Potentiality and Possibility

Barbara Vetter

Corpus Christi College, University of Oxford

Submitted for the degree of Doctor in Philosophy

Trinity Term 2010

Word Count (excluding bibliography): 74 862

Acknowledgments

In writing this thesis, I have had two wonderful supervisors: Timothy Williamson and Antony Eagle. I am greatly indebted to both for their continued support and insightful criticism. Timothy Williamson in particular has supervised me throughout my time as a graduate student, and I have learned more about philosophy from him than from anyone else.

I have greatly benefitted from conversations with, and comments from, various philosophers, including Simona Aimar, Lauren Ashwell, Alexander Bird, Eline Busck-Gundersen, Bill Child, Troy Cross, Natalja Deng, Kit Fine, John Hawthorne, Frank Hofmann, Gail Leckie, Angela Matthies, Katherine Munn, Markus Schrenk, and Alastair Wilson, as well as audiences at Oxford, Geneva, Regensburg, New York, Stockholm, Norwich, Nottingham, Bremen, Berlin, Oslo, and Köln. Markus Schrenk shared and and encouraged my project from the start and discussed it with me over many a coffee. Natalja Deng and Katherine Munn embarked with me on the journey through modal logic. Antony Eagle and Alastair Wilson regularly met with me to discuss dispositions. Benjamin Vauteck, Katherine Munn, and Yoaav Isaacs read a near-to-final draft of this thesis and eliminated several stylistic and orthographic errors. I would like to express my gratitude to all of these philosophers, and to all those that I forgot to mention here.

While working on this thesis, I was financially supported by the AHRC, the Studienstiftung des deutschen Volkes, and the DAAD. I also received a Senior Scholarship from Corpus Christi College, and a Hanfling Scholarship from the Faculty of Philosophy. I would especially like to thank the Hanfling family for their kindness and generosity.

I thank Angela Matthies, Katherine Munn, Natalja Deng, and Gail Leckie for being wonderful housemates and friends (as well as great philosophers). I thank Friedemann Buddensiek for telling me to go to Oxford many years ago. I thank my parents for their support in everything I do, and for bearing with equanimity the fact that what I do is philosophy. Above all, I thank Benjamin Vauteck for being there. Without his love, support, and sense of humour, my life in the past years (and in all likelihood this thesis) would have been very different, and not nearly as good.

Oxford, 15 September 2010

Contents

Chapter 1

Introduction

Individual objects have potentials: paper has the potential to burn, an acorn has the potential to turn into a tree, some people have the potential to run a mile in less than four minutes. Some potentials have names: a vase's fragility is a potential to break, a person's irascibility is her potential to get angry. Some potentials are classified as abilities, such as your ability to read English; others are not, such as the paper's potential to burn. Some potentials are classified, by philosophers, as 'dispositions'; these include the vase's fragility, a person's irascibility, and an atom's disposition to decay. Potentials are often ascribed simply with the auxiliary 'can': paper can burn, the acorn can become a tree, some people can run a mile in less than four minutes. This, in fact, is one simple way of stating what all the properties I have so far mentioned have in common: they concern what a given individual *can* do. I call any such property a potentiality, and I believe that our intuitive grasp of the properties I mentioned already provides some good pre-theoretical understanding of that term. I will provide an alternative way of introducing the term in a moment.

It goes without saying that the notion of potentiality is of Aristotelian pedigree, and I believe Aristotle's *Metaphysics*, book Θ, to be one of the most illuminating treatments of it. With the rise of a certain reductive brand of empiricism and especially under the influence of David Hume, talk of potentiality was regarded as suspect and became less and less commonly appealed to in those strands of philosophy that led to contemporary analytic philosophy – a fate it shared with many other modal notions, such as essence and a metaphysically substantial notion of necessity and possibility.[1] The latter pair had a revival in analytic philosophy with the development of modal logic and the discovery that a semantics of 'possibly' and 'necessarily' can be treated as a special case of the logic of the existential and universal quantifier, as long as we allow the quantifiers to range over an infinity of 'possible worlds'. The philosophical debate about the nature of modality has subsequently focussed largely on the nature of those 'possible worlds'. In the process, the modal notions of an object's potentiality or its essential properties have been either neglected or explained in terms of those of possibility or necessity. To many, they have remained suspect.

In recent years, the philosophical climate has changed somewhat. Kit Fine has provided a theory of essence as irreducible to, and in fact prior to, necessity (see Fine 1994). Anti-Humean accounts of the laws of nature, admitting 'necessary connections' in nature, and some even locating these connections in dispositional properties of fundamental entities, are thriving (see, e.g., Ellis 2001, Bird 2007; Armstrong 1983 has been instrumental in the rise of anti-Humeanism). Dispositions in general have received a fair amount of attention in the recent metaphysical debate, though they tend to be treated in isolation from other kinds of potentiality, such as abilities. The ideas put forward in the pages to follow are part of that changed climate.

[1] I use the term 'modal' in the linguists' sense: everything that is expressed by the modal auxiliaries is a modality. Thus Fine (1994)'s 'non-modal' account of essence is, in my terms, not non-modal: it is merely an account that rejects the reduction of one modality (essence) to another (necessity).

This thesis is a plea for potentiality. It is a plea for recognizing a unified notion of potentiality instead of selectively focussing attention on only some kinds of potentiality; and it is a plea for recognizing potentiality as (at least) equal in ontological rank to possibility.

How exactly does potentiality relate to possibility? We may call the distinction between them a distinction between *global* and *local* modality. Possibility is global: its being possible that such-and-such is not primarily a fact about any one particular object; it is a fact about how the world could have turned out to be. Hence our intuitions about what is possible and what is not can be captured by postulating, for everything that is possible, a world that did turn out to be that way. Potentiality, on the other hand, is local: a potentiality is a property of a particular object. That I have the potential to write this thesis is first and foremost a fact about me; it is a property that I possess. The distinction is paralleled by the relation of essence to necessity. Essence, like potentiality, is local: a property is essential *to* a particular object. Necessity, like possibility, is global: its being necessary that such-and-such is not primarily a fact about one particular object, but a fact about how the world must be. The difference between essence and necessity has been pointed out and studied by Kit Fine (1994).

This provides me with another way of introducing the notion of a potentiality: potentialities are possibilities rooted in objects; they are like possibilities, but they are properties of individual objects. They stand to possibility just as essence (on the Finean view) stands to necessity.

Introducing the notion of potentiality in this way is not meant to provide a definition or reduction, not only because it is metaphorical, but also because (as I will later argue) potentiality should not be defined in terms of, or reduced to, possibility. The introduction is meant only as a heuristic device for philosophers who are more familiar with the global modalities of possibility and necessity. The notion of potentiality itself will remain undefined and unreduced throughout the thesis. My plea for potentiality is to show precisely that taking potentiality as a primitive or basic notion is philosophically fruitful; that we can say a great deal about potentiality without defining or reducing it, and that we can say a great deal about other things in terms of potentiality.

In particular, we can say a great deal about possibility in terms of potentiality. This is the one main use to which I want to put the notion of potentiality: to develop a realist account of possibility (and, thereby, of necessity) that is based entirely on potentiality. Potentiality is, metaphorically speaking, possibility anchored in individual objects; I claim that all possibility is thus anchored in some individual object(s) or other.[2]

The potentiality-based account does entirely without possible worlds, concrete or abstract. I do not argue here that such an account is desirable, but there are various lines of argument to that effect in the literature. Roughly, the reasoning goes as follows: if we give an account of modality in terms of possible worlds, those worlds are either concrete, Lewisian worlds (*locus classicus*: Lewis 1986a) or some kind of abstract entity, such as sets of propositions (*loci classici*: Adams 1974, Stalnaker 1976, Plantinga 1974). As to the former, the 'incredulous stare' (Lewis 1986a) is a strong objection; furthermore, it is hard to see what evidence could be adduced for that initially rather implausible claim; and finally, even if it were true that there are infinitely many concrete universes, that does not seem to be a fact about possibility and necessity, but rather a curious contingent fact about the one actual world, which includes all those 'universes'. If, on the other hand, possible worlds are sets of propositions, we need some way to distinguish those sets of propositions that do from those that do not

[2]Similar ideas have been developed by Mondadori and Morton (1976) and Jacobs (2010), but neither of them has focused on the direct link between potentiality and possibility. Borghini and Williams (2008) have sketched a 'dispositional theory of possibility' that is similar to mine, but based on a different conception of dispositions and spelled out in much less detail.

correspond to genuine possibilities; mere logical consistency is not enough. If abstract possible worlds are supposed to deliver a robust account of metaphysical modality, it is hard to see how they can avoid circularity; if not, then they are simply irrelevant to the metaphysical question of what possibility and necessity are, though they may be useful heuristic and formal devices in studying certain properties of possibility and necessity. (Cf. Williamson 1998 and Jubien 2007 for contemporary versions of this kind of criticism.)

I take these to be good prima facie grounds to provide an alternative account of modality. The potentiality-based account has a further advantage: it provides a metaphysics of modality that does not make the epistemology of modality entirely mysterious. Unlike contemporary accounts of possibility and necessity that appeal to possible worlds, a potentiality-based account is firmly rooted in everyday life and thought. Potentialities are ubiquitous in our ordinary thought about, and dealings with, the world: we handle glasses carefully if they are fragile, we handle people carefully if they are irascible or vulnerable, we learn languages in order to acquire the ability to speak them, and practise the piano to improve our skills at it. A central advantage of my account is that it treats modality not as a philosophers' creature but as something on which we plausibly have a good pre-theoretical grasp.[3]

My aim in this thesis will not be to show that a potentiality-based account of possibility must be adopted, but rather that it can be adopted. I aim to show that given our pre-philosophical grasp on some potentialities, in particular abilities and dispositions, we have the materials to construct a non-reductive account of potentiality; and that given this account, we have the materials to construct a theory of possibility from

[3]For more on the desirability of such an account, see Williamson (2007), whose take on the epistemology of modality is of course a little different.

it. My argument will therefore be mostly constructive rather than destructive. Here is a brief overview of how it will proceed.

In chapter 2, I argue against the standard conception of dispositions as characterized by counterfactual conditionals. On my conception, dispositions are a kind of potentiality, and potentiality is possibility-like. It is possibility-like in that a potentiality is characterized by its exercise, as a potentiality to ..., just as a possibility is a possibility that ... . On the standard model, dispositions are dispositions to ... if ..., akin to counterfactual conditionals rather than to possibility. My argument against the standard model is at the same time, though not explicitly, an argument for a unified conception of potentiality, and of dispositions as a kind of potentiality.

Chapter 3 does not yet provide that conception, but provides further and explicit motivation for a general account of potentiality. I have said that potentiality stands to possibility as essence stands to necessity. Kit Fine has argued that we should take essence, not necessity, to be metaphysically more basic; I sketch an argument to the same effect for potentiality and possibility, but note that its conclusion depends on a suitably general account of potentiality. I have further employed of the modal auxiliary 'can' in introducing the notion of potentiality; in the second part of chapter 3, I propose that 'can' is plausibly construed as ascribing potentiality, contrast my proposal with the standard possible-worlds semantics of 'can', and conclude, again, that a general account of potentiality is needed.

Chapter 4 then goes on to give what I have promised at the beginning: a unified picture of potentiality, covering both dispositions and abilities (and more than that). The guiding question of the chapter is how far we can extend the notion of potentiality from the initial examples of dispositions and abilities. I argue that we can extend it very far indeed. We find that there are both intrinsic and extrinsic potentialities, that the logical behaviour of potentiality is much like that of possibility (e.g., having a potentiality to be F-or-G entails having a potentiality to be F or a potentiality to be G; being F entails having a potentiality to be F; etc.), and introduce the notion of an *iterated* potentiality: a potentiality to have a potentiality to ... . Iterated potentialities greatly increase the 'reach' of any individual object's potentialities; objects have iterated potentialities for many more things than they have non-iterated potentialities for.

With all this in place, I set the notion of potentiality to work. Chapter 5 provides the outlines of a potentiality-based account of possibility. The basic claim of the account is simple: every possibility is 'rooted' in some potentiality, and thereby in some object. Thus it is possible that p just in case there is (at some time) some object (or objects) with an (iterated) potentiality for it to be the case that p. To be a serious contender, the account needs to prove two things: its formal adequacy, and its extensional correctness. Chapter 5 focusses on extensional correctness and, in particular, tries to meet the challenge that there might not be enough potentialities around to ground all the possibilities that we think there are. Rather than putting forward a particular version of the account with a particular solution to the challenge, I provide a survey of the options available to a potentiality-based account of possibility in response to the various versions of the challenge. Such an account has more options available to it than might first be thought. Chapter 5 should establish the potentiality-based view of possibility as a promising research programme.

That research programme must meet the further challenge of formal adequacy: it must make sure that possibility, understood in terms of potentiality, turns out to have the 'right' formal features, i.e. the features studied in modal logic (and validated by possible-worlds based views of modality). In the sixth and final chapter, I meet that challenge by formulating a fully formal treatment of potentiality and deriving from it the logic of possibility, which does in fact turn out to have the right kind of formal features.

No metaphysical question can be treated in isolation from all others. In the course of discussing potentiality and possibility, I will touch on many different issues in metaphysics. In most cases, I try to remain neutral and offer different versions of my claim to go with different views on these issues. This strategy cannot, however, be extended to all cases; on two important questions I have decided to take a stand. I would be interested to see how my arguments could be rephrased so as to be compatible with opposing views, but doing so is beyond the scope of this thesis. The following two assumptions will be made, without further argument.

First: realism about properties. I assume that there are properties, and that a property is the kind of entity that different particulars can share. When two apples are both red, then there is a property that they both possess, the property of being red. Accordingly, different things can share a potential for the same property: the sheets of paper in this thesis all share the potential to burn. Nominalists about properties have different strategies in dealing with our ubiquitous talk of properties shared by things. One is to claim that when we say that two apples share a property, what we say is not in fact true, but something similar is (the two red apples have more or less exactly resembling tropes, or they are both members of the same class of particulars that resemble each other in a certain respect). Another is to claim that what we say is true, but analyse away the apparent appeal to shared properties (what we *really* say is that things have resembling tropes, or that they belong to the same class). I suspect that both strategies could be applied throughout this thesis, making only slight alterations to my central claims. But at the moment, I have no more than that suspicion to offer to the nominalist.

Second: properties and naturalness. I adopt a liberal approach to properties. The properties that things possess include: electric charge, triangularity, being self-identical, being green, being grue, and being such that it is raining. To be sure, some of these properties are less natural than others. The list I just gave was arranged in order of decreasing naturalness (though I am not sure about the place in that order of a logical property such as being self-identical). Naturalness, that is, comes in degrees; some properties are more natural and others less so, but the distinction is one among properties and not one between what properties there are and aren't. It may be thought that there is a privileged set of 'perfectly natural' or 'fundamental' properties, which are natural to the highest degree possible. Maybe so; but I am just as interested in the rest of the spectrum.

Chapter 2

Dispositions

2.1 Introduction

Dispositions are modal properties. In saying of a vase that it is fragile, of a person that she is irascible, or of a disease that it is transmittable, we are not, or not primarily, saying something about what the vase, the person or the disease is *actually* doing, but rather about what they *would* or *could* do. Yet in making these statements, we are not, or not primarily, saying something about a possible course of events, but rather something about what *this object* could or would contribute to such a course of events. Thus a fragile vase remains fragile when packed in styrofoam, a disease remains transmittable even when vaccination is available, and a person remains irascible when removed from all provoking factors. The modality of a disposition is *local* to a particular object.

