 Russell 1923: 148; see also 150.
51 See further Fine 1975: 289.
52 Russell 1923: 151.
53 Russell 1923: 151.
54 See also Rolf 1982: 79–80.
55 Russell 1923: 150.
56 Russell 1923: 153.
57 There are earlier attempts to distinguish vagueness from generality in Russell 1921: 184 ('We may compare a vague word to a jelly and a general word to a heap of shot') and 220–2 ('we may say that a word embodies a vague idea when its effects are appropriate to an individual, but are the same for various similar individuals, while a word embodies a general idea when its effects are different from those appropriate to individuals'). They do not get the matter straight. In particular, they do not allow for the many words which are both vague and general. Note that Russell's admission of general facts does not rescue his account of vagueness, by allowing him to treat vagueness as a kind of ambiguity. The argument in the text shows that he would also need to admit disjunctive facts (and existential ones) and alter his account of generality. Moreover, an ambiguous sentence is not simply verified by more than one fact; each of the facts verifies it *in one of its senses*.
58 Russell 1923: 153.
59 Russell's definition differs from the now standard mathematical definition of isomorphism of structures (= systems), on which correspondence would be simply a one–one correlation of X with Y such that if $x_1, \ldots, x_k$ correspond to $y_1, \ldots, y_k$ respectively then R_i relates $x_1, \ldots, x_k$ (in that order) if and only if S_i relates $y_1, \ldots, y_k$ (in that order). Thus S_j in (vi) is required to be S_i (Bridge 1977: 15). On the standard definition, the identity of a system depends on the order in which its relations are listed, making it obvious that a system is an abstract object. Russell thinks of systems in a more concrete way and leaves no room for an 'artificial' difference between two systems differing only in the order in which the relations are listed. For more on the history of the concept of a structure see Hodges 1985–6.

60 Newman 1928.
61 The identity of the representing relation is assumed to fix the identity of the representing and represented systems.
62 The definitions allow one system to represent another both accurately and inaccurately, if there are two representing relations.
63 Russell 1923: 152. He does not mean that the condition for the representation to be vague is that the representing system should have fewer members than the represented system, for a representing relation can fail to be one–one even when it has the same (infinite) number of members as the represented system. By 'the relation' Russell means 'the representing relation'. Russell's use of the phrase 'one–many' is rather loose. Technically, correspondence is one–many if and only if it satisfies clauses (iii) and (v) above: to each member of the represented system corresponds exactly one member of the representing system (Russell and Whitehead 1910–13, vol. 1: 437). Thus a one–one correlation counts as also one–many. But when Russell calls vague representation one–many, he wants to imply that it is not one–one. He is interested in violations of clause (iv), in one thing representing many, and such a violation will presumably constitute vagueness whether or not the other clauses are satisfied. That leaves cases in which clause (iv) is satisfied but (i), (ii), (iii) or (v) violated: for example, where each element of the representing system represents at most one thing, but some of them represent nothing at all. By ordinary standards, that does not amount to vagueness. There is also a problem in adapting the definition of 'one–one' to Russell's assumption that the represented items, e.g. facts, may have merely possible existence; see Rolf 1982: 69–71.
64 See n.61.
65 The definitions allow one system to represent another both vaguely and precisely; see n.62.
66 Russell 1983–, vol. 8: 176.
67 Russell 1923: 152.
68 See Rolf 1982: 71.
69 Russell 1923: 153.
70 Schiller, who replied to the original paper, objects not to the analysis of precision ('exactness' is his word) as one–one correlation of words and meanings but to the idea that exactness so defined is a good thing (Schiller 1934). It may be noted that Russell 1948: 276–7 (discussed by Sainsbury 1979: 136–7), written at a time when he was less preoccupied with the ideal of a logically perfect language, does not confuse vagueness and generality.
71 Russell 1923: 153.
72 Russell 1927: 280–1.

3 THE REHABILITATION OF VAGUENESS

1 Cohen 1927 and Schiller 1934 are instances of work on vagueness published in the period.
2 Austin 1962: 128. The quotation is taken from a rebuttal of A.J. Ayer's claim that sense datum language is inherently more precise than material object language.
3 Austin 1962: 125–6.
4 Other work on vagueness published between 1937 and 1965 but not discussed in the text includes Aldrich 1937, Copilowish 1939, Benjamin 1939, Burks 1946, Pap 1949: 116–17, Quine 1960: 125–9, Alston 1964: 84–96 and Kaplan 1964: 65–8.
5 Black 1949: 28 (reprinting Black 1937). Peirce's 1902 definition is quoted with approval. Sometimes Black says that the mere conceivability of borderline cases is sufficient for vagueness (e.g. 1949: 30); he is not consistent on the point.
6 Copilowish 1939 replies to Black by trying to subsume vagueness under ambiguity. He argues that borderline cases involve a conflict of semantical rules, and that such a conflict amounts to ambiguity. He admits that the vagueness of an ostensively defined term such as 'red' is hard to explain on this account, but suggests that a borderline case between red and orange involves a conflict between the rules that a shade sufficiently similar to a red shade is red and that a shade sufficiently dissimilar from a red shade is not red. However, he does not show that 'red' really is subject to such rules. More r ecent philosophers have argued that 'red' is subject to the rule that a shade indiscriminable from a red shade is red (Sections 6.3– 6.4), but Copilowish does not equate 'sufficiently similar' with 'indiscriminable'.
7 Black 1949: 35.
8 Black ignores the mathematical difficulties caused by the fact that the consistency of application can be infinite. The problem is quietly solved by Hempel 1939: 165, where $m/(m + n)$ rather than m/n is used, m and n being respectively the number of positive and negative cases.
9 Black measures disagreement in the application of L to x by the limit of twice the minimum of m and n divided by $m + n$ as the latter increases (m and n as in n.8; the wording at Black 1949: 47 implies 'maximum' rather than 'minimum', but this seems to be a mistake). The limit is 0 if there is effective unanimity, 1 if opinions are equally divided, and something in between otherwise. The area under the curve made by this figure as a function of the position of x on the relevant dimension is suggested as a measure of the vagueness of L.
10 Black 1949: 48.
11 Black 1949: 48–9.
12 Higher-order vagueness makes Black's replacement principle suspect. Note also that truth cannot be equated with consistency of application of at least 1,

on pain of contradiction, for otherwise both L and 'not L' will be true of x when speakers of the language are equally divided. Moreover, Black tacitly assumes that L and 'not not L' are used interchangeably; if some speakers prefer to call x 'not L' rather than L, but 'not not L' rather than 'not L', the consistency of application both of L and of 'not L' to x may be less than 1.

13 Black says that his definition preserves the transitivity of exclusion, but exclusion is not transitive. If 'scarlet' excludes 'green' and 'green' excludes 'red', it does not follow that 'scarlet' excludes 'red'.

14 Compare the triangle inequality in the mathematical definition of a metric. Black incorrectly has 'less' in place of 'not greater' (1949: 57). He also introduces a variable for consistency of application into the inequality, treating the meta-linguistic expression '$i(L, M)$', which supposedly designates the degree to which inclusion fails, as though it were a universally quantified conditional of the original vague language with M in the antecedent and L in the consequent. Black has a tendency to mistake pomposity for rigour.

15 If the expression is ambiguous, its different senses must be treated separately. Black claims that the ambiguity can be detected in the consistency profile, in which two drops will indicate two boundaries. This test is not perfectly comprehensive: it will not detect an ambiguity between the senses 'flammable' and 'not flammable'.

16 Black 1949: 54.

17 Black's experiment does not even measure the consistency profile of the word 'natural', for subjects were not given the opportunity to describe no division or more than one as 'natural'. Equally, it fails to model the evolution of linguistic regularities over time, for it involves no interaction between speakers.

18 Hempel 1939: 173.

19 Compare Russell's denial that property possession is gradable, in amplification of his claim that vagueness belongs only to representations (Section 2.4 (1)). For a claim that property possession is gradable, Burks 1946: 483.

20 Hempel 1939: 176.

21 Black 1949: 249–50.

22 The assumption that a language cannot be meaningful unless *we* can understand it is also objectionably anthropocentric. It would be fallacious to defend it on the grounds that 'meaningful' is *our* word. Why should our words apply only to what we can know them to apply to?

23 Hempel 1939: 167, 170.

24 Black 1949: 43.

25 Black 1949: 28.

26 Hempel 1939: 179. Hempel also criticizes Black's use of the steepness of consistency profiles to measure vagueness. It assumes that the objects can be placed in a unique way on a quantifiable dimension, such as height; Hempel points out that there may be no such dimension (165). 'Gnarled' is a vague

word, but we have no linear scale of gnarledness. Hempel proposes a measure of vagueness in terms of consistency of application that does not rely on the scaling assumption (166).

27 Black 1970: 10–11, reprinting Black 1963.

28 'We run the risk of using the rules of logic *blindly*' (Black 1970: 13); a contrast is no doubt intended with Wittgenstein 1953, Section 219: 'I obey the rule *blindly*'. According to Black, logic is sometimes applicable beyond the indisputably clear cases; 'the short man' may denote what would otherwise be judged a borderline case of 'short man' when the only other man in question is taller; 'Short men usually marry short women' may be true on any reasonable way of drawing the line (12). The intended bearing on logic of these examples is hard to guess, for in neither case does Black mention a valid or invalid argument. At least two claims seem to be made: that 'borderline case' is context-relative and that the occurrence of a borderline case does not always produce indeterminacy in truth-value. Perhaps determinacy of truth-value, rather than the absence of borderline cases, is supposed to be the condition for the applicability of logic. Odegard 1965 and Campbell 1974 make relevant criticisms.

29 On a strict Tractarian view, the crystalline purity of language must be shown rather than said.

30 Such claims are developed in Khatchadourian 1962 and Kohl 1969.

31 The idea in the text is not to be confused with the idea that z may be a game at t but not at t^* (for 'game' is used in this sentence in its sense as it now is). The sense of 'number' has developed, but it has always been such that 'If x is a number at t then x is a number at t^*' is true. Conversely, the fact that x may be a child at t but not at t^* does not imply any development in the concept of a child.

32 Marmor 1992: 133.

33 Contrast the opponent in a Peirce–Hintikka game.

34 Sorites paradoxes are mentioned at Wittgenstein 1975 Section 211 (see Geach 1956: 72–3). Wittgenstein notes that we can sometimes box a vague predicate F in between a precise sufficient condition A and a precise necessary condition B, where A entails B but not vice versa. F will therefore have some precise logical properties. Its borderline cases are outside A and inside B. The point is consistent with second-order vagueness, for although all borderline cases for F are cases of the precise conjunction of B with the negation of A, not all cases of the latter need be borderline cases for F.

35 Waismann 1945: 121.

36 Waismann 1945: 126. Contrast Wittgenstein on 'number'.

37 Waismann 1945: 123.

38 Waismann 1951. Pap 1958: 326–7, 346, 355–6 applies Black's account of vagueness to similar effect, introducing degrees of meaning and entailment. If speakers are more reluctant to apply the term 'mature lemon' to a non-yellow mature fruit than to a non-sour one, 'Mature lemons are yellow' is more

analytic than ‘Mature lemons are sour’. An alternative view is that such dispositions tell us more about speakers’ beliefs than about the meaning of ‘mature lemon’.

39 Waismann’s discussion of the logic of sense impressions was related to a debate on the indeterminacy of sense data and mental images; see Price 1932: 149–50; Ayer 1940: 125; Chisholm 1942; for recent discussions, Sorensen 1989 and Tye 1991: 103–16. The relevance of the debate to issues about vagueness is limited both by the peculiarly subjective status of the items at issue and by a tendency to treat unspecificity as vagueness.

40 Waismann 1945–6: 98. He refers to the restriction of excluded middle to decidable propositions in Brouwer’s intuitionistic logic, but denies that the latter is appropriate for sense impressions (compare Black 1949: 36–7). Waismann’s attitude to logic is pragmatic: he thinks that the law turns out not to be useful.

41 For a more sympathetic view of anti-realism about the past see Dummett 1968–9.

42 Waismann 1945–6: 100.

4 MANY-VALUED LOGIC AND DEGREES OF TRUTH

1 ‘p_1’, .. . , ‘p_n’, ‘p’, ‘q’, . . . are sentence variables. Complex formulas specify types of sentence, e.g. $p \vee \sim p$ specifies the type of sentence consisting of the disjunction of a sentence with its negation. ‘A’, ‘B’, . . . are variables over formulas of the formal language.