How does this local modality relate to the 'global' modalities to which philosophers have devoted so much more attention – possibility, necessity, the counterfactual conditional? Is there a 'special relationship' between dispositions on the one hand and some one of those global modalities on the other? If so, which modality is it, and what is the special relationship?

The contemporary literature has a quick answer to parts of this question: there is a special relationship, and it is with the counterfactual conditional. This answer deserves to be called the Standard Model, and it comes in two steps:

The Standard Model.

1. Dispositionality is a two-place operator: a disposition is always a disposition to M if S, where S is the disposition's stimulus condition, and M its manifestation condition.
   (For instance, fragility is the disposition to break if struck.)
2. A disposition to M if S is importantly (though we don't quite know how) related to the counterfactual: if *x* were to S, then *x* would M.
   (For instance, fragility, being the disposition to break if struck, is related to the counterfactual: if *x* were struck, *x* would break.)

Debates have then focussed exclusively on the final part of my question: what is the relationship between dispositions and counterfactual conditionals – is it one of conceptual analysis, metaphysical reduction, or something else? Finks and reverse finks (Martin 1994), masks and mimics (Johnston 1992) and antidotes (Bird 1998) are now widely believed to rule out an analysis of dispositions in terms of counterfactual conditionals, as well as a straightforward reduction to them. Less straightforward reductions have been suggested in the form of a multi-track view of dispositions (one disposition is reduced to many counterfactuals), and recently some have suggested to replace the counterfactual conditional by something a little weaker but still conditional-like, such as a habitual (Fara 2005; a structurally similar proposal has been made by Manley and Wasserman 2008). Even the staunch anti-reductionists, however, subscribe to the Standard Model. For them, the *local* modality of a disposition is not reducible to the *global* modality of the counterfactual, but the link between dispositions and counterfactuals still holds: the irreducibility of dispositions to counterfactuals is taken as sufficient to prove their irreducibility *tout court* (Molnar 2003), and some even suggest that, on the contrary, it is counterfactuals that are reducible to, or 'derived from', dispositions (Bird 2007, Jacobs 2010).

In this chapter, I argue that this entire debate is based on a false assumption: the assumption that the Standard Model is, roughly, correct. I will omit any problems about finks, masks, antidotes and their like and ask the prior question: *which* modality is it that is 'specially related' to, or inherent in, dispositions? I will present three arguments against the straightforward version of the Standard Model, according to which there is one counterfactual (finks and masks aside) to characterize any one disposition (section 2.2), and then consider its more sophisticated versions, the multi-track view and the weaker counterfactuals (section 2.3). Finally, I will present my own view, which provides an entirely different answer to the question: the modality inherent in dispositions is not the counterfactual conditional, but a species of possibility (section 2.4). The Standard Model should be replaced by what we may call the Alternative Model:

The Alternative Model.

1. Dispositionality is a *one-place* operator: a disposition is always a disposition to M, where M is the disposition's manifestation condition.
   (For instance, fragility is the disposition to break.)
2. A disposition to M is importantly (though we don't quite know how) related to the possibility claim: x can M.
   (For instance, fragility, being the disposition to break, is related to the possibility claim: x can break.)

In section 2.5, I address a challenge to my argument that arises from the link between dispositions and laws.

2.2 Three Arguments against the Standard View

2.2.1 Counter-Examples (and a Puzzle)

The literature on dispositions tends to focus on the so-called 'philosophers' dispositions' (Prior 1985): dispositions which are possessed primarily by medium-sized, inanimate objects. The range of examples includes fragility, solubility, inflammability (and not much else). The model that is developed with a view to those examples, however, is usually intended to apply to a much wider range of dispositions: the psychological dispositions of animate beings such as ourselves, on the one hand, and on the other hand the fundamental or close to fundamental dispositions of such entities as atoms, electrons, and so forth. The philosophers' dispositions seem to be seen as a simpler case, one which allows for the formulation of a basic model. But it has been recognized that these two intended areas of application, the (close to) fundamental and the psychological, afford counterexamples to the standard model: dispositions that lack a stimulus condition.

It has become a common observation (going back at least to Molnar 2003, 85, who names the muon) that certain particles have a disposition to decay which is manifested spontaneously, without any triggering event that might be considered its stimulus. On the psychological side, libertarians may want to ascribe to human beings certain powers or dispositions the manifestation of which is not triggered by anything, but is an act of pure spontaneous, free will (again, see Molnar 2003). Less controversially, Manley and Wasserman (2008) have noted that 'talking in any situation at all can manifest loquaciousness, and anger in any situation at all can manifest irascibility.' (Manley and Wasserman 2008, 72)

We need to be careful, however, in seeing just how these cases are counterexamples to the standard model. For there are two rather different ways in which a disposition may be said to lack a stimulus condition.

First, a disposition may manifest spontaneously at any (or some) given occasion, with no stimulus or triggering condition whatsoever. A particle's disposition to decay or the libertarian powers are intended to belong to this category. It is, of course, a controversial question whether or not there are such 'active' dispositions, and I will not here try to convince you that there are. For the present purpose, it is quite sufficient to note that this controversial question *is* a question, and a substantial one – not one that should be settled almost by definition, in the very account we give of what dispositions *are*. Active dispositions are certainly a conceptual, and prima facie also a metaphysical, possibility. So the characterization of dispositions, whether it is intended conceptually or metaphysically, ought not to exclude them from the very start. But that is precisely what the standard model does.

Second, there may, at any occasion of a disposition's manifestation, be a cause of that manifestation – but these causes or triggering conditions may be so radically different that there is no non-trivial common description of them (no common determinable of which they are determinates). Hence there is no one property to serve as the disposition's stimulus condition, and no antecedent for the counterfactual conditional. This, I take it, is the kind of case that Manley and Wasserman envisage: a person's irascibility may, on every occasion of manifestation, be triggered by something or other, but there is nothing general to be said about the conditions that do trigger it, and a manifestation is a manifestation of irascibility quite regardless of what triggered it.

This second class of counterexamples is much less controversial than the first; but it too serves to refute the standard model. Just how large is the class of these less controversial counterexamples? My impression is that counterexamples of the second kind are now generally acknowledged to exist, but they are regarded as some sort of anomaly, a special case that will have to be accommodated eventually but should not be viewed as central. Centre stage is still taken by counterfactuals and those dispositions that appear to conform to the standard model, the philosophers' dispositions. I will have a closer look at those dispositions in the next two sections, but first I would like to adduce some preliminary evidence that my second class of counterexamples is a wide-spread phenomenon indeed.

There is a simple and generally effective way to pick out a disposition, even one that does not have a name yet: take a transitive verb, say, 'wash' or 'read', and suffix it with '-able' or '-ible' to yield an adjective, say, 'washable' or 'readable'. You may have encountered these two adjectives before, but it is easy to form new ones using the familiar suffixes: thus a pot may be smashable, a path walkable, and a game of chess winnable. Do these adjectives, familiar or new, pick out dispositions? It would seem so: for if 'vulnerable', 'irascible' and 'inflammable' succeed in picking out dispositions, there is (prima facie) no reason why 'washable' or 'walkable' should not. In general, it seems that we can reliably form disposition terms following this simple recipe: take a verb, add '-able'/'-ible', and you have a disposition-ascribing adjective. Note, however, that the recipe requires one verb, and one only; the suffixes are one-place operators. But dispositionality, on the standard model, is a two-place operator: to characterize a disposition, we need a stimulus and a manifestation. It seems clear that in our simple recipe, the verb gives us the manifestation, not the stimulus condition. If the dispositions thus picked out have a stimulus condition, where does it come from? Are terms such as 'washable' or 'walkable' systematically ambiguous between different dispositions with different stimuli? That claim would certainly need further support. A more plausible answer is that these dispositions are like loquaciousness and irascibility – that they too have no stimulus condition, even if it may be the case that each one of their manifestations has a cause.

It should be noted that my recipe does not always yield disposition ascriptions: witness such adjectives as 'regrettable', 'admissible', or 'honourable'. A regrettable state of affairs is not one that would or could be regretted (under certain circumstances), but one that may or should be regretted; an honourable person is not disposed to be honoured (under certain circumstances), but rather is such that she may or ought to be honoured. In other words, the suffixes '-able'/'-ible' can take a deontic meaning instead of the circumstantial one that they have when forming disposition terms. And yet there seems to be a close connection between these two uses of the suffixes, and cases that are not very easily assigned to either side: is an adorable baby, for instance, one that is disposed to be adored, or one that may or should be adored? This connection between disposition-ascribing terms and terms expressing deontic modality is a puzzle for the standard model, for it is unclear how the counterfactual conditional relates to deontic modality. A solution to this puzzle will be offered later, in the course of presenting my positive account (section 2.4).

For now, I will continue with my negative argument. For as I said, there is a tendency to view the so-called philosophers' dispositions as central, and take them to be adequately characterized by the standard model and the counterfactual conditional. The following two sections are devoted to showing that, whether or not the philosophers' dispositions are indeed central (I see no reason why they should be), they too are not adequately characterized by the standard model. Indeed, fragility, that favourite example of a philosophers' disposition, falls into my second category of counterexamples to the standard model, together with irascibility and loquaciousness.

2.2.2 Philosophers' Dispositions: A Closer Look at Fragility

Fragility, it has been said, is the disposition to break if struck. I take it that the manifestation is relatively unproblematic. The stimulus, however, is not – or so I will now argue.

A fragile glass may manifest its fragility in breaking upon being hit with a spoon, being dropped onto the floor, being sung to by a soprano, or being subjected to pressure over a period of time. Fragile parchments break upon being merely touched, and a fragile old wooden chair may split when transferred into a different temperature. Only the very first example – hitting with a spoon – would qualify as 'striking' in the ordinary sense.

It may be replied that 'striking' was never intended to be used in the ordinary sense. Rather, it is to be understood as a term of art covering the different yet similar processes that can stimulate the manifestation of fragility. (Some authors, I presume for this reason, use 'stressed' instead of 'struck'.) I happily concede that it does not matter whether or not we have a word that covers all these processes; all that matters is that they have something in common. For if they do not, then again there is no one property to serve as the stimulus condition of fragility, and no antecedent for the counterfactual that characterized it. What, then, do the different processes have in common?

A first and simple answer is that they are all normal or non-extreme causes of objects' breaking; they are clearly different from such abnormal and extreme causes of breaking as being hit by a meteor or being run over by a bulldozer. The fragile objects, on this characterization, will be those that break from normal, as opposed to extreme, causes. But then the stimulus condition is no longer doing any work: the only qualitative characterization that entered into the definition of fragility is breaking. We can say that fragility is the disposition to break, and that to be disposed to F is to F from normal causes of F-ing. I think such a characterization is close to being correct. But it has no role for a stimulus condition; it does not save the standard model.

To save the standard model, then, we would need to find something that these processes have in common independently of their being (normal) causes of breakings. At this point, the defender of the standard model may point out that it is not the philosopher's job to tell us what it is that they have in common: we should leave it up to the physicist, or perhaps the material scientist. The philosopher merely needs to claim that the processes I have mentioned are, at bottom, all species of some uniform kind of physical process; *what* kind of physical process this is, is not a question she needs to answer.

Note that it would not be enough for science to tell us that hitting and dropping, heating and singing are indeed all species of some common genus – of course they are: they are all physical processes. Rather, we need to be told that these processes, or indeed all the processes that can trigger the manifestation of an object's fragility, are the *only* species of their common genus. For if some other kind of process – say, washing – fell under the same genus but did not trigger manifestations of fragility, then that genus would be worthless as a stimulus condition. Take my observation that all the processes I have mentioned are physical processes: this does not enable us to formulate a stimulus condition for fragility, 'being subjected to a physical process', for there are many physical processes that do not trigger manifestations of fragility.

On the strategy under consideration, the defense of the standard model is hostage to actual empirical discovery. If physics or material science tell us that there is, after all, nothing that hitting and dropping have in common with heating and singing (etc.) but do not have in common with anything else, then that strategy fails. And indeed I am told that this is the case: the only thing these processes do have in common is their being causes of breakings (so we are back to the first answer). But I do not need to rest my argument on this empirical observation. For suppose for a moment that science did tell us what the standard model needs: that the processes which trigger manifestations of fragility are (the only) species of some genus of physical process. That would seem to be mere luck. We are perfectly able to pick out properties such as fragility with our dispositional terms, without any presumption as to, or any knowledge of, there being a uniform stimulus condition. Why then should we expect that all of the properties thus picked out do in fact possess a unified stimulus condition? Why should we expect to be lucky in each and every case? There is no good general reason for such an expectation, and a fortiori no reason to build that expectation into our account of dispositions.

2.2.3 Quantities and Counterfactuals

The previous section showed that the stimulus condition for a disposition such as fragility is qualitatively too varied to form one stimulus condition. In this section I will argue that even with any one qualitatively unified stimulus condition – such as striking in the ordinary sense – we will not normally obtain a counterfactual conditional that could adequately characterize the disposition. The reason is that the putative stimulus conditions for philosophers' dispositions are typically quantities: determinable properties that come with a continuous scale of determinates, ordered by a relation of 'greater than'. For instance, striking, if understood as the exertion of mechanical force, is a determinable with such determinates as 'striking with a force of exactly 8.35 N'. In this section, I am going to argue that quantities like striking fail to embed in a counterfactual conditional in the way that would be required for characterizing a disposition.

The argument applies to two kinds of view. First, the view that I have already argued against in the preceding section, according to which the stimulus condition of a disposition is a (qualitatively) uniform kind of process; my argument will show that even a uniform stimulus condition cannot save the standard model, if it is a quantity. Second, the view that accepts my argument in the preceding section but draws from it the conclusion that fragility is a multi-track disposition, characterized by a range of counterfactuals such as 'if x were struck, x would break', 'if x were subjected to pressure over a period of time, x would break', etc. Applied to that second view, my argument will show that if fragility is multi-track, then it has infinitely many 'tracks'. (A view that accepts this conclusion will be discussed in the next section, 2.3.1.)

Take, then, the counterfactual 'if x were struck, x would break' as a (partial) characterization of fragility. That counterfactual is not quite enough. Just about anything would break if struck with some sufficient force (my desk would break if struck by a large hammer); just about nothing would break if struck with any force whatsoever (even a fragile parchment may withstand a very light touch). That is not what was meant, of course: 'struck' just abbreviates 'struck with the appropriate force'. However, that merely shifts the problem: what is 'appropriate'? We may assume that the appropriate force is to set a threshold: a fragile object breaks if struck with *at least* such-and-such force. Presumably, that threshold will be fixed by context, and may be fixed differently by different contexts. What I will have to say holds with regard to any one context; so I will omit mention of contexts in the following and ask the reader to understand all that I say as indexed to one and the same context. Given a particular context, then, where exactly is the threshold to be set?

Note, first, that each fragile object comes with its own (context-independent) threshold. Some very fragile objects would break (finks, masks and antidotes aside) if struck with 1N or more, but not if struck with less; others may break if struck with 5N or more but not if struck with less; and finally, there will be an upper bound, *m*, such that the least fragile objects – those that we are just about still prepared to call fragile (in the context at issue) – would break if struck with *m* or more, but not if struck with less. Anything that would break, finks and masks aside, only if struck with a force greater than *m*, will no longer qualify as fragile. So we get an interval from the most fragile objects' threshold to that of the least fragile ones. Let us pretend, for the sake of a definite example, that this interval goes from 0.1N to 10N. Where in that interval is the threshold to be set for fragility, the property shared by all those many objects?

There are three options.

First option: set the threshold at 0.1N, with the most fragile objects. Fragility then is the disposition to break if struck with at least 0.1N; the counterfactual that holds (finks and masks aside) of any fragile object *x* is: if *x* were struck with at least 0.1N, then *x* would break.

But that counterfactual is not true of any but the most fragile objects. Take a medium-fragile object, one that would break if struck with at least 5N, but not if struck with less. It is not true of that object that it would break if struck with at least 0.1N: there are some situations, e.g. one where it is struck with exactly 2N, where that object is struck with at least 0.1N but does not break. Finks and masks play no role in this.

Second option: set the threshold at 10N, with the least fragile objects. Fragility, then, is the disposition to break if struck with at least 10N; the counterfactual that holds (finks and masks aside) of any fragile object *x* is: if *x* were struck with at least 10N, then *x* would break.

This counterfactual, to be sure, is true of all the fragile objects. It does not, however, adequately capture their shared fragility. Take, again, our medium-fragile object, which would break if struck with at least 5N, but not if struck with less. Strike it with a force of 8N; nothing devious is going on; the object breaks. Has it thereby manifested its fragility? Surely we should say that it has. But on the present proposal, we cannot say that: for fragility, on the present proposal, is the disposition to break if struck with at least 10N, and our object has not been struck with at least 10N. In fact, on the present proposal, it is almost only those breakings which are not specific to fragile objects that count as manifestations of fragility: breaking upon being struck with 10.1N, 11N or 100N. Breaking upon being struck with any force less than 10N is excluded from being a manifestation of fragility; breaking upon being struck with any force of 10N or more are included among the manifestations of fragility. The proposal gets it exactly the wrong way around.

The first option, then, classified too few objects as fragile; the second option classified the wrong breakings as manifestation of fragility. A third option is to set the threshold somewhere in between. But it should be obvious that far from solving the problems of the two earlier options, this will incur the problems of both.

My three options have been exhaustive, assuming (i) that the threshold for fragility must be set somewhere in the interval from 0.1 to 10N, and (ii) that it must be set using 'at least'. I think that (i) is too obvious to be discussed; but (ii), while attractive, is not entirely obvious. So let me look at an alternative to 'at least'.

The obvious alternative is 'at most'. Here we can go through the same three options as above, incurring essentially the same problems.

Saying that the fragile objects are those that break if struck with at most 0.1N is obviously not an option: for this counterfactual is not even true of the very most fragile objects. Saying that the fragile objects are those that break if struck with at most 10N *prima facie* looks better, but has the same undesirable consequence: for this requires breaking upon being struck with *any* force below (and including) 10N, for instance, a force of 0.01N (or even less). But again, this will not be true of any but the very most fragile objects – and perhaps not even of them, since there will likely be some positive force, and a fortiori some positive force below 10N, that cannot even cause the most fragile object to break. And again, setting the threshold anywhere in between will not help.

There may be a few more moves available to the defender of the standard model, but as far as I can see they will all be subject to the exact same problem: quantities and thresholds just do not embed in counterfactuals in the way required for the standard model to work. If the defender of the standard view wants to dispute this, the burden of proof is on her.

2.3 Adjusting the Standard View

2.3.1 Multi-Track Dispositions

My arguments so far have shown that for at least a very large range of dispositions, there is no one counterfactual to characterize any one of those dispositions, and so the modality inherent in those dispositional properties cannot be that of any one counterfactual.

That, of course, is not yet to refute the standard model altogether. My argument has shown that there is no one counterfactual to characterize any one of these dispositions; it has not shown that there are not very many counterfactuals that characterize any one disposition, nor that there are not other dispositions that *are* each characterized by one counterfactual. The defender of the standard model may embrace the conclusion of my argument in the previous section, and say that dispositions are indeed multi-track with infinitely many 'tracks'. Fragility, then, will be characterized not by any one counterfactual such as 'if x were struck, x would break', but rather by infinitely many counterfactuals such as 'if x were struck with a force of at least 8.35N, it would break', 'if x were subjected to a pressure of at least 2.46N over at least 10 hours, it would break', etc. The 'real' dispositional properties will be those that are characterized by only one such conditional: the disposition to break if struck with a force of at least 8.35N, for instance. A certain range of these *specific* dispositions will be characteristically possessed by the objects that we call 'fragile'; but there is nothing substantial that all of these dispositions have in common with each other and nothing else, nothing to set them apart from the many similar dispositions (e.g., the disposition to break if struck with at least 11.67N) that do not count in favour of an object being fragile. What this goes to show, the defender of the multi-track view will say, is merely that our everyday dispositional terms such as 'fragile' do not pick out natural properties (or even nearly natural properties), but a relatively random selection of more natural properties that we happen to be interested in. (For a strategy along these lines, see e.g. Lewis 1997 and Cross 2005.)

It is clear that such a view avoids my second challenge, the lack of a uniform stimulus condition. If there are infinitely many counterfactuals, there may be infinitely many stimulus conditions to go into their antecedents; my observation is built into the view from the start. My final challenge concerning the quantitative nature of any one putative stimulus condition can also be answered. To see how, consider first an unsuccessful strategy for the multi-track view. The strategy is to claim that there is a (non-natural) set of specific dispositions for which 'fragile' is an umbrella term; and that a statement of the form '*x* is fragile' will be true just in case the object referred to by '*x*' possesses all the specific dispositions in this set. That strategy faces a problem of the same kind as the one I have pointed out in the previous section: only the very most fragile object will satisfy the predicate '... is fragile'. A better strategy is suggested by Manley and Wasserman (2008): to satisfy '... is fragile', an object has to possess *most*, but not all, of the specific dispositions (or, equivalently, to satisfy most but not all of the specific counterfactuals).

It is when we look back to my very first objection that the multi-track view has some trouble in solving the standard view's problems. In section 2.2.1, I distinguished between two types of counterexamples: active dispositions and dispositions with no uniform stimulus condition. The latter we have already seen to be solved by the present view. But it is less clear how the view is to accommodate the former: no counterfactuals, no matter how specific, will be true of a disposition the manifestation of which comes about spontaneously. Whatever causal input is given to an object possessing such a disposition, the object might or might not manifest its disposition, hence it is not the case that the object *would* manifest its disposition.

This, then, is a first worry about the multi-track view. It does a fairly good job at answering the objections to the standard view, but it cannot quite complete that job. But there is a second worry about it, which goes deeper than the first; it concerns the motivation of the multi-track view.

Here is a very simple and schematic story of how we come to adopt the multi-track view. When we begin to philosophize about dispositions, we need some mutual understanding of what it is that we are theorizing about. Such mutual understanding is very often achieved by agreeing on certain paradigm examples: 'By a "disposition", I mean a property such as fragility, solubility, irascibility – that kind of thing.' (If we are interested in discussing the metaphysics of dispositions, the mutual understanding had better not be achieved by characterizing the nature of dispositions – for that is exactly what is at issue.) With the paradigm examples in hand, we then develop a first and tentative schema that characterizes them, and which may be hoped to generalize and help us achieve a more general account of dispositions. This schema, of course, has historically been the (counterfactual) conditional, and with it a two-place dispositionality operator: 'Fragility seems to be the disposition to break if struck, so let's think of dispositions in general as dispositions to ... if ...'. It is then discovered that the schema does not quite fit the original examples – as I have argued in sections 2.2.2–2.2.3, fragility is not (adequately characterized as) the disposition to break if struck. The multi-track view then asks us to retain the schema while dropping the initial examples: dispositions in general *are* counterfactual-like dispositions to ... if ..., but fragility is not properly speaking one of them. But what reason do we have to retain the schema if it turns out not to characterize out initial paradigmatic examples? If what we characterize is not fragility and 'that kind of thing', have we not simply changed the topic?

In short: perhaps the multi-track view *can* save a counterfactual-like conception of dispositions (though my first worry should raise some doubts about that); but why *should* we save it at the cost of giving up those dispositions that we started with? Why *should* we take the specific disposition to break if struck with at least 8.35N (at such-and-such an angle, etc.) to be more fundamental than the more familiar disposition of fragility *just because* the former can, and the latter cannot, be adequately characterized by one counterfactual conditional? This would presuppose precisely what is at issue: that a disposition, to be 'real', needs to be counterfactual-like.

2.3.2 Weak Counterfactuals

I have mentioned Manley and Wasserman (2008) as proponents of a multi-track view. Their view, however, is open to a slightly different interpretation, which would evade my doubts about the motivation of the standard view. Roughly, the strategy is this: instead of saying that *most* of a set of counterfactuals have to be true of *x* for *x* to count as fragile, they build the modal force of 'most' into the counterfactual itself. A fragile thing is not one that would break in *all* cases where it is struck (as the counterfactual proper would have it), but rather one that would break in *most* cases where it is struck. A linguistic construction that bears the modal force of 'most' rather than 'all' has been supplied (independently, I believe) by Fara (2005): the habitual '*x* breaks when struck'. Whatever differences of details there may be between Fara's and Manley and Wasserman's accounts, both attempt to solve problems for the standard model by weakening the counterfactual. I will focus on Manley and Wasserman's treatment, which is more explicitly metaphysical.

As the title of their paper indicates, Manley and Wasserman (2008) are concerned with the link between dispositions and conditionals. They quickly give up on a one disposition–one conditional model (though, again, not for the reasons I have mentioned, but rather because of finks and masks). Instead, they take up a suggestion from Lewis (1997) and focus on very specific conditionals such as

(IC) N would break if dropped on Earth from *over* half a metre onto a surface with a Shore measurement *greater* than 50A, through a substance with a density *less than* 50 kg/m^3. (p. 67)

or

(EC) N would break if dropped on Earth from *exactly* one meter onto a surface with a shore measurement of *exactly* 90A, through a substance with a density of *exactly* 1.2 kg/m^3.

I have so far focused on conditionals of the (IC) variety, but Manley and Wasserman concentrate on (EC)-type conditionals.

Clearly, (IC) and (EC) are very much like the specific conditionals that I considered in the preceding section, spelled out in more detail. Manley and Wasserman note, as I did, that dispositions come in degrees, and that the more fragile objects will have a greater number of specific conditionals true of them than the less fragile one. They propose that the fragile objects are those of which *most* or, rather, a 'suitable proportion' of a certain class of conditionals such as (EC) are true. This, of course, sounds just like a version of the multi-track view. But Manley and Wasserman then go on to give a slightly different formulation.

Call the antecedents of conditionals such as (EC) C-cases; thus (EC) gives us the C-case of being dropped on Earth from exactly one metre onto a surface with a Shore measurement of exactly 90A, through a substance with a density of exactly 1.2 kg/m^3. Then we can restate the proposal to say that the fragile objects are those that break in a suitable proportion of the C-cases for fragility; or in general,

(PROP) N is disposed to M when C if and only if N would M in some suitable proportion of C-cases. (p. 76)

Here the 'suitable proportion' is quantifying not over counterfactual conditionals (as it does on the multi-track view), but rather over C-cases. But what exactly is a C-case? Here is what Manley and Wasserman tell us: the term 'C-case', or 'stimulus condition case', is to be used 'for every precise combination of values for heights, Shore measurements, densities of the medium, and so forth.' (p. 74f.) Metaphysically, *C*-cases 'are to be construed less coarsely than worlds: a world may contain many dropping-cases.' (p. 74, fn. 19) The best candidate for the ontology of *C*-cases would seem to be a centered world: a triple of a world, an object, and a time.

C-cases, then, are not general: they are very specific possible situations. (PROP), it seems, does not quantify over conditionals; it quantifies over possible situations. But that is just what a counterfactual does (at least on the standard interpretation). The difference is that a counterfactual applies *universal* quantification to a class of possible situations, or possible worlds: 'if x were struck, then x would break' is true just in case x breaks in *all* the relevant possible situations in which it is struck. (PROP) replaces this universal quantification by a weaker, proportional, kind of quantification: it is not in all, but in most or some suitable proportion of possible striking-situations that a fragile glass breaks.

I have taken such pains to reconstruct the alternative interpretation of Manley and Wasserman's account because it avoids my main objection to the multi-track view. The objection was that the multi-track view gives up on properties such as fragility in favour of a schema the motivation for which was derived from such properties. Unlike the multi-track view, Manley and Wasserman do give us an account of properties such as fragility as more than a random collection of other, better behaved, dispositions. For an object to be fragile is for it to break in most (or in a suitable proportion) of the *C*-cases for fragility. The next question, then, is: can that characterization answer my earlier objections in sections 2.2.1-2.2.3? And in particular, how are the '*C*-cases for fragility' determined?

To begin again with my third objection, the quantitative factorization of qualitatively uniform stimulus conditions that I have pointed out in the preceding section is not a problem. As far as striking is concerned, the threshold may be set as low as we like, say at 0.1N; it will still be true that the fragile objects break in a suitable proportion of cases where they are struck with at least 0.1N.

The problem pointed out in my first and second objection, the lack of a uniform stimulus condition, is noted in general by Manley and Wasserman (though not clearly distinguished from genuinely 'active' dispositions), even as potentially applying to fragility (p. 72), and they claim to have solved it:

> Because the *C*-cases in our domain need not be restricted in any way, absent stimulus conditions are not a problem for (PROP). We can simply allow that N is loquacious just in case N would talk in a suitable proportion of situations—any situations at all. (Manley and Wasserman 2008, 77)

The idea, then, is that where a disposition does not come with a stimulus condition, '*C*' in (PROP) is to be understood as the trivial condition, applying to all cases whatsoever; we may replace it with the trivial truth $\top$.

Manley and Wasserman believe that they have well integrated stimulus-less dispositions into their picture in this way. But the divide between the stimulus-less dispositions, those with no restriction on *C*, and the dispositions that do have a stimulus and a restriction on *C*, is wider than expected. For the 'suitable proportion' of cases, the new modal force that has replaced the old one of the counterfactual conditional, will have to be quite different depending on whether there is a restriction on *C* or not. Suppose for a moment that fragility could after all be given a uniform stimulus condition ('striking'). In that case the *C*-cases will include only cases where an object is subjected to that stimulus condition: cases where it is struck, hit, dropped, etc. A fragile object will break in *most* of these cases. If, on the other hand, I am correct and there is no uniform stimulus condition for fragility, then its *C*-cases will include many more than the ones

I have just mentioned: they will include cases where the object in question is sitting on a shelf or otherwise undisturbed. Even a very fragile object will break only in *a few* of those cases. The modal force of fragility, then, depends very heavily on whether or not it can be given a (uniform) stimulus condition. And in general, the modal force, which is the 'suitable proportion' of cases, will be very different for those dispositions that do have a stimulus condition from those dispositions that do not. It will be very close to 'most' in the case of dispositions with a stimulus, and very close to 'a few' for those without one. And this, I submit, is an entirely ad hoc difference: the modal force of 'fragile' does not depend on whether or not we can formulate a uniform stimulus condition for it, and the modal force of cases that have no uniform stimulus condition is not so vastly different from those that happen to have one.

The problem arises from a tension between two ideas. On the one hand, Manley and Wasserman recognize that we cannot have a stimulus condition, or a restriction on C-cases, for all dispositions. On the other hand, they do postulate a stimulus condition, or a restriction on C-cases, for some dispositions. Adopting both claims leads to the ad hoc difference that I have complained about, and I have argued that the former claim is true. The solution is obvious: abandon the stimulus condition, or the restriction on C-cases, altogether.