2 The new test is mechanical only if the number of values is finite.

3 Fisch and Turquette 1966.

4 A speculative connection (not Lukasiewicz’s) between the supposed indeterminacy in vagueness and the supposed indeterminacy of the future is that it is not yet determined how (or indeed whether) present borderline cases will be resolved. For the history of many-valued logic see Rescher 1969: 1–16.

5 Halldén 1949: 9.

6 Halldén 1949: 76.

7 Halldén 1949: 86–7.

8 Halldén’s tables correspond to those of Bochvar’s ‘internal’ system in an article originally published in Russian in 1939 (Bochvar 1981; Rescher 1969: 29–34). However, Halldén’s work was independent (Bochvar does not appear in his extensive bibliography). They also differ in the values they designate.

9 *Proof:* Let C be the conclusion, $A_1, \ldots, A_m$ the premises involving only sentence variables in C and $B_1, \ldots, B_n$ the other premises. If the argument from $A_1, \ldots, A_m$ to C is classically invalid, some assignment of truth or falsity to the sentence variables in C makes $A_1, \ldots, A_m$ true and C false on the two-valued tables; it can be extended to an assignment that makes every sentence variable not in C ‘meaningless’, which makes $A_1, \ldots, A_m$ true, $B_1, \ldots, B_n$ ‘meaningless’

and C false on Halldén's tables; thus the argument from $A_1, \ldots, A_m, B_1, \ldots, B_n$ to C is Halldén-invalid. Conversely, if the argument is Halldén-invalid, some assignment makes $A_1, \ldots, A_m$ non-false and C false on Halldén's tables; since C is not 'meaningless', no sentence variable in it is 'meaningless', so no sentence variable in $A_1, \ldots, A_m$ is 'meaningless', so $A_1, \ldots, A_m$ are true or false; since they are non-false, they are true; thus the argument from $A_1, \ldots, A_m$ to C is classically invalid.

10 *Proof:* Suppose that the argument from $A_1, \ldots, A_m$ to C is classically valid and that every sentence variable in C is in some A_i. If an assignment makes $A_1, \ldots, A_m$ true on Halldén's tables, they are 'meaningful', so every sentence variable in them, and therefore every such variable in C, is 'meaningful'; since Halldén's tables extend the two-valued ones, the assignment makes $A_1, \ldots, A_m$ true on the latter, so it makes C true on the latter, and so on Halldén's tables. Again, suppose that $A_1, \ldots, A_m$ form a classically inconsistent set. By similar reasoning, no assignment makes them all true on Halldén's tables. Thus if the argument from $A_1, \ldots, A_m$ to C is classically valid and either every sentence variable in C is in some A_i or $A_1, \ldots, A_m$ form a classically inconsistent set, then the argument preserves truth on Halldén's tables. Conversely, suppose that the argument preserves truth on Halldén's tables. Since they extend the two-valued tables, it is classically valid. If $A_1, \ldots, A_m$ form a classically consistent set, some assignment makes them all true on the two-valued tables and therefore on Halldén's while making every sentence variable not in them 'meaningless'; since C is then true and therefore 'meaningful' on Halldén's tables, every sentence variable in it is in some A_i.

11 For technical discussions of Halldén's logic of nonsense see Åqvist 1962 and Segerberg 1965. One response to the problems raised in the text would be to translate sentences of natural language into more complex formulas involving +. It is not discussed for two reasons: it tends to lose what is distinctive in Halldén's approach, and it does not avoid the problem of second-order vagueness (Section 4.6).

12 Halldén 1949: 83–6.

13 A doubt about the tempting claim is raised in Section 3.4.

14 Körner 1955, 1959, 1960 and 1966.

15 Körner 1966: 37–40. As he notes, his tables were anticipated by the logician Kleene for another purpose (Kleene 1938). A recent attempt to apply Kleene's tables to vague languages is Tye, forthcoming.

16 Körner makes a less conventional suggestion, defining an interpreted proposition to be logically true just in case, whenever its neutral components are either all deleted (together with the functors governing them) or all replaced by true or false propositions, the result is classically valid. For example, let p be 'Jack is bald' and q 'Jack is a philosopher'. If p is true or false, the proposition $(p \supset p) \vee q$ is logically true because classically valid; if p is

neutral while q is true or false, $(p \supset p) \vee q$ is not logically true, since the result of deleting its neutral components is q, which is not classically valid. Whether a proposition counts as a logical truth may thus depend on Jack's scalp. This notion makes no use of Körner's three-valued tables. It is not easy to find it a rationale. For detailed discussion of Körner's logic see Kumar 1968, Körner 1968 and Cleave 1970.

17 Haack 1974: 60.

18 Körner 1959: 128; 1960: 164.

19 *Proof:* Let A be a formula made up from sentence variables and the primitive operators ~, $\vee$, &, $\supset$ and $\equiv$. Interpret *A* by letting each variable stand for a neutral proposition. Then some sequence of elections makes *A* true and some sequence makes it false when distinct occurrences of the neutral proposition are subject to inde pendent elections. This can be shown by mathematical induction on the length of *A*, since the two-valued table for each primitive operator has both a T and an F in its right-hand column. Thus any inference form has an instance with true premises and a false conclusion for a suitable sequence of elections, since these are also held independently for each premise and the conclusion. A little logic could be restored by the addition of a constant false symbol $\perp$ to the language, but not much.

20 Halldén 1949: 56–7.

21 Körner 1959: 126; 1960: 160. See also Horgan, forthcoming, and Tye, forthcoming, for discussion of higher-order vagueness in relation to three-valued logic.

22 Imagine the sentence as expressing the speculation of one who cannot see the patch, so that the state of the patch does not have a contextual effect on the reference of 'dark'. 'Darker' is also assumed not to imply 'dark'; the argument can easily be adapted to the opposite assumption.

23 See Lukasiewicz and Tarski 1930. They concentrate on the logic $L(_0)$, the set of sentences valid when the available values are the rational numbers between 0 and 1, rather than $L(_1)$, the set of sentences valid when (as in the text) all real numbers between 0 and 1 are available. However, $L(_1)$ and $L(_1)$ turn out to be equivalent (Theorem 16, proved by Lindenbaum). Thus in the propositional calculus it would make no difference to the set of valid formulas if degrees of truth were ordered like the rationals rather than the reals (the original motivation assumed that moments of time are ordered like the reals; how do we know that they are not ordered like the rationals?). However, the semantics for the quantifiers is well defined only if every set of degrees of truth has a greatest lower bound and a least upper bound. This holds for the reals but not the rationals between 0 and 1.

24 See Rescher 1969: 42. A similar problem is raised by C_α for $0 < \alpha < 1$.

25 For a quite different attempt to adapt set theory to vagueness, using strict finitist ideas, Vopenka 1979: 33–8. The paper that anticipates fuzzy set theory is Kaplan and Schott 1951.

26 Goguen 1969. The existence of fuzzy logic does not contradict the claim in the previous paragraph that applications of fuzzy theory are made within a framework of classical logic and mathematics, for the meta-logic used by fuzzy logicians is classical (e.g. Goguen 1969: 327). Compare Section 4.12.

27 Degree of baldness might be held to decrease with decreasing rapidity as the number of hairs increases. This and similar suggestions affect the letter but not the spirit of the account.

28 If the non-standard conditional $\rightarrow$ were used in place of $\supset$, the sorites argument would be valid in both senses, but all its conditional premises would be perfectly false, leaving their attraction unexplained. One could also define a deviant conditional to be perfectly true provided that its consequent did not drop below its antecedent in degree of truth by more than, say, 1/100,000 (compare Peacocke 1981: 127). All the premises of the sorites argument would then be perfectly true, but *modus ponens* for the deviant conditional would not preserve perfect truth. This suggestion makes the apparent validity of *modus ponens* puzzling. If negated conjunctions were used rather than conditionals, some premises would be little more than half true (Wright 1987: 251–2); the degree theorist may be forced to claim that such a formulation does not capture the intuition behind the sorites paradox; see Section 4.14 for related issues). For more elaborate treatments of sorites paradoxes within some form of fuzzy logic see Goguen 1969 and Machina 1976. Goguen's approach is applied to modal sorites paradoxes in Forbes 1983, 1985. Note that the sorites paradoxes discussed in this chapter turn on vagueness, not on observationality; for the latter, see Sections 6.3–6.4. Not all degree-theoretic treatments of sorites paradoxes are committed to numerical degrees; see Peacocke 1981, Sainsbury 1988–9 and Section 4.13. Relevant material is also to be found in Machina 1972; Sanford 1975; King 1979; Sainsbury 1986, 1992; Burgess and Humberstone 1987; Dubois and Prade 1988.

29 Other relevant constructions are 'more *F* than', 'at least as *F* as', 'more of an *F* than' ('*F*' a noun phrase) and the superlative. The need for a compositional account is emphasized in Klein 1980.

30 If the semantics gives the degree to which '*x* is *F*' is true as the degree to which *x* is *F*, and then explains the latter in comparative terms (e.g. by specifying that two things are *F* to the same degree just in case each is as *F* as the other), circularity threatens the project of explaining the comparative as a compound of the positive (Klein 1980: 5). However, the degree theorist might take the notion of the degree to which something is *F* as primitive.

31 See further Lakoff 1973 and Zadeh 1972, 1975.

32 Klein 1980: 6.

33 Accounts of comparatives not based on fuzzy logic include Cresswell 1976, Klein 1980 and von Stechow 1984. von Stechow denies that the semantics of comparatives depends on vagueness (75). Although Klein allows for truth-value gaps in respect of vagueness, the real work in his account is done by contextually determined comparison classes, a feature independent of

vagueness. For a logic of degrees based on classical logic see Casari 1987. For a general account of degrees that allows a continuum of degrees even where there is a cut-off point see Engel 1989 (he gives the example of acidity).

34 Similar remarks apply to attempts to derive degrees of truth from judgements of prototypicality. There is good evidence that speakers of English regard sparrows as more prototypical birds than penguins, apples as more protypical fruit than figs and 7 as a more prototypical odd number than 447. Such judgements are irrelevant to vagueness; 'Penguins are birds', 'Figs are fruit' and '447 is odd' are all clearly true. This is not to deny that prototypicality is an important feature of many concepts. For the connection, or lack of it, between prototype theory and vagueness see Rosch 1973, Rosch and Mervis 1975, Lakoff 1973, Osheron and Smith 1981, 1982, Zadeh 1982, Armstrong, Gleitman and Gleitman 1983 and Fuhrmann 1988a, 1988b, 1991.

35 A similar problem arises for the theorem $(p \supset q) \vee (q \supset p)$ of the Lukasiewicz logics, and its special case $(p \supset \sim p) \vee (\sim p \supset p)$, a kind of substitute for the law of excluded middle (note that, unlike p and $\sim p$, both disjuncts of $(p \supset \sim p) \vee (\sim p \supset p)$ can be perfectly true on the many-valued approach, in the special case in which p is *exactly* half-true; but since this case will be extremely hard to identify, it makes the decision as to what to assert no easier). The complication is that the critical assumption of the meta-linguistic reasoning is that either $[p] \leqslant [q]$ or $[q] \leqslant [p]$, which depends less on classical logic than on the framework of numerical, linearly ordered degrees of truth; see Section 4.13.

36 'Our models are typical purely exact constructions, and we use ordinary exact logic and set theory freely in their development . . . . It is hard to see how we can study at all rigorously without such assumptions' (Goguen 1969: 327); Machina claims that the assignment of degrees of truth 'would have something of the character of a scientific hypothesis in empirical semantics' (1976: 61), as though the problem is epistemological.

37 The loss of excluded middle in the meta-language means that definitions of operators such as $\rightarrow$, and $J\alpha$ by cases may be incomplete. For example, $[J_1 p] = 1$ if $[p] = 1$ and $[J_1 p] = 0$ in any other case. But to assume that either $[J_1 p] = 1$ or some other case obtains is to assume a problematic instance of excluded middle. The degree theorist may therefore not grant that either $[J_1 p] = 1$ or $[J_1 p] = 0$. No explicit provision has been made for any alternative, but no explicit provision has been made for any alternative to a man's being either tall or not tall. This point creates difficulties for the argument of Wright 1992, especially at 134, defending the paradox of higherorder vagueness in Wright 1987 against Sainsbury 1991 (J_1 corresponds to Wright's definiteness operator and $\sim J_1$ to his broad negation). The degree theorist may treat sorites arguments based on notions such as on-balance-justifiability similarly (compare Wright 1975: 348–51, 1976: 237–40). See also Engel 1992.