It is time now to consider an alternative: one that does away altogether with the stimulus condition in characterizing what a disposition is, and understands dispositions not in terms of a two-place operator 'disposed to ... if ...', but in terms of a one-place operator 'disposed to ...'. The next section will be devoted to formulating this alternative account.

2.4 The Alternative Model: Dispositions and Possibility

2.4.1 Preliminary Remarks

Dispositionality, then, is to be treated as a one-place operator. Fragility is the disposition to break, fullstop. This does not yet answer the question: which modality *is* it that is inherent in, or taht characterizes, dispositions in general and fragility in particular?

The most natural suggestion is that it is some kind of possibility. This suggestion is confirmed, and a special kind of possibility is highlighted, when we look up the standard examples of dispositional terms in the *Oxford English Dictionary*. Fragile things, we are told there, are those that are 'liable to break or be broken; ... easily destroyed'. 'Soluble' means 'capable of being melted or dissolved', and to be 'irascible' is to be 'easily provoked to anger or resentment'.

From the dictionary definitions we can take two suggestions. One is just plain possibility: this is the modality inherent in 'liable to break' or 'capable of being melted'. Not quite metaphysical possibility, perhaps, but possibility restricted by laws of nature and some other constraints to the effect that things take their normal course – the kind of possibility that we express in ordinary language when we say such things as 'This can break', or 'You can grow hydrangeas on this soil'.

Plain possibility, or possibility restricted in this ordinary way, cannot be quite right, however, at least not for the case of fragility. For certainly not everything that *can* break is fragile: plant-pots, brick-stones and even bridges made of steel *can* break, but they are not therefore fragile. (Our standards for fragility may shift with context, but they do not shift in the same way as those for 'can'.) What distinguishes the fragile things from other things that can break? The dictionary definitions suggest that the former, but not the latter, can *easily* break or be broken. Similarly, almost everyone *can* be provoked to anger, but an irascible person can be *easily* provoked. The naturalness of 'easily' in connection with dispositions is confirmed by its slipping into characterizations of a disposition even when given in the stimulus–manifestation scheme: witness Alexander Bird's characterization of fragility as 'the disposition to break easily when stressed' (Bird 2007, 19), or Stephen Mumford's introductory remark in his book that 'although virtually all objects are breakable, it is only those which break easily ... that are called fragile' (Mumford 1998, 5). Notably, the term does not play any role in either author's official characterization of dispositions.

The suggestion, then, is that the modality involved, at least in the dispositions of fragility and irascibility is not just plain possibility, but easy possibility. I will look at easy possibility in some more detail in a moment. First, let me note that, whatever our conception of easy possibility, it is of course not the case that an object is fragile (at time t) just in case there is an easy possibility (at t) that it breaks. Pack a fragile glass in styrofoam, and there is no easy possibility that it breaks; place a rock in front of a bulldozer, and there is an easy possibility that it breaks. This, however, should not be a problem for the proposal: as we have noted, dispositions are properties of individual objects and as such behave differently from a more global modality such as (easy) possibility. Let me offer two ways of stressing this point, one for those without, and one for those with, reductive inclinations.

The non-reductionist version. If you are willing to accept that dispositions are modal properties, and that modal properties should not be reduced to any other modality, then the search for 'the modality involved in dispositions' is not a search for the precise truth-conditions of disposition ascriptions. It is, rather, a search for certain parallels that allow us to formulate a rough, though not an entirely adequate, model for dispositions. You may have thought, until now, that that model was going to be that of counterfactual conditionals: fragile things are such that they would break if struck. Of course, fragility is a matter of how things stand with an object itself, and the counterfactual takes into account not only how things stand with a particular object, but also how things stand outside that object; so we need not be surprised if the truth of 'x is fragile' and the truth of 'x would break if struck' do not always go together. If this has been your view so far, then my recommendation is to simply transfer it onto easy possibility: a fragile thing is one that can easily break; but of course, fragility is a matter of how things stand with an object, and easy possibility concerns not only that but also how things stand outside that object; so we should not be surprised if the truth of 'x is fragile' and the truth of 'x can easily break' do not always go together. If you have been content to treat counterfactuals as providing a rough model and an interesting parallel to dispositions, then I submit you should be content to treat easy possibility as providing merely a rough model and an interesting parallel to dispositions. My claim is that it is the better model and the more illuminating parallel.

The reductionist version. Much of the literature on dispositions and conditionals has been concerned with more than just rough models and illuminating parallels. It has accepted that, as properties of objects, dispositions are somewhat different from just any counterfactual, as is shown by the possibility of finks, masks, antidotes and their like. It has then tried to manipulate the counterfactuals accordingly. Lewis (1997), for instance, builds the condition that the individual retain its physical base for the disposition in question into the antecedent of the counterfactual. If, as I suggest, we drop the stimulus condition and with it the counterfactual, and adopt instead a one-place dispositionality operator akin to possibility, that strategy is of course not an option. But there is a different strategy, also suggested by Lewis and explicitly aimed at a treatment of 'can': instead of supplying implicit content to what is in the scope of a modal operator, the strategy is to manipulate the accessibility or closeness relation.

Following cues from Lewis (1979) and Lewis (1976), we can try homing in on the property-hood of dispositions by treating dispositions as a special kind of restricted possibility: restricted not only by the actual laws of nature and the like, but also by the intrinsic constitution of the object to which we are ascribing the disposition. It will be true that an object *x* is disposed to F just in case the object F's in at least one close world; and into the conditions for a world to count as close we build, among other things, the condition that the intrinsic constitution of the object in question, or even only parts of that intrinsic constitution (the dispositions's physical base), is held fixed, while contingent circumstances external to the object are not. Thus a glass's molecular constitution and the fact that it has a slight crack will be the same in all the close worlds, but the fact that it is packed in styrofoam as opposed to standing near the edge of a table will not.[1]

2.4.2 Easy Possibility and the Case of Fragility

So much for the reductionist or non-reductionist options. Let us now look at the modality that is suggested by the dictionary definitions I have cited: easy possibility.

Easy possibility is not a stranger in philosophy: Timothy Williamson has used it in his account of knowledge (see Williamson 2000, also Williamson 1994 and Sainsbury 1997), and Christopher Peacocke (1999) has appealed to it in his compatibilist account of freedeom. For our purposes, it is enough to say that there is an easy possibility that *p* just in case there is at least one close world in which *p*, where closeness is a matter of similar initial conditions (not of the world, but of a temporal portion of it).

[1] In the next chapter, I will provide reasons to prefer the non-reductionist version, and will also explicitly address the Lewisian semantics of 'can'.

For instance, suppose that a glass is standing close to the edge of a table, and with a careless movement I almost hit it. There is an easy possibility that the glass breaks because in some world with very similar initial conditions – that is, a world in which the glass is in more or less the same position, and I make my movement from more or less the same position with more or less the same lack of care – my movement is just so slightly different that I do hit the glass, it falls, and it breaks.

Is easy possibility – as understood by these philosophers – the right model to characterize the modality inherent in fragility? I will now argue that it is not. This is not to say that the dictionary definitions are incorrect; 'easily' may be used in more than one way in natural language.

A first point to note is that easy possibility, like every species of 'real' possibility (and unlike, for instance, deontic possibility), is *centered* on the actual world: the actual world is always close to itself (its initial conditions will be similar to themselves, whatever measure of similarity we are applying). As a consequence, whenever p is actually the case, it is also the case that it could easily have been that p. Fragility, of course, does not behave in the same way: from an object's breaking in actuality we cannot infer that it was fragile. The non-reductionist may try to explain this away as one of the many inadequacies of the model; the reductionist may try to formulate a different closeness relation, one that is not centered on actuality. Both of these strategies will meet with obstacles, however, for centering is only a symptom of a more general problem: that one world where an object breaks is enough, on the current proposal, for that object to be fragile, and that no restriction on closeness can exclude that there are unwanted breaking worlds among the close worlds.

To see this general problem, consider the case that Manley and Wasserman (2008) call 'Achilles' heel':

> consider a sturdy concrete block that, like Achilles, is almost entirely immune to harm. ... But, like Achilles, the block has a weak spot. If it is dropped onto a *particular* corner at *just* the right angle with *exactly* the right amount of force, an amazing chain reaction will cause it to break. (Manley and Wasserman 2008, 67)

We can easily imagine that the block being dropped in this particular way is an entirely normal type of event – just as normal as dropping it onto any other corner, at any other angle, and with many other forces. On the easy-possibility conception of possibility, that block would then have to count as fragile: whatever exactly we may want to say about closeness, there is one close possibility in which the block breaks, and one is enough. But it seems clear that the imagined block is not fragile.

This is an objection not only to the reductionist version of the easy-possibility view. The non-reductionist takes easy possibility to be only a rough and not fully adequate model; but this is not to say that it can count anything, including Achilles' heels, as one of the many inadequacies of that model. For the motivation for the non-reductionist view is the idea that a disposition is a matter of how things stand with the object (the block) itself, while (easy) possibility concerns how things stand with the object and the world outside it. The block's weak spot, however, *is* how things stand with the block itself; it is not an interference of its surroundings.

Achilles' heels, then, show that the modality of 'fragile' is not that of easy possibility. One close breaking-world is not enough; for an object to be fragile there need to be more ways of breaking it. This suggests a strategy akin to that of Manley and Wasserman: instead of taking the all-or-nothing modalities of possibility or the counterfactual, we can appeal to a quantitative measure, or a proportion-like modality. One world is not enough; but a few worlds are. The proportion-like modality to which we are appealing is itself a kind of possibility, but it is a *graded* possibility. For an object x to be fragile, it has to be not merely possible that x breaks, or x has to break not merely in some world. Rather, it has to be *possible enough*, or sufficiently likely, that x break, or x has to break in sufficiently many of the close worlds.

The close worlds may now again be understood in the standard way; centering is no longer a threat. For an object's breaking in actuality does not guarantee its breaking in sufficiently many close worlds even if the actual world itself is a close one; it might just be the odd one out in this particular case. A more important question is how to understand 'sufficiently many'. It is certainly not to be understood as requiring a *majority* of cases. As I have noted earlier (in section 2.3.2), when there is no stimulus condition to restrict the relevant worlds, the proportion will be closer to 'a few' than to 'most', and accordingly closer to possibility than necessity (or the counterfactual conditional). Just what proportion of worlds is sufficient will be a matter of context. But that is as it should be: disposition terms are context-dependent.

Before I turn to other dispositions, let me briefly review the dialectic of this section up to now. I suggested that the modal force of dispositions was possibility. Ordinary possibility, however, was found wanting. In possible-worlds terms, it is not enough for an object to be fragile that it breaks in some possible world, even some possible world that is close according to ordinary standards. Easy possibility constituted one reaction to this: drastically restrict the worlds that count as close, but retain the modal force of possibility (one world is enough). Achilles' heels gave us a reason to reject this solution, and turn instead to a different reaction: retain the ordinary closeness relation, but revise the modal force. The modal force of 'fragile', I have suggested, is not that of plain possibility, but rather a graded possibility; it is not an existential quantification over possible situations, but requires a certain minimal proportion of them.

2.4.3 More Examples

We have not yet seen any reason to believe that the model of fragility will be transferable to all dispositions. So in looking at other examples, we should keep in mind the three options – plain possibility, easy possibility, graded possibility – that I have set out, and try to establish which of them are best suited to capture the modality of the various dispositions that have been considered in the course of this chapter. There should, of course, be no ad hoc variations in modal force between different dispositions (such ad hoc variation was, after all, one of the core objections raised against Manley and Wasserman). But there may be variations that are grounded in the behaviour of the dispositional terms themselves; in fact, the ability to capture such variations should count in favour of a possibility-based account.

Let me begin by looking at the class of examples that has been emphasized in the earlier parts of my argument (section 2.2.1). It was pointed out there how easily we can form dispositional terms using merely a verb and a suffix -able/-ible: 'readable', 'smashable', 'walkable', 'winnable', 'breakable'. Clearly my account, treating dispositionality in terms of a one-place operator, is better suited to accommodate these examples than the standard model was. But is it plain possibility, easy possibility, or the graded possibility of 'fragile', that is at work here?

It is instructive to contrast 'breakable' with 'fragile'. Consider our block with an Achilles' heel again: it is not fragile, but is it breakable? It is at least much less obvious that the answer is no. Moreover, consider the inference from actual manifestation to disposition that I rejected earlier: from the fact that a thing does break in actuality, we cannot infer that it is fragile; but we can, it seems, infer that it is breakable. The same appears to hold for all or most of the other terms I have mentioned: walking a path proves that it is walkable, smashing a pot that it is smashable, winning a game that it is winnable, and (my) reading a text in fine print proves that it is readable (for me). But this inference is typical of the modal force of possibility proper: an object's F-ing in actuality does not prove that it F-es in a sufficient proportion of cases, but it does prove that it F-es in at least one case.

This suggests that the modal force of disposition terms that are formed according to the 'F + able' schema display a weaker modal force than 'fragile'. Their modal force is possibility proper: one world in which an object is F'ed is enough to make it 'F-able'.

The possibility involved, however, is not easy possibility, as is witnessed by the fact that we can prefix the terms by 'easily'. A chess player, looking at different games at various stages, may categorize them by distinguishing those that are easily winnable, those that are winnable but not easily so, and those that are not winnable at all. An 'F-able' object, then, is just one that can be F'ed, fullstop.

Am I, then, committed to ad hoc variations in modal force, exactly the kind of variations that I complained about when discussing Manley and Wasserman (2008)? No: for my variations are not ad hoc. First, I have given independent evidence that there is variation, in the inferential behaviour of terms such as 'breakable' as opposed to 'fragile'. Second, the variation in modal force that I am postulating are continuous: possibility proper can be understood as a limiting case of a minimal proportion – if *any* positive proportion will do, then one world is enough. All the variation is at one end of the spectrum, the possibility end and a proportion of 'a few'. Manley and Wasserman, on the other hand, had to stipulate a discontinuous variation between the two different ends of the spectrum: a proportion of 'most' for dispositions with a stimulus, and of 'a few' in dispositions without one, and nothing in between.

It is to be hoped that the variation I have admitted will be all the account needs to accommodate the varying modal force of different disposition terms; and that easy possibility, for instance, will not turn out to be the right model for some dispositions that I have not yet discussed.

The proposed account can nicely accommodate two further types of case that have also been mentioned in section 2.2.1. These are (i) active dispositions, and (ii) deontic disposition-like terms.

(i) Active dispositions, I have argued, require a one-place operator. But can we also make the distinction between active and passive dispositions in terms of a one-place operator? We can, and without introducing ad hoc variations in the operator itself: by building the distinction into the argument of that operator. I have so far characterized all dispositions in passive terms: fragility, for instance, as the disposition to *be broken*, which we may understand as the disposition to be *caused* to break. We can also imagine a disposition *to break*, which would be possessed by the fragile objects (as so far characterized), but also by objects which can break spontaneously, with no causal input whatsoever. We can, finally, imagine a disposition to *break spontaneously*, possessed by the latter objects but not the former ones. I take it to be an open question whether fragility is the disposition to break or the disposition to be broken, i.e. whether objects that tend to break spontaneously would thereby be counted as fragile. I observe, however, that the suffix '-able'/'-ible' appears to take only passive complements: to be F-able is always to be disposed to *be F'ed*.

(ii) I have noted that certain apparently dispositional terms pose a puzzle: 'regrettable', 'admissible' or 'honorable', for instance, seem to have a normative dimension. The proposed account can, again, nicely accommodate this by pointing to the well-established observation that in most expressions of possibility, circumstantial and deontic uses are closely tied together (so much so that some linguists use just one term for both of them, 'root possibility'). Once that is noted, it should not come as a surprise that disposition-forming expressions exhibit the same tendency to switch between circumstantial and deontic possibility.

2.4.4 Philosophers' Dispositions Again

We can now, finally, look at the remaining stock examples of the so-called philosophers' dispositions. I have left them until the end because they are the most obvious starting point for an objection to the account I have sketched: if not 'fragile', then perhaps 'soluble', 'inflammable' etc. do have a uniform stimulus condition that my account is unable to accommodate?

To begin with, I happily admit that in the case of solubility and inflammability, we *can* specify a qualitatively uniform stimulus: a (water-)soluble object is one that dissolves when immersed in water, an inflammable object is one that burns when exposed to heat. Both are subject only to the third of my objections against the standard model: their stimulus conditions can be factorized into determinate quantities just as easily as the condition of being struck. How much water does a soluble thing have to be immersed in to dissolve? and for how long? How high a temperature does an inflammable thing have to be exposed to in order to catch fire, and for how long? Of course, I have admitted that Manley and Wasserman's view is not subject to this problem; so in their framework, solubility *can* be given a stimulus condition. That, however, is not decisive: the dialectical situation has changed. I have argued in section 2.3.2 that supplying a stimulus condition for some dispositions but not others leads to adhocery in their respective modal force, and that most dispositions cannot be supplied with a stimulus conditions. If my account can cover the special cases of solubility and inflammability, that is good enough.

Take, then, the disposition of being soluble in water. There is no reason why the qualification ‘in water’ has to be part of a stimulus condition; it may as well be part of the manifestation. Water-solubility, then, is the disposition to dissolve in water; solubility unqualified is the disposition to dissolve, fullstop. Are we losing something essential by dropping ‘if immersed’? No: the manifestation already contains everything that the stimulus could have added. For what is it to dissolve, or to dissolve in water? It is, roughly, to disperse and mix with the surrounding matter, or to disperse and mix with surrounding water. There is no dissolving without immersion; adding ‘if immersed’ to ‘would dissolve’ is inert. Solubility can be understood very well without reference to a stimulus condition. And indeed this is how the Oxford English Dictionary defines it: to be soluble is to be ‘capable of being melted or dissolved’.

To be inflammable, according to the Oxford English Dictionary, is to be ‘capable of being inflamed; ... easily set on fire’. The strategy that I have used in the case of solubility does not seem to apply here: being exposed to heat is not as clearly part of being set on fire, as being immersed in water is part of dissolving in water. The strategy, rather, is one that has been foreshadowed in my discussion of fragility earlier on (section 2.2.2): it is a further fact about the world and the causal connections in it that inflammation is always caused by heat (if it is); if there were other objects that caught fire in a different way but just as easily, we would count those as inflammable too. There is no need to build that further fact into the disposition that we ascribe with the term ‘inflammable’.

So much for the sketch of a defense of my account. We are still left with the question of what the modal force of ‘soluble’, ‘inflammable’, etc. is. Is it plain possibility or the graded possibility that we saw at work in ‘fragile’? The previous sections have given us ways of testing both hypotheses. If from a substance’s being dissolved in water it follows that the substance is soluble, or from an object's being set on fire that it is inflammable, then 'soluble' or 'inflammable' belong with 'breakable' and 'readable', having the modal force of plain possibility. Alternatively, we may construct cases of Achilles' heels for solubility or inflammability: does one possible way of dissolving a substance, or setting an object on fire, count towards their solubility and inflammability, respectively (if suitably anchored in the substance or the object themselves)? These are some of the new questions that my proposal opens for future research. Abandoning the standard model is not only an end, it is the beginning of a promising new enterprise.

2.5 Dispositions and Laws

In my argument for an alternative model of dispositionality, I have concentrated on dispositions, such as fragility, that we ascribe in everyday contexts using ordinary-language dispositional terms. Some of them, such as 'soluble', are at least as important in scientific contexts as they are in everyday ones; but I have nevertheless assumed our grasp on them to be through those everyday terms. It may be objected that we have a very different grasp on dispositional properties: through science itself, through the properties and laws that it discovers. We should be wary to impose results obtained by reflection on our use of natural-language terms such as 'fragile' on properties such as electric charge or mass (or, if these are not themselves dispositional, on the dispositions that come with them). If our best scientific theories force onto us an understanding of these dispositions in accordance with the standard model, then so be it.

One theory that makes this objection particularly pressing is Dispositional Essentialism. According to Dispositional Essentialism, the laws of nature are entailed by the dispositional natures of the fundamental properties. For that entailment to work, the laws must already be encoded in the fundamental properties. And if the laws are of a form that closely corresponds to a conditional ('If something is an F, then it is also a G'), then the properties that encode them should have a conditional-like form too – a form that is provided by the standard model but not by my alternative model. But the worry may be raised not only by dispositional essentialist. Whatever one's views of the metaphysical priorities, there seems to be a close correlation between the dispositions possessed by objects that have a natural property such as charge, and certain laws concerning charge, such as Coulomb's Law. If we think that there are any dispositions, we should think that there are the dispositions that come with properties such as charge (whether or not they are identical with those natural properties), and that encode the laws concerning charge (whether or not they are the metaphysical grounds for those laws).

How, then, could the worry be alleviated? I agree, of course, that we should be careful not to impose any understanding of our ordinary-language terms on the properties that science deals with. I believe, however, that it is the standard model which is guilty of such imposition. The dispositions in question are *not* best understood in terms of the standard model; they are best understood in terms of the alternative model and its one-place dispositionality operator. In arguing for this conclusion, I will focus on Dispositional Essentialism, and in particular on the recent book-length treatment provided of it by Alexander Bird (2007). I will then briefly indicate why I think that the argument generalizes to all versions of the worry I have described.

2.5.1 Dispositional Essentialism and the Laws

Dispositional Essentialism is the view that '[a]t least some sparse, fundamental properties have dispositional essences' (Bird 2007, 45). This minimal characterization can be expanded in various ways, and various dispositional essentialists disagree on just how it should be expanded. Thus Alexander Bird holds, but Brian Ellis denies, that not only *some* but in fact *all* sparse, fundamental properties have dispositional essences; while Brian Ellis holds, and Alexander Bird is neutral on the question whether, the essentially dispositional properties are in turn essential to their bearers. I will here be concerned only with the minimal version of dispositional essentialism, and the argument that Bird (2007) gives for it.

Dispositional Essentialism can be understood in three simple steps:

First step: some fundamental properties have *essences*.

Second step: (some of) these essences are *dispositional*.

Third step: for a property P to be dispositional is for P to be characterized by a stimulus condition S and a manifestation condition M, which behave very much like a counterfactual conditional $Sx \mathrel{\Box\!\!\rightarrow} Mx$ (with a *ceteris paribus* clause).

(The third step obviously corresponds to the standard model of dispositionality.)

Note that nothing in any one step forces us to take the next one: one might hold that some properties have essences, but that these essences are not dispositional; or one might hold that they are dispositional, but give a non-standard account of dispositionality. I believe, however, that these three steps are shared by everyone who commits themselves to Dispositional Essentialism. So in what follows, I shall understand 'Dispositional Essentialism' (DE) as the conjunction of these three steps.

Now to DE's account of the laws. Laws can be viewed as a special kind of regularities. But what is it that makes a given regularity a law? According to the Lewisian best system account, it is merely its place as an axiom in the best system describing the world (Lewis 1994; for more on the best system account, see chapter 5.6). There is nothing special about the regularity itself; taken on its own, it does not differ from regularities that do not qualify as laws. Realists such as the dispositional essentialists (but also, for instance, Armstrong 1983) object to this: laws must be *explanatory*, and what they explain is precisely regularities. If the laws themselves are nothing more than regularities, then they cannot also explain regularities; nothing can explain itself. The realist's own solution is to provide some metaphysical grounding for the regularities that qualify as laws, some 'lawmakers'. It is the lawmakers that make the law-qualifying regularities special; and it is the lawmakers that fulfil the explanatory job expected of a law. The lawmakers according to DE are the dispositional essences of fundamental properties. They make some regularities special: a regularity is a law just in case it is 'grounded' in the dispositional essences of fundamental properties. They also explain the regularities in question, simply and elegantly, by entailing them. Here is how the explanation works.

According to DE, the fundamental properties that figure in the laws have dispositional essences, characterized by a stimulus and a manifestation condition. Where P is any fundamental property, DE says that P is essentially, and hence necessarily, a disposition to yield a particular manifestation M in response to a particular stimulus S. Applying the third step towards DE (and ignoring for the moment the *ceteris paribus* clause), this gives us

(I) $\Box(Px \rightarrow (Sx \mathrel{\Box\!\!\rightarrow} Mx))$,

which in a few simple steps of first-order modal logic with modus ponens for the counterfactual – assume $Px \wedge Sx$, derive Mx, discharge the assumption – leads to the statement of a nomic generalization:

(V) $\forall x((Px \wedge Sx) \rightarrow Mx)$.

(See Bird 2007, 46.) (V) is a statement of a regularity; DE explains this regularity in terms of its derivation from (I). (V) is also a statement of a law, and we now have a simple way of distinguishing a regularity that is a law from a regularity that is not a law: the former, but not the latter, can be derived from a (true) characterization of a dispositional essence, of the form of (I). The 'grounding' of laws in essentially dispositional properties has now a rather precise meaning, and coincides with their explanatory function: both can be cashed out in terms of logical entailment.

Presented in this abstract manner, the account seems to work out just fine. But let us fill in the schema and look at an example. One of the dispositional essentialists' favourite examples for a fundamental property that has a dispositional essence is charge. The law that characterizes charge, or rather: that should be grounded in charge, is Coulomb's Law. Coulomb's Law states the relation between any given determinate charge Q and any other charge q_i, the distance r_i between Q and q_i, and the attractive or repulsive force F_i that is exerted:

(CL) $F_i = \epsilon \frac{Q q_i}{r_i^2}$

Clearly, (CL) does not look anything like (V). A first and minor point of dissimilarity is that (V) is, and (CL) is not, stated in the form of a conditional. What is more important is that (CL) states a *function*, and the variables in it range over *quantities*. Charge, force, and distance are quantities: they are determinable properties that come with an ordered range of determinates. An object is not merely charged (positively or negatively), it has a particular determinate charge, say, charge e or charge $2e$. The same holds of exerting a force and being at a distance from something. Coulomb's Law, accordingly, states not merely *that* an object with a (positive or negative) charge will manifest a certain kind of force (attractive or repulsive) in response to a certain kind of stimulus condition (say, another charge at some distance from it); it states exactly *how* *much* force the object will exert in response to exactly *how much* charge at *how great* a distance.

(I) and (V), on the other hand, appear rather to be designed for qualitative properties; they can tell us merely that *if* such-and-such properties are instantiated, *then* such-and-such other properties will be instantiated too. But as Bigelow and Pargetter (1988) note, with quantities 'the simple "on" or "off" of being instantiated or not being instantiated seems to leave something out' (Bigelow and Pargetter 1988, 287). Can the quantitative nature of charge and Coulomb's Law be integrated into the derivation of (V) from (I)?

To see how this might be done, we need to fix the minor dissimilarity and formulate Coulomb's Law in the form of a conditional. Now, (CL) is a function with several variables, and we do not want any free variables in a (V)-like statement of a law. There are two things we can do with the free variables: we can fill in determinate values for them, or we can quantify over them. For simplicity's sake, let us focus on a particular determinate charge, say, electric charge (charge e). Then the first strategy yields an infinity of rather specific statements, one of which is

(V-1) $\forall$x ((x has charge $e \wedge x$ is 5.3×10^{-11}m from a charge of 1.6×10^{-19}C) $\rightarrow x$ exerts a force of 8×10^{-8}N).

The second strategy, on the other hand, yields only one, multiply quantified, statement:

(V-$\forall$) $\forall x \forall r_i \forall q_i$ ((x has charge $e \wedge x$ is at a distance of r_i from a charge q_i) $\rightarrow x$ exerts a force of $F_i = \epsilon \frac{e q_i}{r_i^2}$).

Note that (V-∀) is closer to Coulomb's Law, but (because of its multiple quantification) is not quite of the same form as (V). (V-1), on the other hand, is an instance of (V). Accordingly, (V-1) can be derived from an instance of (I):

(I-1) $\Box$ (x has charge $e \rightarrow$ (x is 5.3 $\times 10^{-11}$m from a charge of 1.6 $\times 10^{-19}$C $\Box\!\!\rightarrow$ x exerts a force of 8 $\times 10^{-8}$N)).

(V-∀), on the other hand, can be derived from the following:

(I-∀) $\Box$ (x has charge $e \rightarrow \forall$ charges q_i $\forall$ distances r_i (x is at r_i from q_i $\Box\!\!\rightarrow$ x exerts force $F_i = \epsilon \frac{e q_i}{r_i^2}$))

Again, (I-1), while an instance of (I), is not a characterization of charge e; and (I-∀), while being a better candidate for characterizing the dispositional essence of electric charge, is not quite of the same form as (I). There is a tension here: we can either keep strictly to the form of Bird's derivation, or we can capture the law (Coulomb's Law) and the property (electric charge) that we intuitively want to capture; but we cannot do both. I will argue that this tension is no mere appearance: there is a real conflict here, and it cannot be resolved without giving up some part of DE.[2]

The question that I put to the dispositional essentialist, then, is this: which is the fundamental property, and which is the law to be derived: (I-1) and (V-1), or (I-∀) and (V-∀)?

2.5.2 Multi-Track Dispositions

Bird's answer to this question is hidden in an early and somewhat peripheral section on multi-track dispositions (Bird 2007, 21-24), and is curiously forgotten in the rest of the book. In the section on multi-track dispositions, Bird argues for the priority of

[2] Alice Drewery has raised worries very similar to the ones I am going to discuss in a talk given at the Metaphysics of Science conference in Nottingham, September 2009.

(I-1) and (V-1) over (I-∀) and (V-∀). Here is what he says (slightly rephrased to fit my set-up).

(I-∀) is equivalent to a conjunction of (infinitely many) statements such as (I-1). Given the nature of dispositionality as outlined in the third step towards DE, it would seem that (I-1) characterizes *some* disposition. The question then is, which of them is more fundamental, (I-1) and its cognates, or (I-∀)?

Bird defines a 'pure disposition' as one 'which can, in principle, be characterized ... as a relation between a stimulus and a manifestation' (Bird 2007, 22). (I-1) gives the characterization of a pure disposition; (I-∀), as we have seen, is equivalent to a conjunction of pure dispositions. And Bird continues:

> It is my view that all impure dispositions are non-fundamental. Fundamental properties cannot be impure dispositions, since such dispositions are really conjunctions of pure dispositions, in which case it would be the conjuncts that are closer to being fundamental. (Bird 2007, 22)

The fundamental properties are those characterized by (I-1) and its many cognates, and not properties such as charge *e*, which is better characterized by (I-∀). It turns out not only that charge, or even charge e, is not, after all, one of the fundamental properties. It also turns out that Coulomb's Law, or even (V-∀), is not, after all, a law that can be derived from the dispositional essences of the fundamental properties. What can be so derived are the infinitely many law-like statements such as (V-1).

What, then, is the status of Coulomb's Law, or even the instance of it expressed in (V-∀)?