38 A related problem is that, in order to accommodate second-order vagueness, one seems obliged to say things like: '$[p] = 0.61$' is true to degree 0.8, but '$[p] = 0.67$' is true to degree 0.9. How can any function assign more than one value

to p (Rolf 1984: 222–3)? The assignments suggested are certainly problematic. If '$[p] = 0.67$' is true to degree 0.9, then '$[p] \neq 0.67$' is true only to degree 0.1; but it should be at least as true as '$[p] = 0.61$', which is supposedly true to degree 0.8. However, all that follows is that no two statements assigning distinct degrees of truth to $[p]$ can be true to more than degree 0.5. A conjunction such as '$[p] = 0.61$ & $[p] = 0.67$' can still receive a non-zero degree of truth, but this is just an extension of a point about the treatment of first-order vagueness, that a contradiction can receive a non-zero degree of truth (Section 4.14).

39 See Goguen 1969: 350–1; Forbes 1985: 175.

40 Compare Goguen 1969: 355; Forbes 1985: 176 deviates from Goguen's componentwise definitions by making $p \supset q$ perfectly true whenever $[p]$ and $[q]$ are incomparable. This has the unfortunate result that $p \supset r$ can be less than perfectly true when $p \supset q$ and $q \supset r$ are perfectly true, e.g. if$[p] = 1, 1, 0$, $[q] = 0, 0, 1$ and $[r] = 1, 0, 0$.

41 Technically, Goguen allows the truth set to be any *closg*, defined as a complete lattice with an additional binary operation that is associative, distributes over arbitrary joins and has the lattice infinity as an identity (1969: 354).

42 *Proof:* Let N and N^* be the functions (as sets of ordered pairs) on [0, 1] taking α to $1 - \alpha$ and $(1 - \alpha^{2})^{1/2}$ respectively. Now $\alpha, \beta \in N^*$ just in case $\alpha^2, \beta^2 \in N$. Since squaring is an order-preserving mapping of [0, 1] onto itself, N and N^* obey the same constraints in terms of $\leqslant$. Yet $1/2, 1/2 \in N$ while $1/2, 1/2 \notin$ N*.

43 See Peacocke 1981: 137.

44 Section 7.2 will argue that $(\sim_1)$ makes the denial of bivalence for p incoherent, a result the degree theorist may not welcome.

45 See for example Bellman and Zadeh 1977. They describe their definition of *true* as 'entirely subjective as well as local in nature' (118). See also Haack 1978: 165–9, 1979.

46 See Fine 1975: 269–70, Kamp 1975: 131, Urquhart 1986: 108–9.

47 Contrast Forbes 1985: 173.

48 Bloch 1954: 173.

49 Although the axiomatization was found by Wajsberg in 1935, the first published proof is in Rose and Rosser 1958. See Rosser 1960 for a general survey of results in this area, and Morgan and Pelletier 1977 for a philosophical discussion.

50 Scarpellini 1962.

51 Compare Morgan and Pelletier 1977: 89–90.

52 Pavelka 1979 and Novák 1989: 128–43; also Gerla and Tortora 1990.

53 For a recent discussion see Shapiro 1991, which rebuts philosophical objections to second-order logic.

54 Several of the objections to fuzzy logic in Morgan and Pelletier 1977 assume the presence of the J_α operators in the object language. Without it, they say, 'the apparent many-valuedness is only illusory, since we cannot *say* anything in a

many-valued way' (86). However, what we say might *have* an intermediate degree of truth without being *about* intermediate degrees of truth.

55 Morgan and Pelletier 1977: 92.

56 Rosser and Turquette 1952.

57 Another difficulty with the idea that degrees of truth are identical if and only if they are indiscriminable is that indiscriminability, unlike identity, is non-transitive.

5 SUPERVALUATIONS

1 If cases for which no specification has been made *can* occur, the vagueness is *intensional*. If such cases *do* occur, it is also *extensional*. The distinction goes back to Carnap 1955. Compare Waismann on open texture and vagueness (Section 3.4).

2 The connections of theoretical terms to vague observational terms will also induce a vagueness in the former.

3 An interpretation in the relevant sense is not an attempt to say what is 'really meant', for we know that such an interpretation is more precise than what is 'really meant'. The word 'interpretation' may therefore be found misleading. In the case of vagueness, 'precisification', 'sharpening', 'delineation' and 'specification' have been used instead. Nevertheless, Mehlberg's choice of 'interpretation' will be followed; it carries a salutary reminder of the recalcitrance of the material.

4 A more general approach would take into account ways of making the language more but not perfectly precise. Indeed, one can secure the results of supervaluationism without appealing to ways of making the language perfectly precise, by assuming that a way of making the language more precise can always be extended enough to settle the truth-value of a given sentence (Fine 1975: 280). The arguments in the text can be adapted without much difficulty to these subtler forms of supervaluationism.

5 Mehlberg 1958: 277.

6 Mehlberg 1958: 257–8.

7 Not all uses of supervaluations in treating referential indeterminacy in science depend on a strict distinction between theory and observation. Later applications of the method in philosophy of science include Przelecki 1969: 20, 90–5 and Field 1973. Mehlberg attacks the view that vague statements are either true or false in borderline cases by arguing that it makes their truth-values unknowable and therefore truth an 'occult quality' (1958: 256). On his view, 'The number of trees in Toronto is even' remains verifiable, in the sense that it could have been known to be true or false, for that requires only that it could have been true or false, not that it actually is (Canada could have contained just two trees, both in the centre of Toronto). According to Mehlberg, when ordinarily vague statements lack a truth-value, they do so

contingently. In contrast, a completely vague statement essentially lacks a truth-value, and is therefore unverifiable.

8 Mehlberg 1958: 259.

9 Mehlberg 1958: 258.

10 Mehlberg 1958: 325.

11 van Fraassen 1966, 1968, 1971.

12 Lewis 1970: Appendix, Kamp 1975, Dummett 1975, Fine 1975, Przelecki 1976.

13 See Fine 1975: 269–70 and Kamp 1975: 131. Fine speaks of 'penumbral connections', arguing that supervaluationism is the only view to accommodate them all (278–9). It might be objected that a conjunction whose conjuncts are not wholly false is not itself wholly false (Forbes 1985: 173). However, 'wholly false' is slippery; one might say that a conjunction with some obviously true conjuncts and some obviously false ones is not wholly false – but it is still false.

14 In a trivial way, supervaluations can be seen as a form of many-valued semantics in which the value of a sentence is the function taking each admissible valuation to the truth-value of the sentence under that valuation. However, such values are not *degrees* of truth.

15 Dummett 1975 uses local validity; Fine 1975 uses global validity. Note also that some invalid arguments have counterexamples only in counterfactual situations, so truth preservation on all admissible interpretations is not sufficient for validity. Consider the argument from 'Oxford is a large city' to 'Birmingham is a large city'. Since Birmingham meets any reasonable requirement for being a large city met by Oxford, the conclusion is true on any admissible interpretation on which the premise is true. Yet the argument is not deductively valid, for in some logically possible situation Oxford is the largest city in the world and Birmingham the smallest. Thus validity depends on at least two dimensions of variation: between admissible variations and between possible situations.

16 Compare Dummett 1991: 77–80.

17 Inspection of the counterexample shows that contraposition does not hold in the logic of supervaluations even in the special case where no auxiliary premises are allowed. Other examples involve auxiliary premises. Similar remarks apply to the rules discussed in the following paragraph of the text.

18 Fine 1975: 290.

19 The failure of instances of conditional proof and *reductio ad absurdum* is used against supervaluationism by Machina 1976: 52–3.

20 Sorensen 1988b: 237–8 gives examples of sorites arguments in which the major premise is neither supertrue nor superfalse. Thus if the minor premise is supertrue, the conclusion is not superfalse. Such examples constitute no objection to supervaluationism.

21 For more detail on the objections to supervaluationism considered in this section see Fine 1975: 284–7, Sanford 1976: 205–7, Rolf 1984: 228–33 and Sorensen 1988b: 238–9.

22 Lewis 1983: 229 (acknowledging David Kaplan); Kamp 1975: 137–45; Edgington 1992.

23 On the crude account of comparative truth, '*a* is at least as *F* as *b*' is equivalent to 'Definitely if *b* is *F* then a is *F*'; since '*a* is *F*er than *b*' is equivalent to '*a* is at least as *F* as *b* and *b* is not at least as *F* as *a*', it too can be defined in terms of 'definitely'.

24 See also Pinkal 1983a.

25 The envisaged semantics provides putative counterexamples to conditional proof, argument by cases and *reductio ad absurdum.* The inference from 'Saul is brave' to 'Saul is at least as brave as David' is valid, but the conditional conclusion 'If Saul is brave then Saul is at least as brave as David' is invalid. The inferences from 'Saul is brave' and from 'Saul is not brave' to 'Either Saul is at least as brave as David or Saul is not brave' are valid, but the inference from 'Saul is brave or Saul is not brave' to the same conclusion is invalid. The inferences from 'Saul is brave and not at least as brave as David' to 'Saul is at least as brave as David' and to 'Saul is not at least as brave as David' are valid, but 'Saul is not both brave and not at least as brave as David' is invalid.

26 An alternative would be to count a sentence as 'true' if it is true on 'enough' admissible interpretations, a vague matter. Thus even if the admissible interpretations on which 'Saul is brave' is true are a proper subset of those on which 'David is brave' is true, there might be enough of the former to make 'Saul is brave' true. Unfortunately, this approach allows the conjunction of two true sentences not to be true; if '*A*' is true on 90 per cent of admissible interpretations and '*B*' on 90 per cent, '*A* and *B*' might be true on only 80 per cent. According to Lewis, we do treat 'true enough' as 'true', and occasionally get into trouble because some forms of inference preserve the property of being true without preserving the property of being true enough (1983: 244). The inference from two sentences to their conjunction would be an example. Moreover, this approach mishandles 'acute', for 'An angle of 89° is acute' is simply true and 'An angle of 89° is not acute' simply false, yet (by the natural measure) the latter is true on 89/90 of the admissible interpretations required to handle 'more acute than' and the latter on only 1/90 of them.

27 The best supervaluationist account of higher-order vagueness is Fine 1975: 287–98.

28 Conversely, the schema 'Either definitely definitely *A* or definitely not definitely *A*' is valid in S5.

29 Burgess 1990b argues heroically that vagueness stops at a finite order. However, the argument assumes that, for example, 'seems red', 'genuine doubt' and 'accepted provisionally into a conversation in the language in question' are precise (440, 445). One can think of borderline cases for each.

30 In effect, McGee 1991 treats admissibility as consistency with meaning postulates, and definiteness as provability from those postulates. The results are counter-intuitive, for reasons connected with Gödel's second incompleteness theorem. 'If definitely A then A' comes out invalid, and 'If definitely not definitely *A* then definitely both *A* and not *A*' valid (205–8). Yet *x* can be a definite borderline case of '*F*', even though sharpening of '*F*' would make *x* cease to be a borderline case. If the meaning postulates entail that *x* is a borderline case, the supervaluationist must hold that admissible interpretations can be inconsistent with meaning postulates; if the meaning postulates do not entail that *x* is a borderline case, they do not exhaust what is definite about the semantics of the language. Either way, admissibility is not a matter of consistency with the meaning postulates. Note also that, in the presence of higher-order vagueness, it is vague which sentences are meaning postulates. McGee does not provide an account of higher-order vagueness.

31 If the picture of semantic rules were maintained, one problem for supervaluationism is that a semantic rule might forbid certain borderline cases to be decided. Both a statement and its denial would then count as misuses.

32 There is a delicate issue about the supervaluationist definition of validity when the S4 schema fails. Global validity is supertruth-preservation. '*A*' is supertrue at a point *s* just in case '*A*' is true at all points admitted by *s*. Since admitting is non-transitive, '*A*' may be supertrue at a point when 'Definitely *A*' is not, contrary to the claim in Section 5.3 that the inference from '*A*' to 'Definitely *A*' is globally valid. Formally, it is more elegant to define validity in terms of all points in the space, but this is not supertruth-preservation if 'Definitely' is the object-language expression of supertruth. Related issues are discussed later in the text. Note that even if the inference from '*A*' to 'Definitely *A*' is not globally valid, the rules of contraposition, conditional proof, argument by cases and *reductio ad absurdum* still do not preserve supervalidity (one can show this by considering the inference from '*A* and not definitely *A*' to '*A* and not *A*', which is certainly globally valid).

33 Although the construction validates the S4 principle for 'definitely*', it does not validate the S5 schema, for it does not validate the Brouwerian schema 'If *A* then definitely* not definitely* not *A*', which follows from the S5 schema given the T schema. The Brouwerian principle requires admitting* to be symmetric, and this has not been guaranteed. Thus a kind of higher-order vagueness might remain. However, the Brouwerian principle for 'definitely' has some plausibility: 'If *A* then definitely not definitely not *A*'. It corresponds to the condition that a point admits only those points which admit it; why should *x* regard *y* as reasonable if *y* regards *x* as unreasonable? If admitting is a matter of points not differing by too much, one might expect it to be symmetric. The Brouwerian principle for 'definitely' implies that for 'definitely*' (a symmetric relation has a symmetric ancestral). Even if the Brouwerian principle fails for 'definitely', one could replace the non-symmetric 'admits' by the symmetric 'admits or is admitted by' and take the

ancestral of the latter, although this construction does not correspond to an explicit definition in terms of 'definitely'. A failure of the Brouwerian principle does not seem to be at the heart of the problem. Salmon 1989 discusses formally parallel issues for 'necessarily'.

34 Fine 1975: 297. Fine does not commit himself to the workability of the suggestion; he allows as an alternative that the set of admissible specifications is 'intrinsically vague'.

35 See also Heller 1990: 82–4, Sayward 1989, Day 1992.

36 Note the qualifications at n.33.

37 Fine 1975: 296.

38 Dummett 1991: 74 seems to claim a connection between objectivity and determinacy. That discussion neglects higher-order vagueness in other respects too. 'Definitely' is assumed to have an S5 semantics (78). An assertion of 'Definitely *A*' is said to be permissible just when the assertion of '*A*' is mandatory (79); but what is the condition for an assertion of 'Definitely *A*' to be mandatory? The importance of higher-order vagueness is acknowledged in Dummett 1975.

39 Liar-like paradoxes may cause one to doubt that any property meets the Tarskian condition. However, problems of self-reference and the like are too distant from those of mundane vagueness to make it plausible that the Tarskian condition cannot be met for a vague language without semantic vocabulary (see McGee 1991: 217 and Section 7.2).

40 If the classically valid patterns are maintained because they are locally valid, the equation of supertruth with truth is further undermined, for that standard of validity is not equivalent to supertruth-preservation.

6 NIHILISM

1 The distinction between global and local nihilism is not related to that between global and local validity in Chapter 5.

2 For discussion of views close to global nihilism see Dummett 1975; Wright 1975, 1976, 1987, 1991, 1992; Wheeler 1975, 1979; Unger 1979a, b, c, 1980 (contrast Unger 1990: 321–3); Sanford 1979; Quine 1981; Grim 1982, 1983, 1984; Abbott 1983; Rolf 1984; Heller 1988, 1990; Sorensen 1988b: 226–9; Sainsbury 1991; Horgan, forthcoming.

3 A sorites argument is typical if it is relevantly like the Heap and the Bald Man. By everyday standards, it should be formulated without equivocation, the minor premise should be obvious, the key expression vague, the difference between successive members of the series sufficiently small, and so on.

4 Some writers with global nihilist tendencies offer more positive arguments. Thus Wheeler 1975, 1979 combines a causal theory of reference with the claim that vague predicates are not law-governed in the way needed for them to be causally related to properties. 'All tall men are over five foot three' is rejected

as a law because it would be true even if the extension of 'tall' were slightly varied (1975: 373). It is not clear how many scientific laws would pass this test.

5 The validity of the sorites argument is not at issue in this context. The charge of invalidity depends on the idea that if the major premise is less than perfectly true, the conclusion can have a lower truth-status than either premise. If the major premise is perfectly true, each conditional instantiating it will be perfectly true, so its consequent will have at least as high a truth-status as its antecedent on any reasonable semantics.

6 The all-or-nothing view may acknowledge some limits. There is a sorites argument that all colours are red, perhaps not that all sounds are.

7 The qualification 'without change of meaning' will be understood throughout.

8 The individuation of properties is assumed to be coarse-grained (e.g. by necessary equivalence).

9 One might say that 'integer not much greater than 104 and not much less than 106' is a precise expression defined in terms of vague ones (Sorensen 1988a, b: 229; Rolf 1980). However, this would not show that the global nihilist's stipulations can begin with vague expressions, for if 'not much greater' and 'not much less' are incoherent, so are expressions composed out of them.

10 For discussion of related forms of local nihilism see Dummett 1975, 1979; Wright 1975, 1976, 1987, 1991; Peacocke 1981; Travis 1985; Burns 1986, 1991; Schwartz 1989.

11 Compare Jackson 1977: 113–14.

12 The observer might be held not to be warranted in asserting that even x_0 is *perfectly* square. The minor premise of the envisaged sorites argument would then be false. This would just make the example a bad one. The underlying problem could still be raised with 'not square' in place of 'square', taking the series in reverse order, for the observer is certainly warranted in asserting that x_n is not square.

13 Notions of the look of a thing for which naked eye indiscriminability is necessary but not sufficient can be developed using the methods of Williamson 1990a: 65–87, 109–13. It allows the looks of *x* and *y* to be identical even if, for some *z*, *x* is discriminable from *z* and *y* is not. Note that a notion of the look of a thing for which indiscriminability was sufficient but not necessary would be too broad to be of interest. One can develop similar notions of identity of apparent shape for which apparent identity of shape is necessary but not sufficient. However, indiscriminability in shape and apparent identity of shape are not equivalent; in the Müller–Lyer illusion two lines falsely appear to be of unequal length; they are not really discriminable in length – they only appear to be. Sorites arguments turning on the non-transitivity of indiscriminability are treated in detail in Williamson 1990a: 43–7, 88–103.

14 'If' in (1)–(3) is to be read as the material conditional, whose transitivity is not here in dispute; if (1) and (2) are perfectly true, so is (3).

15 The assumption in the text is more radically simplistic than is suggested there; see Johnston 1992 for some relevant examples. Note, however, that the

assumption does not entail that something is not red if and only if in normal conditions all normal observers always classify it as not red. More generally, such equivalences for colour terms do not entail the corresponding equivalences for their negations. Such assumptions are therefore consistent with bivlence and the law of excluded middle, even given that in normal conditions normal observers sometimes disagree about marginal cases.

16 The remark in the text applies to a set of responses that includes 'This is is red' and is closed under truth-functions and the operator 'For all I know'; it obviously does not apply to responses such as 'I have looked at this for at least a minute'.

7 VAGUENESS AS IGNORANCE

1 As noted in Section 1.2, the epistemic view probably goes back to the Stoics. More recently, it has been defended in Cargile 1969, 1979, Campbell 1974, Scheffler 1979, Sorensen 1988b: 217–52, 1991, Horwich 1990: 81–7. Przelecki 1979 (a paper originally published in Polish in 1964) discusses an apparently similar view. For critical discussion see Heller 1990: 89–106, Simons 1992 and Horgan, forthcoming.

2 One can strengthen (B) to a universal generalization by universally quantifying '*u*' and '*P*'. If the quantification is not substitutional, it can avoid the restriction to what is expressible in English. The argument in the text does not need this added strength. Strawson 1950 and McDowell 1982 defend similar conceptions of bivalence.

3 Tarski quotes *Metaphysics* Γ 7.27 in 'The concept of truth in formalized languages' (1983: 155).

4 Horwich 1990: 80 has a somewhat similar argument.

5 A disquotational schema for truth in a vague language is endorsed by, for example, Evans and McDowell 1976: xi; Machina 1976: 75; Peacocke 1981: 136–7. See also Section 5.7.

6 Tarski 1983: 197 derives bivalence (which he calls 'the principle of excluded middle') as Theorem 2 of 'The concept of truth in formalized languages'. The proof uses the law of excluded middle (in the present sense).

7 A formal version of the argument is as follows. Each formula '*P*' is assigned a semantic value $[P]$. The semantics values form a lattice under a partial ordering $\leqslant$, i.e. each pair of values has a greatest lower bound (glb) and least upper bound (lub). $[P \wedge Q] = \text{glb}\,\{[P], [Q]\}$; $[P \vee Q] = \text{lub}\{[P], [Q]\}$; if $[P] \leqslant [Q]$ then $[{\sim}Q] \leqslant [{\sim}P]$. These assumptions are met by standard classical, supervaluational, intuitionist and many-valued treatments, and others. It is

then easy to show that $[T(u)] = [P]$ and $[F(u)] = [\sim P]$ imply $[\sim|T(u) \vee F(u)|] \leqslant [\sim P \wedge \sim\sim P]$.

8 See for example Simons 1992: 175, n.16.

9 Strictly speaking, the remark in the text requires the biconditionals to be clear, not just true.

10 Compare Dummett's condition for the objectivity of truth (1991: 74).

11 The intuitionist is assumed to equate 'true' with 'provable' rather than with 'proved'.

12 A sorites paradox might move some to deny that all sentences of the form '*n* grains make a heap' are bivalent without denying bivalence for any particular *n* (although for some *n* they would refrain from asserting it), but see Section 7.2 (b).

13 Putnam 1983 suggests that intuitionist logic is appropriate for handling sorites paradoxes. However, he admits that intuitionist semantics is inappropriate (Read and Wright 1985, Putnam 1985). The use of the logic therefore seems *ad hoc*, since it is not connected to an account of the meaning of vague terms. For further discussion see Schwartz 1987, 1990, Schwartz and Throop 1991, Putnam 1991: 413–14 and Rea 1989. Although Black 1949: 37 (reprinting Black 1937) has been credited with making the first connection between vagueness and intuitionist logic, he treats the latter simply as another example of an alternative logic, and contrasts its state of development (favourably) with that of his own logic of vagueness.

14 Dummett 1991: 74–82 gives a contrary view.

15 The claim that utterances that say something are true or false may be compared with the acceptance of bivalence for formulas that express propositions in Kripke 1975.

16 Compare Dummett 1991: 348–51.

17 A more teleological question would be 'Which properties did this mechanism evolve to register?'. Considerations like those in the text would still apply.

18 See Sorensen 1991.

19 Understanding a sentence may involve more than knowing what it means, e.g. knowing what its constituents mean. Moreover, such knowledge about a sentence is not in general sufficient for knowledge of what was said by an utterance of it; the latter may depend on the context too. These points do not affect the argument in the text.

20 The account of understanding has similarities to that of Soames 1989, on which to understand the sentences of a language is to satisfy conventional standards for their use. However, an epistemic view of vagueness is also consistent with some accounts of understanding that assign primacy to knowledge of reference (e.g. Higginbotham 1992 argues that understanding is a matter of knowing *what one is expected to know* about reference).

21 Hart 1991–2, which supplies many necessary qualifications: e.g. the stability should not depend on the grains all being cubes.

22 Sperber and Wilson 1985–6.

23 For accounts emphasizing connections between vagueness and context dependence see Lewis 1979, Kamp 1981, Pinkal 1983b, Burns 1991.

8 INEXACT KNOWLEDGE

1 The conditionals in the argument may be read as truth-functional.

2 Compare Nozick 1981: 240–2. The relevant notion is Nozick's knowledge^, knowledge expanded enough to satisfy deductive closure, rather than his knowledge*, knowledge contracted enough to satisfy it. However, on Nozick's definition, if one knows that *A* and knows that *A* entails *B*, then one knows^ that *B*, even if one has not drawn the conclusion that *B* (perhaps because the two items of knowledge have been kept separate) and fails to believe that *B*. The argument in the text requires only a concept that extends to believed propositions deduced from known ones.

3 I do not know what the number *n* is for which I fail the KK principle. That is not surprising, for of necessity one cannot knowingly identify a particular failure of the KK principle in the first person present tense. If I know that I both know that *A* and do not know that I know that *A*, then I must know the first conjunct of that conjunction (given that knowledge of a conjunction entails knowledge of its conjuncts), i.e. I must know that I know that *A*, so the second conjunct is false, so I do not know the conjunction (since knowledge entails truth and the truth of a conjunction entails the truth of its conjuncts). This point may help to explain the seductiveness of the KK principle. For a survey of earlier discussion of the principle, see Lenzen 1978: 69–77. Later work includes Wiggins 1979, Sorensen 1987, 1988b: 242–3, 313–17 and Mellor 1991: 261–3.