(V-∀) certainly states a regularity; but does it, on the present account, state a law? We have seen that a regularity will qualify as a law just in case it is entailed by the dispositional essence of a fundamental property, as expressed in (I). But there is no such dispositional essence to entail (V-∀); (I-∀), which does entail it, does not state the dispositional essence of any fundamental property. So (V-∀) does not seem to state a law. Similarly, there is no dispositional essence to entail Coulomb's Law in its full generality; so Coulomb's Law is not a law.

The same point can be made the other way around: assuming that Coulomb's Law, or at least (V-∀), is a law, what is this law grounded in? By the explication of grounding, it would have to be electric charge, i.e. the disposition expressed in (I-∀); but that disposition has questionable ontological standing. If the property expressed by (I-∀) is 'really' just a conjunction, as Bird tells us, then it is 'really' nothing but its conjuncts. It is, metaphysically speaking, nothing over and above the conjuncts. But if it is nothing over and above the conjuncts, then it cannot provide the ontological grounding for anything over and above what its conjuncts provide the grounding for.

If regularities are to be explained by the laws (or the lawmakers), then dispositional essentialism has not delivered the explanation it had promised. Its laws, (V-1) and its cognates, can explain *some* regularities – the regularity, for instance, that a given object exerts a force of 8 $\times 10^{-8}$N whenever it is 5.3 $\times 10^{-11}$m from a charge of 1.6 $\times 10^{-19}$C. But the crucial regularity is the similarity *between* these specific regularities: the fact that they all exhibit the same mathematical correlations between stimulus and manifestation.

How is that regularity to be explained? It is clear that the 'impure disposition' (I-∀) will have to play a role. And indeed Bird says that '[w]hile it is possible to gerrymander impure dispositions of all sorts, it is clear as regards the cases we are interested in, [such as] charge [...], that the conjunctions are natural or non-accidental.' (Bird 2007, 22) That is clear indeed; but it is far less clear how the dispositional essentialist is to account for the naturalness of these conjunctions. It will not do to say that (V-∀) is grounded in a non-fundamental disposition captured by (I-∀). For as we have seen, that non-fundamental disposition could not be, and hence could not ground, anything over and above the fundamental conjuncts of which it is made up.

If a single determinate charge such as charge e and the instance of Coulomb's Law that is expressed in (V-∀) cannot be captured by DE, then even less will the determinable property charge and Coulomb's Law be captured. So I conclude that, with Bird's preference for (I-1) over (I-∀), DE fails to accomplish its explanatory task when it comes to the property of charge and Coulomb's Law. To see just how damaging this conclusion is, we need to consider how far the argument I have given generalizes beyond the one property and the one law that I have been considering.

How far, then, does the argument generalize? The answer is: very widely. It may be, in fact it is rather likely, that charge and Coulomb's Law are not fundamental. But my argument did not turn on any specific features of charge, or of Coulomb's Law. It turned only on charge being a quantity, and Coulomb's Law being a functional law that states mathematical correlations between quantities. Every candidate fundamental property that participates in a functional law will be subject to the same line of argument. Bird conjectures, I believe for the reason stated, that the really fundamental laws will not have any constants, such as ϵ in Coulomb's Law (Bird 2009). But not only is that an undesirably strong prediction; it will not be enough to avoid the kind of argument I have given. The argument applies to properties and laws that involve *any* mathematical operation – multiplication or division is enough. And Bird should certainly not predict that there will be no mathematical operations whatsoever in the fundamental laws. That would not only be a daringly strong prediction. It would make it utterly mysterious what *could* count as a fundamental law at all.

We can see now that Bird's argument for the fundamentality of (I-1) rather than (I-∀) has very damaging consequences for DE. The natural solution is to give up the multi-track view and adopt (I-∀) as fundamental instead. That, of course, means that we must reject Bird's argument for the priority of 'pure' dispositions. Let us see how that can be done.

2.5.3 Beyond Multi-Track Dispositions

Bird said that dispositional properties such as the one expressed in (I-∀), or electric charge, are 'really conjunctions of pure dispositions' (Bird 2007, 22). Now, what is the status of that 'really'? What makes the pure dispositions any more fundamental than the 'impure' ones such as that characterized by (I-∀)?

The answer lies in what I have called the third step towards DE: the characterization of what it is for a property to be dispositional. Dispositionality, we have seen, is understood in DE as connecting a stimulus property and a manifestation property, the connection amounting to something like a counterfactual conditional (with a *ceteris paribus* clause). The 'pure' dispositions are pure simply in that they perfectly conform to that characterization.

If we look at (I-∀), however, the characterization of dispositionality fails: we cannot take it apart in the way we can take apart (I-1), factoring it into two separate properties. With (I-1), the stimulus condition is: being $5.3 \times 10^{-11} m$ from a charge of $1.6 \times 10^{-19} C$; the manifestation condition is: exerting a force of $8 \times 10^{-8} N$. In (I-∀), we have quantifiers ranging over q_i and r_i in both the stimulus and the manifestation condition. If we try to separate them and specify one in separation from the other, we lose the very correlation which (I-∀) has been formulated to capture. But not only can the properties not be separated from each other; the counterfactual conditional is not doing any work.

What (I-∀) says is that everything with charge *e* will always exert a force that stands in a certain mathematical correlation to whatever other charges are present and their distance from it.

We can see now why Bird went for his multi-track view.[3] It is the third step towards DE, the characterization of dispositionality, that forced it on him. If we think that a disposition has to come with separable stimulus and manifestation conditions, related in a counterfactual-like way, then the best (I-∀) can hope for is to count as a conjunction of such dispositions. If (I-∀) is to stand on its own, a new conception of dispositionality will be needed: to begin with, dispositionality will look more like a one-place operator, but one that takes complex functions as its complements. Dispositions will have to be construed not as dispositions to ... if ..., but as dispositions to ..., fullstop – in other words, the standard model will have to be discarded, and my alternative model adopted.

Exactly how dispositions such as electric charge fit into the schema I have described earlier is a question that I have to leave open for further research. To begin with, an account will be needed of how electric charge, the dispositional property characterized by (I-∀), is related to such specific sub-dispositions as (I-1) if it is to be more than just a conjunction of them. Note that, while the apparatus of determinables and determinates may enter in some way or other, it cannot be used to characterize the relation between electric charge and the many sub-dispositions. For the determinates of one determinable exclude each other: while having charge entails having some determinate charge, having any one determinate charge entails not having any other determinate charge. In the case at hand, however, having electric charge entails not merely having *some*, but indeed having *all* the many parallel sub-dispositions.

[3]Note that Ellis (2001) does not seem to subscribe to the multi-track view: he ascribes to dispositional properties or 'causal powers' the 'required unifying power ... to explain an infinite range of quantitatively distinct dispositions' (Ellis 2001, 132). However, as far as I see, Ellis does not tell us *how* causal powers unify these distinct dispositions. When he explains how the laws are grounded in dispositional properties (Ellis 2001, 130), he too uses a conditional with determinate values.

I do not claim to know how these questions should be answered. But they are questions to be asked within the scope of the alternative model. My closer look at DE has shown that DE has no more reason to adhere to the standard model than does the philosopher who is trying to understand the dispositions, such as fragility, that we ascribe with the dispositional terms of ordinary language.

I believe that this result generalizes beyond DE, though I can do no more than sketch that generalization here. The objection that I imagined to be brought against my alternative model was this: the dispositional properties that play a role in scientific thought should be understood so as to 'encode' the laws that correspond to them. (Whether those dispositions are identical to the natural properties, such as charge or mass, that figure in the laws, does not matter for this point; nor does it matter whether those dispositions are the ontological grounds for, or whether they themselves are derivative from, the laws.) If this 'encoding' is best understood by using the standard model, then that model should not be dismissed.

I agree with the conditional but reject the antecedent. The laws to be encoded are laws such as Coulomb's Law, and the standard model does not allow for an 'encoding' of these laws that retains their most important feature, the quantitative correlations. To encode a law such as Coulomb's Law, a disposition must be of the kind that is expressed in (I-∀), not of the kind that is expressed by (I-1); and that means that it must be of the kind that conforms to the alternative model, not the standard model. This point, though I have argued for it with DE in mind, is perfectly general.

Of course, dispositions that are in accordance with the standard model can be stipulated here as elsewhere (recall the multi-track strategy in section 2.3.1), but we have no more reason to take them to be the 'real' dispositions than we had reason to do the same with the specific fragility-dispositions – not, that is, unless we have already subscribed to the standard model. Science does not force them onto us; science is better understood by adopting a conception of dispositions in line with the alternative model.

Chapter 3

Towards an Account of Potentiality

3.1 Introduction

In chapter 2.4.1, I distinguished two versions of the view I proposed, a reductionist and a non-reductionist version. On the reductionist version, dispositionality is just a species of possibility: x has a disposition to F just in case x F's in some relevant world, with some special condition for relevance. On the non-reductionist version, possibility of any kind provides no more than a rough model and an illuminating parallel for dispositionality. My argument of the previous chapter remained neutral between these two versions. It is now time to take a stand: I want to provide a non-reductionist account of dispositions and other related properties, which I call 'potentialities'. In this chapter, I provide some preliminary motivation for such an account, by following two different threads: the link between potentiality and essence, and the link between potentiality and the modal auxiliary 'can'. First, let me make a few brief introductory remarks on 'potentiality'.

Dispositions, I have said earlier, are modal properties: they do not, or not primarily, concern what objects are *actually* doing, but rather what they *can* do. A fragile object is one that can break or be broken easily; a breakable object is one that can be broken. Adjectives ending with '-able'/'-ible', I have suggested, appear to take passive complements: an F-able object is one that can be F'ed. With active verbs, 'can' is often used to express a different kind of modal property, abilities: I can play the piano, you can speak English, etc. Like dispositions, abilities are properties of individual objects; and like dispositions, they are possibility-like properties. We can extend the list of such possibility-like properties: there are potentials, tendencies, propensities, inclinations, capacities, and capabilities. In order to have a somewhat technical term that covers all such properties, I call them 'potentialities'. All potentialities are properties of individual objects that concern what the objects *can* do. In section 3.3, I will suggest that the link between potentiality and 'can' is an even more intimate one: that 'can' in general is used to ascribe potentiality, not to express possibility. Before I do so, I will look at another link: that between potentiality and essence.

Potentiality is the 'local' analogue of possibility; potentialities are, as it were, possibilities rooted in individual objects. Potentiality stands to possibility as essence stands to necessity: both are 'local' versions of a 'global' modality. Potentialities and essences are potentialities and essences *of* particular objects: x has a potentiality to do so-and-so, and y is essentially such-and-such. Possibility and necessity are possibilities *that* and necessities *that*; there is no question *whose* possibility or necessity is at issue. In section 3.2, I will look at Kit Fine's arguments concerning the relation between essence and necessity, and provide a parallel argument for the case of potentiality and possibility. According to Fine, essence cannot be reduced to necessity, but is in fact prior to it: necessity is constituted by the essences of objects. I will suggest that for similar reasons, potentiality plausibly cannot be reduced to possibility, and that perhaps we should think of possibility as constituted by the potentialities of individual objects. The second part of the suggestion must remain tentative, at least as long as no general and non-reductive account of potentiality has been given. I will proceed to give such an account in chapter 4, and then apply it to an account of possibility in chapter 5. The two threads of this chapter serve to motivate these two accounts.

3.2 Essence, Potentiality, and Possibility

3.2.1 Fine on Essence and Necessity

Kit Fine has famously argued that the notion of a thing's essence cannot be reduced to, or derived from, that of metaphysical necessity (cf. especially Fine 1994, but also Fine 1995b and Fine 1995c). Using Fine's notation $\Box_x p$ for 'it is true in virtue of the essence of x that p', the *modal account* of essence says that the equivalence

$$\Box_x Fx \equiv \Box Fx$$

not only holds, but also characterizes what it is for any object x to be essentially F. (The equivalence may be complicated considerably, most obviously by conditionalizing the right-hand side on x's existence; but for present purposes it is sufficient to employ the simplest version.)

Fine argues that the modal account fails because the equivalence does not hold: it has false instances when read from right to left. Thus it is true that Socrates is necessarily a member of {Socrates}, and that Socrates is necessarily such that $2+2 = 4$;

but neither of these properties is essential to Socrates. The modal account fails to distinguish necessary from essential properties.

Fine gives a general diagnosis of this result:

> Given the insensitivity of the concept of necessity to variations in source, it is hardly surprising that it is incapable of capturing a concept which is sensitive to such variation. Each object, or selection of objects makes its own contribution to the totality of necessary truths; and one can hardly expect to determine from the totality itself what the different contributions are. ... Indeed, it seems to me that far from viewing essence as a special case of metaphysical necessity, we should view metaphysical necessity as a case of essence. (Fine 1994, 9)

In this passage, Fine is claiming more than merely that essence does not reduce to modality. He is suggesting further that, on the contrary, necessity reduces to (or is constituted by, based on, or what have you) essence. Let me label these two claims for further reference:

First claim: Essence does not reduce to necessity. (Necessity is not prior to essence.)

Second claim: Necessity is constituted by essence. (Essence is prior to necessity.)

In the passage I have quoted, the second claim is offered as an *explanation* of the first. If the essences of individual objects each contribute their share to the totality of necessity, we should not expect to be able to trace back the individual contributions when given merely the totality (any more than you can determine the lines of each particular singer by listening to a piece of complicated choral music). If the second claim can be given a satisfactory and non-metaphorical formulation, then the first claim will provide evidence for its truth.

How, then, can the second claim be spelled out? How is necessity to be understood as 'a case of essence'? Fine has an answer to that question. First, we need to adjust the essence operator: instead of names, it will now take predicates as a subscript: $\Box_F p$ says that it is true in virtue of the nature of (all) objects which are F that p. (The singular case is included as a special case: $\Box_{\lambda x.x=a} p$ says that p is true in virtue of the nature of (the objects identical to) a.) The propositions p such that $\Box_F p$ are the F-necessities:

> [E]ach class of objects ... will give rise to its own domain of necessary truths, the truths which flow from the nature of the objects in question. The metaphysically necessary truths can then be identified with the propositions which are true in virtue of the nature of all objects whatsoever. (Fine 1994, 9)

(The idea of 'domains' or 'spheres' of necessary truths is spelled out in more detail in Fine 2002.) So we get the following equivalence:

$$\Box p \equiv \Box_{\lambda x.x=x} p$$

Unlike the equivalence proposed by the modal account of essence, this one is true; however, it can serve as a definition not of essence but of (metaphysical) necessity. Note that in explicating the right-hand side of the equivalence, we use a universal quantifier: the necessities are those propositions that are true in virtue of the nature of *all* objects. The structural connections between the universal quantifier and necessity have been studied and employed by accounts of necessity that appeal to possible worlds; there is no reason why, in a different way, it should not also be used by the essentialist account.

So much for a quick summary of Fine's essentialist views on essence and necessity. Potentiality, I have said, stands to possibility as essence stands to necessity. It will be interesting to see, therefore, whether an analogue of Fine's two claims can be developed for the case of potentiality and possibility. The next section will suggest that it can.

3.2.2 Potentiality and Possibility

Fine's second claim, that essence is prior to necessity, was motivated as an explanation of the first claim, that essence is irreducible to necessity. Let me therefore begin by considering such an argument for a potentiality-analogue of the first claim: potentiality is irreducible to possibility.