4 It can be proved that, in a simple logic in which knowledge is assumed only to be deductively closed and truth-entailing, it is consistent to suppose 'There are exactly *k* people in the stadium and I have i_j iterations of knowledge that there are not exactly j people, i_{j+1} iterations of knowledge that if there are exactly $j + 1$ then I do not know that there are not exactly $j, \dots,$ and i_k iterations of knowledge that if there are exactly *k* then I do not know that there are not exactly $k - 1$' if and only if for some n, $i_n < k - n$ ($j \leqslant n \leqslant k$; Williamson 1992a: 240–1).

5 The argument assumes that reliability is a necessary condition for knowledge. It does not assume that the concept of knowledge can be analysed in terms of the concept of reliability, for it does not assume that further necessary conditions can be non-circularly stated in the presence of which the reliability condition is sufficent as well as necessary for knowledge (reliability may be sufficient for knowledge granted some background conditions, but the latter need not be necessary for knowledge). Someone who wants to analyse the concept of knowledge in terms of (for example) the concept of justification

should regard the reliability condition not as unnecessary for knowledge but as a consequence of the justification condition. Moreover, ‘knowledge’ may resemble most other words in expressing a concept with no non-trivial analysis. Note also that although reliability may be a probabilistic notion, a given margin for error is preserved under logical consequence (contrast the probabilistic objection rebutted in Section 8.2); see n.7.

6 For simplicity, the text temporarily ignores all obstacles to knowledge other than inexactness.

7 If every point less than an inch from the point of impact is within each of several areas, every such point is within the intersection of those areas, and so within any area which includes that intersection. Correspondingly, if each of some propositions leaves a margin for error, then so do their joint logical consequences.

8 The whole wall is treated as having no margin, the possibility that the bullet lands off the wall being disregarded.

9 The argument in the text shows why it would be futile to restore the KK principle for a new concept, knowledge^{+}, such that one knows^{+} that *A* just in case one knows that *A*, knows that one knows that *A*, knows that one knows that one knows that *A*, ... One knows^{+} almost no non-trivial empirical propositions.

10 Margin for error principles are applied to the Paradox of the Surprise Examination and to backward induction arguments about Iterated Prisoner’s Dilemmas in Williamson 1992a.

11 One cannot know physical measurements with perfect precision, but the problem with ‘thin’ would remain even if one could.

12 As ‘word’ is used in the text, words are not individuated by their meanings. Such a use is clearly legitimate (although it is not the only one), as when we say that a word has changed its meaning since 1600.

13 As always ‘*n* grains make a heap’ is to mean that there can be a heap of *n* grains.

14 The relation between vagueness and the non-transitivity of indiscriminability is discussed in Dummett 1975, 1979; Wright 1975, 1976, 1987; Platts 1979; Peacocke 1981; Travis 1985; Burns 1986, 1991; Schwartz 1989; Williamson 1990a.

15 Indirect discrimination is used, and connected with vagueness, in Russell 1927: 280–1. For more explicit accounts see Russell 1950: 104–5; Goodman 1951: 196–200; Dummett 1975. Burns 1986 and 1991 make much of indirect discrimination.

16 Discrimination is analysed as the activation of knowledge of distinctness in Williamson 1990a: 5–8.

17 The notion of ‘most worlds’ does not depend solely on the number of worlds in the relevant sets; it depends, like the notion of probability, on a measure of the size of sets.

18 See Williamson 1992a: 237–9 for a two-dimensional model of inexact beliefs.

9 VAGUENESS IN THE WORLD

1 Dummett 1978: 260.
2 Dummett 1981: 440.
3 Russell 1923 (1983–, vol. 9: 148).
4 Bradley 1893: 359–400, 'Degrees of truth and reality'.
5 Burks 1946: 483.
6 The usual suggestion is that identity can be indeterminate, rather than that it comes in degrees. The two ideas share a rejection of the view that things must be either perfectly identical or perfectly distinct.
7 Parfit 1971, 1984: 238–41.
8 Kripke 1972: 345 n.18. Kripke also discusses the view that strict identity does not allow change in composition, and is not vague, what allows change in composition and is vague being a relation not of identity but of counterparthood, perhaps involving the strict identity of some part of the composing matter. He doubts that one could find the required level of basic particulars at which identity is not vague.
9 Evans 1978; Salmon 1982: 243–6, 1984, 1986a: 110–14. See Thomason 1982; Noonan 1982, 1984, 1990; Broome 1984; Over 1984, 1989; Wiggins 1986; Rasmussen 1986; Cook 1986; Parsons 1987; Garrett 1988, 1991; van Inwagen 1988, 1990; Lewis 1988; Stalnaker 1988; Pelletier 1989; Burgess 1989, 1990a; Johnsen 1989; Tye 1990; Cowles and White 1991; Engel 1991: 196–8, 213–15; Howard-Snyder 1991; Zemach 1991.
10 'Paris' is assumed for the sake of argument to be determinate in reference.
11 A supervaluational treatment goes well with Evans's analogy between the operators 'Definitely' and 'Indefinitely' and the modal operators 'Necessarily' and 'Possibly'. He calls the former pair duals, as the latter are. Inconsistently, he also treats 'Indefinitely *A*' as 'It is not determinate whether *A*'; to preserve duality and the modal analogy, it should be read as 'It is not determinate that not *A*'. He also extends the argument by appeal to a logic corresponding to the modal logic S5; this would make difficulties for the treatment of higher-order vagueness.
12 Someone who holds 'Europe' to be indeterminate in reference between several precisely demarcated regions is not, however, automatically debarred from treating '"Europe" refers to Europe' as determinately true. Supervaluations might be applied to the meta-language under a constraint linking the precisification of 'Europe' to the precisification of 'refers to'.
13 For more on related issues see Rolf 1980; Heller 1988, 1990; Sainsbury 1989; Burgess 1990a; van Inwagen 1990: 213–83.
14 Since representations can themselves be represented, the notions of vagueness* and precision* apply to representations too. However, one should not assume that all vague representations are vague* or that all precise ones are precise*. Whether a representation is vague or precise depends on how it represents; whether it is vague* or precise* depends on how it is represented.

15 One could also require that the removal of the unclarity in question not be at the expense of introducing a new unclarity.

16 Strictly speaking, the *de re*/*de dicto* distinction must be relativized to particular occurrences of a term and a sentence functor in a particular sentence. It is obvious in what follows which relativization is intended.

17 Even if the inference from *de dicto* to *de re* were valid, the two would still be distinguished by the failure of the converse inference.

18 The notion of 'knowing which' is highly sensitive to context. Examples like those in the text suggest that it is nevertheless not wholly illusory.

19 Knowing of 1453 that it was before 1460 is not claimed to entail knowing that 1453 was before 1460. Perhaps someone did the former without doing the latter when in 1460 she thought of a year that was in fact 1453 by remembering her experiences in that year without being able to date it.

20 The existential quantifier implicit in the form *de re* makes the equivalences not automatic. However, they can still be proved from the definitions and the corresponding equivalences *de dicto*. For 'unclear', the argument uses the factiveness of 'clear'.

21 Since 'a' and 'b' express ways of thinking *de re* of x, the non-equivalence of 'It is clear whether a is F' and 'It is clear whether b is F' goes beyond the failures of substitution in the initial sketch of the *de re*/*de dicto* distinction, which supposed one of the terms ('the year Constantinople fell') not to be associated with a way of thinking *de re* of the year. Failures of the former kind do seem to occur. If 'Hesperus' and 'Phosphorus' are defined by pointing at Venus in the evening and morning respectively, both are associated with ways of thinking *de re* of Venus; nevertheless, it may apparently be known that Hesperus is visible in the evening and not known that Phosphorus is visible in the evening. Some direct reference theorists who hold that names refer directly resist this conclusion, holding that 'Hesperus is visible in the evening' and 'Phosphorus is visible in the evening' express the very same knowledge (e.g. Salmon 1986b). However, even these theorists allow that the two sentences are associated with different *guises* of the proposition; it is known under one and not under the other. For present purposes, it could be stipulated that 'It is clear that A' means that the proposition that A is clear under the guise associated with 'A' (the context would therefore be implicitly quotational). This will allow for the failure of substitution even on the direct reference view. 22 To revise the explanation of unclarity *de re* by adding the requirement that if $x = y$ then 'a' and 'b' express the same way of thinking would be *ad hoc* and pointless and would lose a useful distinction captured by the present notation.

22 Suppose that Stac Polly is a mountain (a parallel argument applies if not). Let 'F' abbreviate 'such that it is a mountain if and only if in the actual world Stac Polly is a mountain'. It is clear that Stac Polly is F, for it is *a priori* that Stac Polly is such that it is a mountain if and only if in the actual world Stac Polly is a mountain. In the actual world world Stac Polly is a mountain, so it is necessary that *in the actual world* Stac Polly is a mountain. Hence, necessarily,

'*F*' (as uttered in the actual world) is true of all and only things that have the property of mountainhood. On the view in question, to understand '*F*' is therefore to think *de re* of the property of mountainhood. Thus it is clear of the property of mountainhood that Stac Polly has it.

APPENDIX: THE LOGIC OF CLARITY

1 Hughes and Cresswell 1984: 28–9 (B = KTB).
2 Williamson 1992b.
3 Apply Williamson 1992b or Chellas and Segerberg 1994 to the theorem in the text.
4 Hughes and Cresswell 1984: 28 (T = KT).
5 It is sufficient if all countable variable margin models are intended models of the logic of clarity, for since KT has the finite model property (Hughes and Cresswell 1984: 142–3), *W* in the completeness leg of the proof of the theorem may be taken to be finite, in which case *W** will be countable by construction.
6 Williamson 1990b. The result does not extend to KTB.

References

Abbott, W.R. 1983. ‘A note on Grim’s sorites argument’, *Analysis* 43: 161–4.

Aldrich, V. 1937. ‘Some meanings of “vague”’, *Analysis* 4: 89–95.

Alston, W.P. 1964. *Philosophy of Language*, Englewood Cliffs, N.J., Prentice-Hall.

Annas, J. 1990. ‘Stoic epistemology’, in S. Everson (ed.) *Companions to Ancient Thought 1: Epistemology*, Cambridge, Cambridge University Press.

Åqvist, L. 1962. ‘Reflections on the logic of nonsense’, *Theoria* 28: 138–57.

Armstrong, S.L., Gleitman, L.R. and Gleitman, H. 1983. ‘What some concepts might not be’, *Cognition* 13: 263–308.

Austin, J.L. 1962. *Sense and Sensibilia*, ed. G.J. Warnock, Oxford, Oxford University Press.

Ayer, A.J. 1940. *The Foundations of Empirical Knowledge*, London, Macmillan.

Bain, A. 1870. *Logic*, London, Longman.

Baldwin, J.M. 1901–2. *Dictionary of Philosophy and Psychology*, 2 vols, London, Macmillan.

Barnes, J. 1982. ‘Medicine, experience and logic’, in J. Barnes, J. Brunschwig, M.F. Burnyeat and M. Schofield (eds) *Science and Speculation*, Cambridge, Cambridge University Press.

Bellman, R.E. and Zadeh, L.A. 1977. ‘Local and fuzzy logics’, in J.M. Dunn and G. Epstein (eds) *Modern Uses of Multiple-Valued Logic*, Dordrecht, Reidel.

Benjamin, A.C. 1939. ‘Science and vagueness’, *Philosophy of Science* 6: 422–31.

Beth, E.W. 1959. *The Foundations of Mathematics*, Amsterdam, North-Holland.

Black, M. 1937. ‘Vagueness: an exercise in logical analysis’, *Philosophy of Science* 4: 427–55; reprinted in his *Language and Philosophy*, Ithaca, N.Y., Cornell University Press, 1949.

Black, M. 1949. *Language and Philosophy*, Ithaca, N.Y., Cornell University Press.

Black, M. 1963. ‘Reasoning with loose concepts’, *Dialogue* 2: 1–12; reprinted in his *Margins of Precision*, Ithaca, N.Y., Cornell University Press, 1970.

Black, M. 1970. *Margins of Precision*, Ithaca, N.Y., Cornell University Press.

Bloch, M. 1954. *The Historian’s Craft*, tr. P.Putnam, Manchester, Manchester University Press.

Bochvar, D.A. 1981. 'On a three-valued logical calculus and its application to the analysis of the paradoxes of the classical extended functional calculus', tr. M. Bergmann, *History and Philosophy of Logic* 2: 87–112.

Bradley, F.H. 1893. *Appearance and Reality*, London, George Allen.

Bridge, J. 1977. *Beginning Model Theory*, Oxford, Clarendon Press.