In the debate about dispositions, arguments against a reduction of dispositions (to the counterfactual conditional) often appeal to finks, masks, and antidotes, mechanisms which prevent the manifestation of a disposition even if its stimulus is present, as well as reverse finks and mimics – mechanisms which elicit the manifestation given the stimulus in the absense of the disposition (see, for instance, Martin 1994, Bird 1998, Molnar 2003, Everett 2009). In this section, I will give a similar argument, but I must adapt it in two ways. First, I am not concerned with reducibility to counterfactuals; I have argued in the previous chapter that the modality we should be thinking about is possibility. Secondly, I am not concerned with the reducibility of dispositions only, but rather with properties of a wider category, potentialities. Potentialities include more than dispositions, they include abilities, and for all I have said, they might include even more. (I will argue later that they do.) So it is difficult to prejudge the extension of the concept of potentiality, and accordingly difficult to provide (relatively uncontroversial) counterexamples of the form 'it is possible that x is F, but x has no potentiality to be F'. One such counterexample should be safe to rely on even without a firm grasp on the extension of 'potentiality': nothing has a potentiality not to exist, hence I do not have a potentiality not to exist; but it is surely possible that I did not exist. But rather than make much of this one counterexample, I will concentrate on a more pervasive phenomenon: potentialities have degrees, and there appears to be no way to capture these degrees correctly in a reductive account of potentiality in terms of possibility.

Potentialities come in degrees: some glasses are more fragile than others, and some of us possess the ability to play the piano to a much greater degree than others. Moreover, those degrees are contingent: a given glass might have taken a little crack and thus been more fragile than it is; had I practised the piano more when I was a child, I would now possess the ability to play it to a much greater degree.

Incidentally, this shows that potentiality is more than just the dual of essence. We might have thought that having a potentiality to F was equivalent to not being essentially not-F, just as 'possibly *p*' is equivalent to 'not necessarily not *p*'. Whether that equivalence holds, I cannot yet determine since I have not said enough about the extension of the term 'potentiality'; it is certainly not the case that an object has an ability or disposition to F whenever F-ing is not precluded by its essence. But even if the equivalence were to hold (I argue below in chapter 4.8 that it does not), it would fail to capture an important aspect of potentiality: its degrees. With potentiality and possibility, things are a little more complicated.

Treating potentiality along with possibility, and understanding possibility in terms of possible worlds, we would have to say: an object x has a potentiality to F just in case x F's in some possible world. That approach would have to be extended to accommodate degrees of potentialities; and of course it can be so extended. In particular, possible worlds talk allows us to introduce degrees of possibility, thus promising to satisfy the constraint that we must capture the gradability of potentiality. The degree to which an object x has a potentiality to F will then simply be the degree of the possibility that x F's.

There are two ways to understand degrees of possibility on a possible-worlds framework: one is in terms of closeness to the actual world, the other is in terms of a probability-like measure or proportion of possible worlds. Let me look at these two options in turn.

First, closeness. On this conception, to what degree it is possible, at *w*, that *p* depends on how close to *w* the closest *p*-world is. Closeness of worlds is determined by contextually determined similarity relations between worlds. (Classic statements: Lewis 1973, Lewis 1986a.) The degree of a possibility, on this account, is contingent like that of a potentiality: which worlds are close differs from world to world; and as time progresses, worlds that were initially close may diverge and others that were initially farther away may converge. So far, all is well. But while closeness does deliver contingency, it delivers the *wrong contingency*.

The problem is known as 'accidental closeness' (Manley and Wasserman 2008). On the present proposal, the degree of a glass's fragility just is the degree of the possibility that the glass breaks, which in turn depends on how close to the actual world the closest world is in which the glass does break (times being held constant). Now take two glasses that are equally fragile, store one of them safely on a shelf above a soft carpet, and place the other one at the edge of a table on a stone floor. On any reasonable closeness relation, there is a close world where the second glass breaks – one careless movement is enough – but no equally close world where the first glass breaks. But by hypothesis the two glasses were equally fragile. Hence the closeness account gave us the wrong degree for at least one of them (or more likely, for both: too low for the first, too high for the second).

Second, proportion. On this account, to what degree it is possible (at *w*) that *p* is determined by the measure of worlds at which *p*, or the proportion of *p*-worlds to non-*p*-worlds. This is not subject to problems of accidental closeness: any possible circumstance, stone floors as well as soft carpets, will be among the worlds measured, no matter how close to actuality. However, we can see immediately from the characterization above that degrees of possibility on this account are not contingent: there is no relativization to *w* in the right-hand side (and no space for such relativization). If degrees of potentiality were identical to degrees of possibility on this conception, then worlds where this glass had taken an additional crack would count towards determining its degree of fragility, and worlds where I had violin lessons (which in fact I have not had) would count towards the degree of my ability to play the violin.

If closeness yields the wrong contingency, and a measure/proportion account yields none at all, how about combining the two? Let us say that the degree of the possibility that *p* is determined by the measure of reasonably close worlds where *p*. Then closeness may take care of the contingency of degrees of potentialities, while the proportion rids us of the problem of accidental closeness. Among the close worlds, there should be those with stone floors and those with soft carpets, and in general enough variation of external circumstances to preclude accidental factors such as the glass's actual storage from mattering. Not everything must be varied, however. We want a measure of breaking-worlds out of all the worlds where the glass has taken all and only the cracks that it has actually taken, and of violin-playing-worlds out of all the worlds where I have had exactly the amount of violin lessons I have had (that is, none). In general, we want to measure the manifestation-worlds out of the worlds where the potentiality-possessing object is intrinsically just as it is in the actual world (including, if you like, its history). This is the closeness relation that should be at work: closeness is determined by similarity to the actual world with respect to the intrinsic constitution of the object whose potentialities are at issue. The degree of *x*'s potential to F is determined by the proportion of worlds, out of those where x is intrinsically just like it is in the actual world, where x F's. (This proposal is close to what Manley and Wasserman 2008 suggest for the case of dispositions, cf. ch. 2.3.2.)

Yet this is still not fine-grained enough. For the requirement of intrinsic similarity is both too strong and too weak.

It is too strong because there are extrinsic potentialities. To determine the degree of my vulnerability you have to consider exactly those worlds where not only am I intrinsically just as I am actually, but where certain features of my surroundings are too. Which features? Well: those which count towards my vulnerability. In general, the relevant worlds have to include those and only those where the potentiality-possessing object *and* the relevant features of its environment are just like they are in the actual world. Which are the relevant features of the environment? They are those, if any, that possession of the potentiality in question depends on.

Moreover, and more importantly, the requirement is too weak. Suppose I have had piano lessons, but have made a firm decision never to play the piano. That decision is part of my intrinsic make-up, just as much as my ability to play the piano. So considering only worlds where I am intrinsically just as I am in actuality (in the supposed scenario) will yield very few if any worlds where I do play the piano, thus misrepresenting the degree of my ability. This is not a far-fetched kind of case: anyone who is irascible but well-bred, disposed to cry but disciplined, or prone to overeat but strictly following a diet will serve as an example. The suggested treatment would have us say that by adhering to a diet, one loses the tendency to overeat – a claim which goes strongly against the phenomenology of dieting. In general, that the potentialities I have mentioned are not lost but only masked in such cases can be seen from the effort that it takes to avoid or 'mask' their manifestation.[1] To capture all and only the relevant worlds for an object's potential to F, we have to include worlds with some variations in the object's intrinsic make-up. But only *some* variation: in fact, variation in anything but the potentiality itself. Again, the specification of the close worlds cannot help but refer to the very potentiality that is to be reduced – and that cannot constitute a reduction. (This line of argument will be pursued in more detail, though from a different angle, in section 3.3.)

My argument has been sketchy, but it suggests that an analogue to Fine's first claim can be maintained for potentiality: potentiality plausibly does not reduce to possibility. We thus have some motivation to consider an analogue of Fine's second claim: could it be that possibility is in fact 'constituted' in some sense by potentiality? If so, how can that claim be spelled out?

Fine's characterization of necessity in terms of essence made use of the structural similarity, studied and used in possible-worlds semantics, between necessity and the universal quantifier; we might similarly make use of the structural similarity between possibility and the existential quantifier. One thing's potentiality is enough to create a possibility; or less metaphorically: it is possible that p just in case something has a potentiality for p. If more than one thing has the relevant potentiality, then the degree of the possibility will be determined somehow by all of them.

We could then paraphrase Fine's comments and explain the first claim, that potentiality is not reducible to possibility, by the second claim: if the potentialities of individual objects each contribute their share and degree to the totality of possibilities and their degrees, we should not expect to be able to trace back the individual contri-

[1] It will be obvious to everyone familiar with the literature on dispositions that I am appealing to intrinsic masks or antidotes here. I do not say they exist at the fundamental level, but it seems obvious that they exist in complex beings such as ourselves. Recent arguments for the existence of intrinsic finks, masks, etc. can be found in Clarke 2008, Everett 2009 and Ashwell 2010.

butions when given merely the totality (any more than you can determine the lines of each particular singer by listening to a piece of complicated choral music).

However, such an account of possibility would face many immediate questions. First, what is a 'potentiality *for p*'? Have we not understood potentialities so far as potentialities to ... ? A second, and more pressing, question is one that I have noted earlier: what is the extension of 'potentiality'? If dispositions and abilities are the only potentialities there are, they will not suffice to deliver all the possibilities that we intuitively think there are. (It is possible that my granddaughter becomes President of the United States. Who should have a disposition or an ability to that effect?) Both questions point us in the same direction: we need a general account of potentiality. And if we have been convinced by the argument of this section, we will want that account to be a non-reductive one; we should understand potentiality on its own terms, not through the apparatus of possible worlds. I will provide such an account of potentiality in chapter 4, and return to the second claim, a potentiality-based account of possibility, in chapter 5. But first, let me look at another important aspect of potentiality: its link with the modal auxiliary 'can'. It will provide us with further motivation for a general account of potentiality.

3.3 The Semantics of 'Can'

3.3.1 'Can' and Potentiality

In chapter 2, I have linked dispositions closely to the expression 'can': a breakable thing is one that can be broken, a fragile think is one that can easily be broken. The link with 'can' holds even more obviously for abilities: the most natural way to ascribe to oneself an ability to play the piano is to say 'I can play the piano'. I have so far treated 'can' as philosophers (and philosophically inclined linguists) standardly do: namely, as expressing possibility. On that standard construal, 'x can F' is true iff it is possible that x F's, or equivalently, iff there is some (relevant) possible world in which x F's.

Given the surface grammar of 'can', however, that standard treatment is somewhat surprising. 'Can' is a predicate modifier: it takes a verbal expression (e.g. 'play the piano') to form a new, more complex verbal expression ('can play the piano'). As a predicate modifier, 'can' contrasts with sentence-modifying adverbs such as 'possibly'. A sentence-modifying adverb can be prefixed to a whole declarative sentence to yield another declarative sentence: 'possibly, I play the piano' is a declarative sentence; 'can I play the piano' is not.

If we go by surface grammar, then, the form of 'x can F' is not 'Can(Fx)', but rather '(Can(F))(x)'. In 'I can play the piano', the complex predicate 'can play the piano' is predicated of me. If the semantics of 'can' follows surface grammar, it should construe the sentence as ascribing a property to me that is a function of the semantic values of 'can' and 'play the piano'. That property, unsurprisingly, is most naturally taken to be a potentiality: 'The vase can break' ascribes to the vase a potentiality, the disposition of being breakable; 'I can play the piano' ascribes to the speaker a potentiality, the ability to play the piano. Or so the surface grammar suggests.

The standard semantics of 'can' has it that surface grammar is misleading: semantically (so the story goes), 'can' behaves like the sentence-modifying adverb 'possibly'. Deep down, on the level of logical form, it *is* a sentence modifier. Or so it is thought.

To be sure, surface grammar may be misleading; there may be good reason to go against it in semantics. But the default option should be to explore a semantics that is in line with surface grammar; the burden of proof is on those who are contravening it.

The standard semantics is so deeply ingrained in philosophical discourse that this burden is rarely felt. However, it is relatively clear why it is that the standard semantics ignores surface grammar: the treatment of 'can' as, deep down, a sentence modifier goes best with a possible-worlds semantics. The possible-worlds semantics in turn comes with a range of advantages that may not be shared by a potentiality-based semantics. Here is a list of them.

(i) With possible-worlds semantics, we can provide a unified semantics of modal expressions of all categories, syntactic (verbs, adverbs, etc.) and semantic (circumstantial, deontic, epistemic).

(ii) Possible-worlds semantics provides the most elegant account of the context-dependence of 'can'.

(iii) According to Humean supervenience, potentialities are not part of the basic furniture of the world; any expressions (such as 'can') which seem to ascribe them should be analysed in terms of some Humeanly acceptable resources. Possible-worlds semantics does just that.

(iv) Possible-worlds semantics is 'wide' enough to account for all sentences containing 'can'. It is not clear that potentialities can do the same: which potentiality is ascribed, for instance, in an utterance of 'Children can be hit by a car when playing on this street'?

I will address these considerations in the order in which I have listed them, though in a somewhat scattered way. (i) will be addressed in this section, (ii) and (iii) together are the subject of sections 3.3.2 to 3.3.5. (iv) can be addressed only by providing an account of potentiality that is wide enough to accommodate all uses of 'can', a task which is to be accomplished in the next chapter. I will therefore briefly return to the issue of a semantics of 'can' at the end of that chapter.

First, let me address (i).

Modality is standardly partitioned into epistemic, deontic, and circumstantial modality. Roughly, epistemic modality is about what is compatible (or not) with our knowledge, deontic modality is about permission and obligation, and circumstantial modality is about developments that are open (or not) given how things really are. In linguistic treatments, circumstantial and deontic modality are often grouped together as 'root' or 'agent-oriented' modality, but I will not consider deontic modality in this section. While being mostly associated with circumstantial modality, 'can' also has some deontic and epistemic uses. It is sometimes used, like 'may', to express permission, as in 'You can go now.' In its negated form, it is regularly used epistemically: 'This cannot be true!' (For more on linguistic treatments of the modalities, see, e.g., Coates 1983, Kratzer 1991 and the articles in Frawley 2006.)

The uses of 'can' in which I am interested are only the circumstantial ones. I am going to argue that there is an important contrast between the circumstantial 'can' on the one hand and epistemic modals on the other which suggests that the latter, but not the former, are 'deep down' sentence modifiers. The contrast thereby tells against an overly unified approach to the semantics of circumstantial and epistemic modals that assimilates the former to the latter by treating both as sentence modifiers.

Thomason and Stalnaker (1973) have proposed a test to establish whether a given adverb, 'Q-ly', is a sentence modifier or not:

> *Criterion 4*
>
> Only if *Q*-ly occurs as a sentence modifier can one paraphrase the sentence

> by deleting the adverb and prefacing the resulting sentence by *It is Q-ly true that.* (Thomason and Stalnaker 1973, 205)

With adverbs, the criterion seems to work well (try 'unfortunately' as an example of a sentence modifier, and 'slowly' as an example of a predicate modifier). Before applying it to the case I am here interested in, we should note two caveats. First, the Thomason/Stalnaker test is proposed as a criterion for application to adverbs, not auxiliaries like 'can'. Second, it is formulated explicitly as a sufficient condition, not a necessary one, for being a sentence modifier, though Thomason and Stalnaker do suggest that it 'comes close to being a necessary and sufficient condition' (Thomason and Stalnaker 1973, 206). With this in mind, let us apply the test:

1.a She can play the piano.

1.b It can be true that she plays the piano.

2.a The child can be hit by a car.

2.b It can be true that the child is hit by a car.