Brock, J. 1979. 'Principal themes in Peirce's logic of vagueness', *Peirce Studies* 1: 41–9.

Brock, J. 1980. 'Peirce's anticipation of game-theoretic logic and semantics', in M. Herzfield and M.D. Lenhart *Semiotics 1980*, New York, Plenum.

Broome, J. 1984. 'Indefiniteness in identity', *Analysis* 44: 6–12.

Burge, T. 1990. 'Frege on sense and linguistic meaning', in D. Bell and N. Cooper (eds) *The Analytic Tradition*, Oxford, Blackwell.

Burgess, J.A. 1989. 'Vague identity: Evans misrepresented', *Analysis* 49: 112–19.

Burgess, J.A. 1990a. 'Vague objects and indefinite identity', *Philosophical Studies* 59: 263–87.

Burgess, J.A. 1990b. 'The sorites paradox and higher-order vagueness', *Synthese* 85: 417–74.

Burgess, J.A. and Humberstone, I.L. 1987. 'Natural deduction rules for a logic of vagueness', *Erkenntnis* 27: 197–229.

Burks, A.W. 1946. 'Empiricism and vagueness', *Journal of Philosophy* 43: 477–86.

Burns, L.C. 1986. 'Vagueness and coherence', *Synthese* 68: 487–513.

Burns, L.C. 1991. *Vagueness: An Investigation into Natural Languages and the Sorites Paradox*, Dordrecht, Kluwer.

Burnyeat, M.F. 1982. 'Gods and heaps', in M. Schofield and M.C. Nussbaum (eds) *Language and Logos*, Cambridge, Cambridge University Press.

Campbell, R. 1974. 'The sorites paradox', *Philosophical Studies* 26: 175–91.

Cargile, J. 1969. 'The sorites paradox', *British Journal for the Philosophy of Science* 20: 193–202.

Cargile, J. 1979. *Paradoxes: A Study in Form and Predication*, Cambridge, Cambridge University Press.

Carnap, R. 1955. 'Meaning and synonymy in natural languages', *Philosophical Studies* 6: 33–47; reprinted in his *Meaning and Necessity*, Chicago, Ill., University of Chicago Press, 2nd edn, 1956.

Carruthers, P. 1990. *The Metaphysics of the Tractatus*, Cambridge, Cambridge University Press.

Casari, E. 1987. 'Comparative logics', *Synthese* 73: 421–49.

Chellas, B.F. and Segerberg, K. 1994. 'Modal logics with the MacIntosh rule', *Journal of Philosophical Logic* 23.

Chisholm, R.M. 1942. 'The problem of the speckled hen', *Mind* 51: 368–73.

Cleave, J.P. 1970. 'The notion of validity in logical systems with inexact predicates', *British Journal for the Philosophy of Science* 21: 269–74.

Cohen, M.R. 1927. 'Concepts and twilight zones', *Journal of Philosophy* 24: 673–83.

Cook, M. 1986. 'Indeterminacy of identity', *Analysis* 46: 179–86.

Copilowish, I.M. 1939. 'Border-line cases, vagueness, and ambiguity', *Philosophy of Science* 6: 181–95.

Cowles, D.W. and White, M.J. 1991. 'Vague objects for those who want them', *Philosophical Studies* 63: 203–16.

Cresswell, M.J. 1976. 'The semantics of degree', in B.H. Partee (ed.) *Montague Grammar*, New York, Academic Press.

Day, T.J. 1992. 'Excluded middle and bivalence', *Erkenntnis* 37: 93–7.

Deas, R. 1989. 'Sorensen's sorites', *Analysis* 49: 26–31.

Dubois, D. and Prade, H. 1988. 'An introduction to possibilistic and fuzzy logics', in P. Smets, E.H. Mamdani, D. Dubois and H. Prade (eds) *Non-Standard Logics for Automated Reasoning*, London, Academic Press.

Dummett, M.A.E. 1968–9. 'The reality of the past', *Proceedings of the Aristotelian Society* 69: 239–58.

Dummett, M.A.E. 1975. 'Wang's paradox', *Synthese* 30: 301–24; reprinted in his *Truth and Other Enigmas*, London, Duckworth, 1978.

Dummett, M.A.E. 1978. *Truth and Other Enigmas*, London, Duckworth.

Dummett, M.A.E. 1979. 'Common sense and physics', in G.F. Macdonald (ed.) *Perception and Identity*, London, Macmillan.

Dummett, M.A.E. 1981. *The Interpretation of Frege's Philosophy*, London, Duckworth.

Dummett, M.A.E. 1991. *The Logical Basis of Metaphysics*, London, Duckworth.

Edgington, D. 1992. 'Validity, uncertainty and vagueness', *Analysis* 52: 193–204.

Engel, P. 1992. 'Les concepts vagues sont-ils sans frontières?', *Revue*

Engel, P. 1991. *The Norm of Truth*, London, Harvester Wheatsheaf. *Internationale de Philosophie* 46: 527–38.

Engel, R.E. 1989. 'On degrees', *Journal of Philosophy* 86: 23–37.

Engel-Tiercelin, C. 1986. 'Le vague est-il réel? Sur le réalisme de Peirce', *Philosophie* 10: 69–96.

Evans, M.G.J. 1978. 'Can there be vague objects?', *Analysis* 38: 208; reprinted in his *Collected Papers*, Oxford, Clarendon Press, 1985.

Evans, M.G.J. 1982. *The Varieties of Reference*, Oxford, Clarendon Press.

Evans, M.G.J. and McDowell, J.H. (eds) 1976. *Truth and Meaning*, Oxford, Oxford University Press.

Field, H. 1973. 'Theory change and the indeterminacy of reference', *Journal of Philosophy* 70: 462–81.

Fine, K. 1975. 'Vagueness, truth and logic', *Synthese* 30: 265–300.

Fisch, M.H. and Turquette, A.R. 1966. 'Peirce's triadic logic', *Transactions of the Charles S. Peirce Society* 2: 71–85; reprinted in M.H. Fisch (ed.) *Peirce, Semeiotic, and Pragmatism*, Bloomington, Ind., Indiana University Press, 1986.

Forbes, G. 1983. 'Thisness and vagueness', *Synthese* 54: 235–59.

Forbes, G. 1985. *The Metaphysics of Modality*, Oxford, Clarendon Press.

Fraassen, B.C. van 1966. 'Singular terms, truth-value gaps and free logic', *Journal of Philosophy* 63: 481–95.

Fraassen, B.C. van 1968. 'Presupposition, implication and self-reference', *Journal of Philosophy* 65: 136–52.
Fraassen, B.C. van 1971. *Formal Semantics and Logic*, London, Macmillan.
Frede, M. 1983. 'Stoics and skeptics on clear and distinct impressions', in M.F. Burnyeat (ed.) *The Skeptical Tradition*, Berkeley, Calif., University of California Press.
Frege, G. 1960. *Translations from the Philosophical Writings of Gottlob Frege*, ed. P. Geach and M. Black, 2nd edn, Oxford, Blackwell.
Frege, G. 1979. *Posthumous Writings*, ed. H. Hermes, F. Kambartel and F. Kaulbach, tr. P. Long and R. White, Oxford, Blackwell.
Frege, G. 1980. *Philosophical and Mathematical Correspondence*, ed. G. Gabriel, H. Hermes, F. Kambartel, C. Thiel, A. Veraart and B. McGuinness, tr. H. Kaal, Oxford, Blackwell.
Fuhrmann, G. 1988a. '"Prototypes" and "fuzziness" in the logic of concepts', *Synthese* 75: 317–47.
Fuhrmann, G. 1988b. 'Fuzziness of concepts and concepts of fuzziness', *Synthese* 75: 349–72.
Fuhrmann, G. 1991. 'Note on the integration of prototype theory and fuzzy-set theory', *Synthese* 86: 1–27.
Garrett, B.J. 1988. 'Vagueness and identity', *Analysis* 48: 130–4.
Garrett, B.J. 1991. 'Vague identity and vague objects', *Noûs* 25: 341–51.
Geach, P.T. 1956. 'The law of excluded middle', *Aristotelian Society* suppl. 30: 59–73.
Gerla, G. and Tortora, R. 1990. 'Fuzzy natural deduction', *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 36: 67–77.
Goguen, J.A. 1969. 'The logic of inexact concepts', *Synthese* 19: 325–73.
Goodman, N. 1951. *The Structure of Appearance*, Cambridge, Mass., Harvard University Press.
Grim, P. 1982. 'What won't escape sorites arguments', *Analysis* 42: 38–43.
Grim, P. 1983. 'Is this a swizzle stick which I see before me?', *Analysis* 43: 164–6.
Grim, P. 1984. 'Taking sorites arguments seriously: some hidden costs', *Philosophia* 14: 251–72.
Haack, S. 1974. *Deviant Logic*, Cambridge, Cambridge University Press.
Haack, S. 1978. *Philosophy of Logics*, Cambridge, Cambridge University Press.
Haack, S. 1979. 'Do we need "fuzzy logic"?', *International Journal of Man–Machine Studies* 11: 437–45.
Halldén, S. 1949. *The Logic of Nonsense*, Uppsala, Uppsala Universitets Arsskrift.
Hamilton, W. 1860. *Lectures on* Logic, 2 vols, ed. H.L. Mansel and J. Veitch, Edinburgh, Blackwood.
Hart, W.D. 1991–2. 'Hat-tricks and heaps', *Philosophical Studies* (Dublin) 33: 1–24.
Hegel, G.W.F. 1892. *Lectures on the History of Philosophy*, tr. E.S. Haldane and F.H. Simson, London, Kegan Paul.
Hegel, G.W.F. 1975. Logic, tr. W. Wallace, Oxford, Clarendon Press.

Heijenoort, J. van (ed.) 1967. *From Frege to Godel*, Cambridge, Mass., Harvard University Press.

Heijenoort, J. van 1985. 'Frege and vagueness', in L. Haaparanta and K.J.J. Hintikka (eds) *Frege Synthesized*, Dordrecht, Reidel.

Heller, M. 1988. 'Vagueness and the standard ontology', *Noûs* 22: 109–31.

Heller, M. 1990. *The Ontology of Physical Objects: Four-Dimensional Hunks of Matter*, Cambridge, Cambridge University Press.

Hempel, C.G. 1939. 'Vagueness and logic', *Philosophy of Science* 6: 163–80.

Higginbotham, J. 1992. 'Truth and understanding', *Philosophical Studies* 65: 3–16.

Hilpinen, R. 1982. 'On C.S. Peirce's theory of the proposition: Peirce as a precursor of game-theoretical semantics', *The Monist* 65: 182–8.

Hintikka, K.J.J. 1962. *Knowledge and Belief*, Ithaca, N.Y., Cornell University Press.

Hintikka, K.J.J. 1973. *Logic Language-Games and Information*, Oxford, Clarendon Press.

Hodges, W. 1985–6. 'Truth in a structure', *Proceedings of the Aristotelian Society* 86: 135–51.

Hookway, C. 1985. *Peirce*, London, Routledge & Kegan Paul.

Hookway, C. 1990. 'Vagueness, logic and interpretation', in D. Bell and N. Cooper (eds) *The Analytic Tradition*, Oxford, Blackwell.

Horgan, T., forthcoming. 'Robust vagueness and the forced march sorites paradox', *Philosophical Perspectives.*

Horwich, P. 1990. *Truth*, Oxford, Blackwell.

Howard-Snyder, F. 1991. '*De re* modality entails *de re* vagueness', *Pacific Philosophical Quarterly* 72: 101–12.

Hughes, G.E. and Cresswell, M.J. 1984. *A Companion to Modal Logic*, London, Methuen.

Humberstone, I.L. 1988. 'Some epistemic capacities', *Dialectica* 42: 183–200.

Inwagen, P. van 1988. 'How to reason about vague objects', *Philosophical Topics* 16: 255–84.

Inwagen, P. van 1990. *Material Beings*, Ithaca, N.Y., Cornell University Press.

Jackson, F. 1977. *Perception*, Cambridge, Cambridge University Press.

Jardine, L. 1977. 'Lorenzo Valla and the intellectual origins of humanist dialectic', *Journal of the History of Philosophy* 15: 143–64.

Jardine, L. 1982. 'Humanism and the teaching of logic', in N. Kretzmann, A. Kenny and J. Pinborg (eds) *The Cambridge History of Later Medieval Philosophy*, Cambridge, Cambridge University Press.

Jardine, L. 1983. 'Lorenzo Valla: academic skepticism and the new humanist dialectic', in M.F. Burnyeat (ed.) *The Skeptical Tradition*, Berkeley, Calif., University of California Press.

Jardine, L. 1988. 'Humanistic logic', in C.B. Schmitt, Q. Skinner and E. Kessler (eds) *The Cambridge History of Renaissance Philosophy*, Cambridge, Cambridge University Press.

Johnsen, B. 1989. 'Is vague identity incoherent?', *Analysis* 49: 103–12.

Johnston, M. 1992. ‘How to speak of the colors’, *Philosophical Studies* 68: 221–63.

Kamp, H. 1975. ‘Two theories about adjectives’, in E.L. Keenan (ed.) *Formal Semantics of Natural Language*, Cambridge, Cambridge University Press.

Kamp, H. 1981. ‘The paradox of the heap’, in U. Mönnich (ed.) *Aspects of Philosophical Logic*, Dordrecht, Reidel.

Kaplan, A. 1964. *The Conduct of Inquiry*, San Francisco, Calif., Chandler.

Kaplan, A. and Schott, H.F. 1951. ‘A calculus for empirical classes’, *Methodos* 3: 165–88.

Kerferd, G.B. 1978. ‘What does the wise man know?’, in J.M. Rist (ed.) *The Stoics*, Berkeley, Calif., University of California Press.

Khatchadourian, H. 1962. ‘Vagueness’, *Philosophical Quarterly* 12: 138–52.

King, J.L. 1979. ‘Bivalence and the sorites paradox’, *American Philosophical Quarterly* 16: 17–25.

Kleene, S.C. 1938. ‘On a notation for ordinal numbers’, *Journal of Symbolic Logic* 3: 150–5.

Klein, E. 1980. ‘A semantics for positive and comparative adjectives’, *Linguistics and Philosophy* 4: 1–45.

Kneale, W. and Kneale, M. 1962. *The Development of Logic*, Oxford, Clarendon Press.

Kohl, M. 1969. ‘Bertrand Russell on vagueness’, *Australasian Journal of Philosophy* 47: 31–41.

Körner, S. 1955. *Conceptual Thinking*, Cambridge, Cambridge University Press.

Körner, S. 1959. ‘On determinables and resemblance’, *Aristotelian Society*, suppl. 33: 125–40.

Körner, S. 1960. *The Philosophy of Mathematics*, London, Hutchinson.

Körner, S. 1966. *Experience and Theory*, London, Routledge & Kegan Paul.

Körner, S. 1968. ‘Reply to Mr Kumar’, *British Journal for the Philosophy of Science* 18: 323–4.

Kripke, S.A. 1972. ‘Naming and necessity’, in D. Davidson and G. Harman (eds) *Semantics of Natural Language*, Dordrecht, Reidel; reprinted in S.A. Kripke, *Naming and Necessity*, Oxford, Blackwell, 1980.

Kripke, S.A. 1975. ‘Outline of a theory of truth’, *Journal of Philosophy* 72: 690–716.

Kumar, D. 1968. ‘Logic and inexact predicates’, *British Journal for the Philosophy of Science* 18: 211–22.

Lakoff, G. 1973. ‘Hedges: a study in meaning criteria and the logic of fuzzy concepts’, *Journal of Philosophical Logic* 2: 458–508.

Lenzen, W. 1978. ‘Recent work in epistemic logic’, *Acta Philosophica Fennica* 30: 1–219.

Lewis, D.K. 1970. ‘General semantics’, *Synthese* 22: 18–67; reprinted in his *Philosophical Papers*, vol. 1, Oxford, Oxford University Press, 1983.

Lewis, D.K. 1979. ‘Scorekeeping in a language game’, *Journal of Philosophical Logic*, 8: 339–59; reprinted in his *Philosophical Papers*, vol. 1, Oxford, Oxford University Press, 1983.

Lewis, D.K. 1983. *Philosophical Papers*, vol. 1, Oxford, Oxford University Press.

Lewis, D.K. 1988. 'Vague identity: Evans misunderstood', *Analysis* 48: 128–30.

Long, A.A. and Sedley, D.N. 1987. *The Hellenistic Philosophers*, Cambridge, Cambridge University Press.

Lukasiewicz, J. and Tarski, A. 1930. 'Untersuchungen über den Aussagenkalkül', *Comptes rendus des séances de la Société des Sciences et des Lettres de Varsovie*, cl. 3, 23: 1–21, 30–50; reprinted in A. Tarski, *Logic Semantics, Metamathematics*, tr. J.H. Woodger, ed. J. Corcoran, Indianapolis, Ind., Hackett, 2nd edn, 1983.

Machina, K.F. 1972. 'Vague predicates', *American Philosophical Quarterly* 9: 225–33.

Machina, K.F. 1976. 'Truth, belief and vagueness', *Journal of Philosophical Logic* 5: 47–78.

Marmor, A. 1992. *Interpretation and Legal Theory*, Oxford, Clarendon Press.

Masson-Oursel, P. 1912. 'Esquisse d'une theorie comparée du sorite', *Revue de Metaphysique et de Morale* 20: 810–24.

McDowell, J.H. 1977. 'On the sense and reference of a proper name', *Mind* 86: 159–85.

McDowell, J.H. 1982. 'Truth-value gaps', in L.J. Cohen (ed.) *Proceedings of the Sixth International Congress for Logic, Methodology and Philosophy of Science*, Amsterdam, North-Holland.

McGee, V. 1991. *Truth, Vagueness and Paradox*, Indianapolis, Ind., Hackett.

Mehlberg, H. 1958. *The Reach of Science*, Toronto, University of Toronto Press.

Mellor, D.H. 1991. *Matters of Metaphysics*, Cambridge, Cambridge University Press.

Mignucci, M. 1993. 'The Stoic analysis of the sorites', *Proceedings of the Aristotelian Society* 93: 231–45.

Moline, J. 1969. 'Aristotle, Eubulides, and the sorites', *Mind* 78: 393–407.

Monfasani, J. 1990. 'Lorenzo Valla and Rudolph *Agricola*', *Journal of the History of Philosophy* 28: 181–200.

Morgan, C.G. and Pelletier, J.F. 1977. 'Some notes concerning fuzzy logics', *Linguistics and Philosophy* 1: 79–97.

Nadin, M. 1980. 'The logic of vagueness and the category of synechism', *The Monist* 63: 351–63.

Newman, M.H.A. 1928. 'Mr Russell's "Causal Theory of Perception"', *Mind* 37: 137–48.

Noonan, H.W. 1982. 'Vague objects', *Analysis* 42: 3–6.

Noonan, H.W. 1984. 'Indefinite identity: a reply to Broome', *Analysis* 44: 117–21.

Noonan, H.W. 1990. 'Vague identity yet again', *Analysis* 50: 157–62.

Novák, V. 1989. *Fuzzy Sets and their Applications*, Bristol, Adam Hilger.

Nozick, R. 1981. *Philosophical Explanations*, Oxford, Clarendon Press.

Odegard, D. 1965. 'Excluding the middle from loose concepts', *Theoria* 31: 138–44.

Osheron, D.N. and Smith, E.E. 1981. ‘On the adequacy of prototype theory as a theory of concepts’, *Cognition* 9: 35–58.
Osheron, D.N. and Smith, E.E. 1982. ‘Gradedness and conceptual combination’, *Cognition* 12: 299–318.
Over, D.E. 1984. ‘The consequences of direct reference’, *Philosophical Books* 25: 1–7.
Over, D.E. 1989. ‘Vague objects and identity’, *Analysis* 49: 97–9.
Pap, A. 1949. *Elements of Analytic Philosophy*, New York, Macmillan.
Pap, A. 1958. *Semantics and Necessary Truth*, New Haven, Conn., Yale University Press.
Parfit, D.A. 1971. ‘Personal identity’, *Philosophical Review* 80: 3–27.
Parfit, D.A. 1984. *Reasons and Persons*, Oxford, Oxford University Press.
Parikh, R. 1983. ‘The problem of vague predicates’, in R.S. Cohen and M.W. Wartofsky (eds) *Language*, *Logic and Method*, Dordrecht, Reidel.
Parsons, T. 1987. ‘Entities without identity’, *Philosophical Perspectives* 1: 1–19.
Pavelka, J. 1979. ‘On fuzzy logic, I, II, III’, *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik* 25: 45–52, 119–34, 447–64.
Peacocke, C.A.B. 1981. ‘Are vague predicates incoherent?’, *Synthese* 46: 121–41.
Peirce, C.S. 1905. ‘Issues of pragmaticism’, *The Monist* 15: 481–99; reprinted in C.S. Peirce, *Collected Writings of Charles Sanders Peirce*, ed. C. Hartshorne, P. Weiss *et al.,* vol. 5, Cambridge, Mass., Harvard University Press, 1931–56.
Peirce, C.S. 1931–56. *Collected Writings of Charles Sanders Peirce*, ed. C. Hartshorne, P. Weiss *et al*., 8 vols, Cambridge, Mass., Harvard University Press.
Pelletier, F.J. 1989. ‘Another argument against vague objects’, *Journal of Philosophy* 86: 481–92.
Pinkal, M. 1983a. ‘Towards a semantics of precization’, in T.T. Ballmer and M. Pinkal (eds) *Approaching Vagueness*, Amsterdam, North-Holland.
Pinkal, M. 1983b. ‘On the limits of lexical meaning’, in R. Bäuerle, C. Schwarze and A. von Stechow (eds) *Meaning*, *Use and Interpretation*, Berlin, de Gruyter.
Platts, M.B. 1979. *Ways of Meaning*, London, Routledge & Kegan Paul.
Prantl, C. 1855–70. *Geschichte der Logik im Abendlande*, 4 vols, Leipzig, Fock.
Price, H.H. 1932. *Perception*, London, Methuen.
Przelecki, M. 1969. *The Logic of Empirical Theories*, London, Routledge & Kegan Paul.
Przelecki, M. 1976. ‘Fuzziness as multiplicity’, *Erkenntnis* 10: 371–80.
Przelecki, M. 1979. ‘The semantics of open concepts’, in J. Pelc (ed.) *Semiotics in Poland*, Warsaw, Polish Scientific Publications.
Putnam, H. 1983. ‘Vagueness and alternative logic’, *Erkenntnis* 19: 297–314.
Putnam, H. 1985. ‘A quick Read is a wrong Wright’, *Analysis* 45: 203.
Putnam, H. 1991. ‘Replies and comments’, *Erkenntnis* 34: 401–24.
Quine, W.V. 1960. *Word and Object*, Cambridge, Mass., MIT Press.
Quine, W.V. 1981. ‘What price bivalence?’, *Journal of Philosophy* 77: 90–5.
Rasmussen, S.A. 1986. ‘Vague identity’, *Mind* 95: 81–91.
Rea, G. 1989. ‘Degrees of truth versus intuitionism’, *Analysis* 49: 31–2.

Read, S.L. and Wright, C.J.G. 1985. 'Hairier than Putnam thought', *Analysis* 45: q 56–8.
Rescher, N. 1969. *Many-Valued Logic*, New York, McGraw-Hill.
Rolf, B. 1980. 'A theory of vagueness', *Journal of Philosophical Logic* 9: 315–25.
Rolf, B. 1982. 'Russell's theses on vagueness', *History and Philosophy of Logic* 3: 69–83.
Rolf, B. 1984. 'Sorites', *Synthese* 58: 219–50.
Rosch, E.H. 1973. 'On the internal structure of perceptual and semantic categories', in T.E. Moore (ed.) *Cognitive Development and the Acquisition of Language*, New York, Academic Press.
Rosch, E.H. and Mervis, C.B. 1975. 'Family resemblances: studies in the internal structure of categories', *Cognitive Psychology* 7: 573–605.
Rose, A. and Rosser, J.B. 1958. 'Fragments of many-valued statement calculi', *Transactions of the American Mathematical Society* 87: 1–53.
Rosser, J.B. 1960. 'Axiomatization of infinite-valued logics', *Logique et Analyse* 3: 137–53.
Rosser, J.B. and Turquette, A.R. 1952. *Many-Valued Logics*, Amsterdam, North-Holland.
Russell, B.A.W. 1921. *The Analysis of Mind*, London, Allen & Unwin.
Russell, B.A.W. 1923. 'Vagueness', *Australasian Journal of Philosophy and Psychology* 1: 84–92; reprinted in B.A.W. Russell, *The Collected Papers of Bertrand Russell*, ed. E. Eames, J. Slater *et al.*, vol. 9, London, Allen & Unwin/ Unwin Hyman, 1983–.
Russell, B.A.W. 1927. *The Analysis of Matter*, London: Kegan Paul.
Russell, B.A.W. 1948. *Human Knowledge: Its Scope and Limits*, London, Allen & Unwin.
Russell, B.A.W. 1950. *An Inquiry into Meaning and Truth*, London, Allen & Unwin.
Russell, B.A.W. 1983–. *The Collected Papers of Bertrand Russell*, ed. E. Eames, J. Slater *et al.*, many vols, London, Allen & Unwin/Unwin Hyman.
Russell, B.A.W. and Whitehead, A.N. 1910–13. *Principia Mathematica*, 3 vols, Cambridge, Cambridge University Press.
Saarinen, E. (ed.) 1979. *Game-Theoretical Semantics*, Dordrecht, Reidel.
Sainsbury, R.M. 1979. *Russell*, London, Routledge & Kegan Paul.
Sainsbury, R.M. 1986. 'Degrees of belief and degrees of assent', *Philosophical Papers* 15: 97–106.
Sainsbury, R.M. 1988–9. 'Tolerating vagueness', *Proceedings of the Aristotelian Society* 89: 33–48.
Sainsbury, R.M. 1989. 'What is a vague object?', *Analysis* 49: 99–103.
Sainsbury, R.M. 1991. 'Is there higher-order vagueness?', *Philosophical Quarterly* 41: 167–82.
Sainsbury, R.M. 1992. 'Sorites paradoxes and the transition question', *Philosophical Papers* 21: 177–90.
Salmon, N.U. 1982. *Reference and Essence*, Oxford, Basil Blackwell.

Salmon, N.U. 1984. 'Fregean theory and the Four Worlds Paradox: a reply to David Over', *Philosophical Books* 25: 7–11.
Salmon, N.U. 1986a. 'Modal paradox: parts and counterparts, points and counterpoints', *Midwest Studies in Philosophy* 11: 75–120.
Salmon, N.U. 1986b. *Frege's Puzzle*. Cambridege, Mass., MIT Press.
Salmon, N.U. 1989. 'The logic of what might have been', *Philosophical Review* 98: 3–34.
Salmon, N.U. 1990. 'A Millian heir rejects the wages of *Sinn*', in C.A. Anderson and J. Owens (eds) *Propositional Attitudes: The Role of Context in Logic, Language and Mind*, Stanford, Calif., Center for the Study of Language and Information.
Sanford, D.H. 1975. 'Borderline logic', *American Philosophical Quarterly* 12: 29–40.
Sanford, D.H. 1976. 'Competing semantics of vagueness: many values versus super-truth', *Synthese* 33: 195–210.
Sanford, D.H. 1979. 'Nostalgia for the ordinary: comments on papers by Unger and Wheeler', *Synthese* 41: 175–84.
Sayward, C. 1989. 'Does the law of excluded middle require bivalence?', *Erkenntnis* 31: 129–37.
Scarpellini, B. 1962. 'Die Nicht-Axiomatisierbarkeit des unendlichwertigen Prädikatenkalküls von Lukasiewicz', *Journal of Symbolic Logic 27*: 159–70.
Scheffler, I. 1979. *Beyond the Letter*, London, Routledge & Kegan Paul.
Schiller, F.C.S. 1934. 'How is "exactness" possible?', *Actes du Huitième Congrès International de Philosophie*; reprinted in F.C.S. Schiller, *Our Human Nature*, New York, Columbia University Press, 1939.
Schmitt, C.B. 1983. 'The rediscovery of ancient skepticism in modern times', in M.F. Burnyeat (ed.) *The Skeptical Tradition*, Berkeley, Calif., University of California Press.
Schwartz, S.P. 1987. 'Intuitionism and sorites', *Analysis* 47: 179–83.
Schwartz, S.P. 1989. 'Vagueness and incoherence: a reply to Burns', *Synthese* 80: 395–406.
Schwartz, S.P. 1990. 'Intuitionism versus degrees of truth', *Analysis* 50: 43–7.
Schwartz, S.P. and Throop, W. 1991. 'Intuitionism and vagueness', *Erkenntnis* 34: 347–56.
Sedley, D.N. 1977. 'Diodorus Cronus and Hellenistic philosophy', *Proceedings of the Cambridge Philological Society* 23: 74–120.
Sedley, D.N. 1982. 'On signs', in J. Barnes, J. Brunschwig, M.F. Burnyeat and M. Schofield (eds) *Science and Speculation*, Cambridge, Cambridge University Press.
Sedley, D.N. 1984. 'The negated conjunction in Stoicism', *Elenchos* 5: 311–16.
Segerberg, K. 1965. 'A contribution to nonsense-logics', *Theoria* 31: 199–217.
Shapiro, S. 1991. *Foundations without Foundationalism: A Case for Second-order Logic*, Oxford, Clarendon Press.

Sillitti, G. 1984. 'Aristotele e l'aporia del sorite. A proposito di due studi recenti', *Elenchos* 5: 211–20.
Simons, P. 1992. 'Vagueness and ignorance', *Aristotelian Society*, suppl. 66: 163–77.
Smith, J.W. 1984. 'The surprise examination on the paradox of the heap', *Philosophical Papers* 13: 43–56.
Soames, S. 1989. 'Semantics and semantic competence', *Philosophical Perspectives* 3: 575–96.
Sorensen, R.A. 1985. 'An argument for the vagueness of "vague"', *Analysis* 45: 134–7.
Sorensen, R.A. 1987. 'The vagueness of knowledge', *Canadian Journal of Philosophy* 17: 767–804.
Sorensen, R.A. 1988a. 'Precisification by means of vague predicates', *Notre Dame Journal of Formal Logic* 29: 267–74.
Sorensen, R.A. 1988b. *Blindspots*, Oxford, Clarendon Press.
Sorensen, R.A. 1988c. 'Vagueness, measurement and blurriness', *Synthese* 75: 45–82.
Sorensen, R.A. 1989. 'The ambiguity of vagueness and precision', *Pacific Philosophical Quarterly* 70: 174–83.
Sorensen, R.A. 1991. 'Vagueness within the language of thought', *Philosophical Quarterly* 41: 389–413.
Sperber, D. and Wilson, D. 1985–6. 'Loose talk', *Proceedings of the Aristotelian Society* 86: 153–71.
Stalnaker, R. 1988. 'Vague identity', in D.F. Austin (ed.) *Philosophical Analysis: A Defence by Examples*, Dordrecht, Kluwer.
Stechow, A. von 1984. 'Comparing semantic theories of comparison', *Journal of Semantics*, 3: 1–77.
Strawson, P.F. 1950. 'On referring', *Mind* 59: 320–44.
Tarski, A. 1983. *Logic, Semantics, Metamathematics*, 2nd edn, tr. J.H. Woodger, ed. J. Corcoran, Indianapolis, Ind., Hackett.
Thomason, R. 1982. 'Identity and vagueness', *Philosophical Studies* 43: 329–32.
Thompson, M. 1953. *The Pragmatic Philosophy of C.S. Peirce*, Chicago, Ill., Chicago University Press.
Travis, C. 1985. 'Vagueness, observation, and sorites', *Mind* 94: 345–66.
Tye, M. 1990. 'Vague objects', *Mind* 99: 535–57.
Tye, M. 1991. *The Imagery Debate*, Cambridge, Mass., MIT Press.
Tye, M., forthcoming. 'Sorites paradoxes and the semantics of vagueness', *Philosophical Perspectives.*
Unger, P.K. 1979a. 'There are no ordinary things', *Synthese* 41: 117–54.
Unger, P.K. 1979b 'I do not exist', in G.F. Macdonald (ed.) *Perception and Identity*, London, Macmillan.
Unger, P.K. 1979c. 'Why there are no people', *Midwest Studies in Philosophy* 4: 177–222.

Unger, P.K. 1980. ‘The problem of the many’, *Midwest Studies in Philosophy* 5: 411–67.

Unger, P.K. 1990. *Identity*, *Consciousness and Value*, Oxford, Oxford University Press.

Urquhart, A. 1986. ‘Many-valued logic’, in D. Gabbay and F. Guenthner *Handbook of Philosophical Logic*, vol. III: *Alternatives to Classical Logic*, Dordrecht, Reidel.

Valla, L. 1982. *Repastinatio Dialectice et Philosophie*, ed. G. Zippel, 2 vols, Padua, Antenore.

Vopnka, P. 1979. *Mathematics in the Alternative Set Theory*, Leipzig, Teubner.

Waismann, F. 1945. ‘Verifiability’, *Aristotelian Society*, suppl. 19: 119–50.

Waismann, F. 1945–6. ‘Are there alternative logics?’, *Proceedings of the Aristotelian Society* 46: 77–104.

Waismann, F. 1951. ‘Analytic–synthetic III’, *Analysis* 11: 49–61.

Weiss, S. 1976. ‘The sorites fallacy: what difference does a peanut make?’, *Synthese* 33: 253–72.

Wheeler, S.C. 1975. ‘Reference and vagueness’, *Synthese* 30: 367–80.

Wheeler, S.C. 1979. ‘On that which is not’, *Synthese* 41: 155–94.

Wiggins, D.R.P. 1979. ‘On knowing, knowing that one knows and consciousness’, in E. Saarinen, R. Hilpinen, I. Niiniluoto and M. Provence Hintikka (eds) *Essays in Honour of Jaakko Hintikka*, Dordrecht, Reidel.

Wiggins, D.R.P. 1980. *Sameness and Substance*, Oxford, Blackwell.

Wiggins, D.R.P. 1986. ‘On singling out an object determinately’, in P. Pettit and J.H. McDowell (eds) *Subject*, *Thought*, *and Context*, Oxford, Clarendon Press.

Williamson, T. 1990a. *Identity and Discrimination*, Oxford, Blackwell.

Williamson, T. 1990b. ‘Verification, falsification and cancellation in KT’, *Notre Dame Journal of Formal Logic* 31: 286–90.

Williamson, T. 1992a. ‘Inexact knowledge’, *Mind* 101: 217–42.

Williamson, T. 1992b. ‘An alternative rule of disjunction in modal logic’, *Notre Dame Journal of Formal Logic* 33: 89–100.

Williamson, T. 1992c. ‘Vagueness and ignorance’, *Aristotelian Society*, suppl. 66: 145–62.

Wittgenstein, L. 1953. *Philosophical Investigations*, tr. G.E.M. Anscombe, Oxford, Blackwell.

Wittgenstein, L. 1961a. T*ractatus Logico-Philosophicus*, tr. D. Pears and B. McGuinness, London, Routledge & Kegan Paul.

Wittgenstein, L. 1961b. *Notebooks 1914–1916*, ed. G.H. von Wright and G.E.M. Anscombe, Oxford, Blackwell.

Wittgenstein, L. 1975. *Philosophical Remarks*, ed. R. Rhees, tr. R. Hargreaves and R. White, Oxford, Blackwell.

Wright, C.J.G. 1975. ‘On the coherence of vague predicates’, *Synthese* 30: 325–65.

Wright, C.J.G. 1976. ‘Language mastery and the sorites paradox’, in M.G.J. Evans and J.H. McDowell (eds) *Truth and Meaning*, Oxford, Oxford University Press.

Wright, C.J.G. 1987. 'Further reflections on the sorites paradox', *Philosophical Topics* 15: 227–90.

Wright, C.J.G. 1991. 'The sorites paradox and its significance for the interpretation of semantic theory', in N.L. Cooper and P. Engel (eds) *New Inquiries into Meaning and Truth*, London, Harvester Wheatsheaf.

Wright, C.J.G. 1992. 'Is higher order vagueness coherent?', *Analysis* 52: 129–39.

Zadeh, L.A. 1965. 'Fuzzy sets', *Information and Control* 8: 338–53.

Zadeh, L.A. 1972. 'A fuzzy set-theoretic interpretation of linguistic hedges', *Journal of Cybernetics* 2: 4–34.

Zadeh, L.A. 1975. 'Fuzzy logic and approximate reasoning', *Synthese* 30: 407–28.

Zadeh, L.A. 1982. 'A note on prototype theory and fuzzy sets', *Cognition* 12: 291–7.

Zemach, E.M. 1991. 'Vague objects', *Noûs* 25: 323–40.

Ziff, P. 1974. 'The number of English sentences', *Foundations of Language* 11: 519–32.

Ziff, P. 1984. *Epistemic Analysis*, Dordrecht, Reidel.

Index