Are 1.b and 2.b successful paraphrases of 1.a and 2.a, respectively? It seems to me that they are not. Neither 1.b nor 2.b sounds very natural in English. But on the (by far) most natural reading, 1.b expresses an *epistemic* possibility of her playing the piano, which would be more idiomatically expressed by 'It can be that she is playing the piano'. 2.b, which would sound more natural with a past tense '... that the child has been hit by a car', similarly expresses an epistemic possibility that the child is being (or has been) hit by a car. Note how naturally 1.b and 2.b are followed by '... but I'm not sure'; the same is not true of 1.a and 2.a.

In general, it seems that when we turn 'can' into a sentence modifier, then on the most natural reading, we get (if anything) epistemic instead of circumstantial possibility. The epistemic uses of ‘cannot’, on the other hand, do seem to pass the test for being a sentence modifier, as does the epistemic ‘might’. Consider

3.a John cannot be the murderer.

3.b It cannot be true that John is the murderer.

4.a John might be the murderer.

4.b It might be true that John is the murderer.

Here the paraphrase sounds successful, suggesting with epistemic ‘cannot’ and ‘might’ we do have good reason to contravene surface grammar and construe them as being sentence modifiers ‘deep down’. No such reason has emerged for the circumstantial use of ‘can’.

The phenomenon generalizes: circumstantial possibility tends to be expressed by predicate modifiers (the verb ‘be able to’, the suffixes ‘-ible’/‘-able’), epistemic possibility by sentence modifiers (‘perhaps’, ‘it is possible that’). This observation is not limited to English. Witness Angelika Kratzer about modality in German:

> [t]he distinction [between epistemic and circumstantial modality] is clearly marked in the vocabulary. Verbs with inherent modality, modal adjectives on *-lich* and *-bar* [‘-ible’/‘-able’] and phrases like *imstande sein* or *in der Lage sein* [‘to be able to’] never express epistemic modality.
> Sentence adverbs like *wahrscheinlich* or *möglicherweise* [‘perhaps’ and ‘possibly’] always express epistemic modality – if they express modality at all. (Kratzer 1981, 56)

In a similar vein, DeRose (1991) has noted that the sentence modifier ‘It is possible that...’ is used in English to express epistemic, not circumstantial, possibility. Circumstantial possibility is expressed by ‘It is possible for ... to ...’ (e.g., ‘It is possible for me to go’), which is, again, most naturally construed as a predicate-modifying expression.

If all this is correct (and I have provided no more than good but defeasible evidence), then a unified treatment of circumstantial and epistemic modality which takes all modal expressions as ‘deep down’ sentence modifiers may sacrifice, for the sake of formal elegance, important insights about their differences. In particular, it overlooks the divide between them which suggests that circumstantial modality *in general* is most naturally construed as ascribing properties to objects, while epistemic modality *in general* is most naturally construed as applying to whole propositions. That divide is best explained if we think of expressions of circumstantial modality as ascribing modal properties (and in particular, of ‘can’ as ascribing potentialities) to objects, while construing expressions of epistemic modality as expressing the epistemic standing of propositions. What is compatible with our knowledge, like knowledge itself, is propositional, so it is unsurprising that it (but not the modal properties ascribed by circumstantial modals) should be expressed by a sentence modifier.

Consideration (i), then, is weaker than it might at first look. For a unified treatment of circumstantial and epistemic modals as sentence modifiers sacrifices one feature that is distinctive of circumstantial modality as opposed to epistemic modality: its being expressed, quite generally, by predicate modifiers.

I now turn to (ii) and (iii) on my list. To address them, I first need to say more about the possible-worlds semantics of ‘can’. My argument will not refute these considerations. Rather, I will show that they are in tension with each other: adherence to Humean supervenience prevents us from giving the most informative and elegant account of the semantics of ‘can’ that possible worlds afford, for that account requires appeal to potentialities. I will concentrate in particular on one kind of potentialities, abilities.

3.3.2 The Kratzer/Lewis Semantics in Outline

Work on the formal semantics of modal expressions has been shaped by Angelika Kratzer's approach, as outlined in Kratzer (1977), Kratzer (1981) and Kratzer (1991). According to Kratzer, the semantics of any modal expression is determined along three axes: its modal force (roughly: possibility or necessity and various gradations in between), a modal base, and an ordering source. An expression's modal force is the only fixed component; the modal base and ordering source are determined by context (though some modal expressions also come with a restriction on which modal bases and ordering sources they admit).

A modal base is a function which assigns to every world *w* a set of worlds, the worlds accessible to *w* in the relevant context; it is only these worlds that the modal force is then applied to. An ordering source is a function from a world to an 'ideal' (a world which may or may not itself be accessible); the accessible worlds are then ranked by their closeness to that ideal, and those that are too remote from it may be excluded from consideration.

Modal bases come in two main kinds, epistemic and circumstantial. An epistemic modal base comprises the worlds in which all that we know (or perhaps, all that we do or could easily come to know, or so DeRose 1991 suggests) is true; a circumstantial modal base comprises the worlds in which certain contextually selected features of the world are as they actually are.

The English 'can' (like the German 'können'), according to Kratzer, expresses possibility and goes with any modal base and ordering source. A sentence '*x* can F' is true at a world w relative to a modal base $f(w)$ and an ordering source g iff x does F in some world $w' \in f(w)$ that is ranked high enough by g. (In the circumstantial cases I am concerned with, the ordering source will typically be a so-called 'stereotypical' one, which excludes worlds where abnormal things happen – sorcery, quantum blips, what have you – as too remote from the ideal. I will omit mention of ordering sources from now on as they play no part in my argument.)

In linguistics, the treatment of the modals that I have outlined so far is associated with Angelika Kratzer; in philosophy, with David Lewis. Lewis's formalism is slightly different from Kratzer's, but there is no difference in principle. Lewis (1979) sketches a way to deal with the 'relative modality' expressed by terms such as 'can'. Lewis compares the items that enter into the context in a conversation or other linguistic interaction with the score of a baseball game. Sentences depend on the score for their truth-value or acceptability, and other expressions for their intension or extension. In the case of modals like 'can',

> [t]he boundary between the relevant possibilities and the ignored ones (formally, the accessibility relation) is a component of the conversational score, which enters into the truth condition of sentences with "can" or "must" or other modal verbs. (Lewis 1979, 247)

The 'relevant' possibilities, or the accessible worlds, are none other than Kratzer's modal base. In another context, Lewis speaks of compossibility with a set of facts, or of holding fixed certain facts (Lewis 1976). This again amounts to the same: conversational score determines the facts to be held fixed, and thereby the relevant or accessible worlds: they are those worlds where the said facts are 'fixed', i.e. are as they actually are. '*x* can F' is true iff *x* F's at some accessible world, i.e. at some world where the contextually determined facts hold.[2]

But how *does* the modal base (the accessibility relation, the relevant possibilities, or the facts held fixed) get selected? Neither Lewis nor Kratzer gives a fully systematic answer to this question, and presumably neither of them believed there could be such an answer. But Lewis (1979) tells us that the conversational score at any time in a conversation is determined by such factors as what has been said before, what has been conspicuous to all participants, and importantly by a 'rule of accommodation':

> conversational score does tend to evolve in such a way as is required in order to make whatever occurs count as correct play. [...] Here is a general scheme for rules of accommodation for conversational score:
>
>> If at time *t* something is said that requires component s_n of conversational score to have a value in the range *r* if what is said is to be true, or otherwise acceptable; and if s_n does not have a value in the range *r* before *t*; and if such-and-such further conditions hold; then at *t* the score-component takes some value in the range *r*.
>
> (Lewis 1979, 347)

Kratzer too postulates a rule of accommodation: cf. Kratzer (1981, 61).

3.3.3 Kratzer on Dispositions and Abilities

Intuitively, a sentence such as 'I can play the piano' ascribes an ability. What does the possible-worlds semantics make of such uses of 'can'? The answer is suggested by some remarks made in passing by Kratzer (1981) about dispositional terms. Typical

[2]Kratzer's ordering source has no analogue in these Lewisian sketches; since this difference is irrelevant to my discussion in the following, I will treat Kratzer and Lewis as defending the same account in this respect.

expressions of dispositions are adjectives ending in '-ible'/'-able' or '-ile', and Kratzer counts these suffixes among the modal expressions to be covered by her account. The dispositional 'fragile', for instance, in a sentence 'This cup is fragile' expresses 'purely circumstantial' possibility: 'It is in view of certain properties inherent in the cup, that it is possible that it breaks. The ordering source seems to be empty.' (Kratzer 1981, 64) The modal base, then, consists of those worlds where the relevant properties are possessed by the cup.

Which are the relevant properties that determine the modal base? In the case of the glass, we might just want to say: its internal constitution. Is the same sufficient for the apparently ability-ascribing uses of 'can'?

Consider the following story (a variation of the unwilling piano player in chapter 3.2.2): Jenny was trained to be a pianist from a very early age; in fact, she was a child prodigy on the piano. Then terrible things happened, which were connected to her playing the piano and which severely traumatized her. Now when Jenny sits down at a piano the memories overcome her and she is paralyzed, unable to move, let alone play the piano. When asked to sit down and play something on the piano, Jenny may truly say

(¬C1) I cannot do it [play the piano].

When we learn about her past, however, we can equally truly say

(C1) Jenny can play the piano.

The suggested treatment can account for (¬C1) but not for (C1): in all the close possible worlds where Jenny has the same internal constitution as she actually does (including both her mental and her physical properties), she does not play the piano. How, then, is the modal base picked for (C1)? It must be more fine-grained: some of

Jenny's properties must enter into it – her ability to play the piano and all the physical and mental properties associated with it – while others must not – most importantly, the traces of her traumatic experience and her disposition to become paralyzed whenever confronted with a piano. It is difficult to see how we could separate the factors which do enter into the modal base from those which do not, except by reference to the very ability that is being ascribed: the modal base comprises those worlds where Jenny's ability to play the piano is as it actually is, while other factors, external and internal, may vary.

Angelika Kratzer might very well agree with what I have said so far. In fact, she explicitly includes modal properties such as abilities and dispositions among the factors that enter into the modal base. They facilitate the semantics, by providing the right kind of modal base; and a systematic formal semantics is all that she is concerned with.

3.3.4 Lewis on Relevant Facts and Conversational Score

David Lewis is pursuing a more ambitious programme. According to his well-known doctrine of Humean Supervenience, 'all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another ... we have an arrangement of qualities. And that is all.' (Lewis 1986b, x) Abilities, dispositions and their like are not part of the basic furniture of our world; they supervene on the Humean basis, which consists solely of local, intrinsic, and purely categorical properties. Any expression that makes reference to abilities or dispositions can in principle be replaced by one that does not. And it is one central task of modal semantics to show how this replacement is achieved. Thus Lewis must have a rejoinder to my argument in the preceding section; and he does. Here it is.

> To say that something can happen means that its happening is compossible with certain facts. *Which* facts? That is determined, but sometimes not determined well enough, by context. An ape can't speak a human language—say, Finnish—but I can. Facts about the anatomy and operation of the ape's larynx and nervous system are not compossible with the ape's speaking Finnish. The corresponding facts about my larynx and nervous system are compossible with my speaking Finnish. But don't take me along to Helsinki as your interpreter: I can't speak Finnish. My speaking Finnish is compossible with the facts considered so far, but not with further facts about my lack of training. What I can do, relative to one set of facts, I cannot do, relative to another, more inclusive, set. (Lewis 1976, 77)

This does not directly pertain to the kind of case that I have been considering, although we get a similar pair of statements:

(C2) David Lewis can speak Finnish.

(¬C2) David Lewis cannot speak Finnish.

(C2) is true because David Lewis's speaking Finnish is compossible with certain facts about his larynx and nervous system. Or more precisely: (C2) is true because the conversational score at this point has it that the accessible worlds are those where facts about Lewis's larynx and nervous system are held fixed, and Lewis does speak Finnish at some such world. (¬C2) is true because his speaking Finnish is not compossible with the former facts *and* facts about his lack of training. Or more precisely: (¬C2) is true because the conversational score at this point has it that the accessible worlds are those where facts about Lewis's larynx and nervous system *and* about his lack of training are held fixed.

Analogously, Jenny can play the piano – (C1) is true – because her playing the piano is compossible with certain facts about her nervous system, her hands, and her early training. (¬C1) is true because Jenny's playing the piano is not compossible with facts about all of the above *and* facts about her later, traumatizing, experiences. (Precisify as before. Note, incidentally, that we have some leeway in choosing to speak about her history – the training – or her intrinsic features – the neurological traces of said training; nothing much seems to depend on which one we choose. Lewis seems to prefer the former.)

Now, how *does* context determine, in any one of the cases Lewis and I have described, which facts are to be 'held fixed' across the accessible worlds? Why is it aspects of David Lewis's biology on one occasion, aspects of his history on another? Why one part of Jenny's history (piano lessions) on one occasion, and another part (her trauma) on another?

For the answer, we need to look back to Lewis (1979) and the 'rule of accommodation' that I have described earlier: as Lewis told us, 'conversational score does tend to evolve in such a way as is required in order to make whatever occurs count as correct play' (Lewis 1979, 347). Thus (C2) requires for its truth (I will neglect other aspects of 'correct play') that it is facts about David Lewis's larynx and nervous system (but not his lack of training) that are held fixed across the relevant worlds; and so they are. (¬C2) requires for its truth that facts about David Lewis's lack of training be held fixed in addition to the above; and so they are. And so on for (C1) and (¬C1).

However, that is not quite true: the truth of, say, (C2) does not *require* that the relevant facts be facts about David Lewis's larynx and nervous system: other facts, say, facts about his stomach, would do just as well. Those facts are certainly compossible with David Lewis's speaking Finnish. Similarly, in the case of (C1), holding fixed facts about Jenny's stomach or her hair-colour would make the statement come out true. If the rule of accommodation could select only what is *required* for the truth of a given statement, it could not select at all.

In response to this, it may seem that all we need to do is rephrase the rule of accommodation and say: conversational score tends to evolve in such a way as is *sufficient* to make whatever occurs count as correct play, and in particular, to make an utterance come out true. (Perhaps that is precisely the import of Lewis's appeal to a 'range' of values in his formulation of the rule.)

That, however, only raises another question: how does the conversational score 'choose' among the different eligible facts: why should the conversational score select facts about David Lewis's larynx rather than his stomach, or about Jenny's hands rather than her hair-colour? Nothing in the conversations we were given had made larynxes or hands more salient (by mentioning them or making them otherwise conspicuous) than stomachs and hair-colours.

Lewis's account seemed plausible because because we already know that larynxes and training are relevant to one's ability to speak Finnish, and hands and (another type of) training are relevant for one's ability to play the piano. It would lose much of its appeal if it turned out that those facts had been arbitrary examples and Lewis might just as well have referred to stomachs and hair-colours. For a semantic account of a type of sentence, it is not enough that all the right sentences (or utterances of them) come out true or false, as the case may be; they should do so for the right reasons. If (C2) comes out true because David Lewis's speaking Finnish is compossible with facts about his stomach, then something has gone wrong.

So far, Lewis has given his semantic mechanism nothing to go by but the rule of accommodation, and that rule, properly rephrased, has no resources to distinguish between stomachs and larynxes when it comes to the truth of (C2). Intuitively, it is clear what makes larynxes good candidates and stomachs bad ones when it comes to the truth conditions for (C2) (and similarly for the other examples): facts about David Lewis's stomach or Jenny's hair-colour are irrelevant to their abilities to speak Finnish or play the piano, respectively; but the larynx and the nervous system are the physical base of the ability to speak Finnish, and training is causally related to one's ability to speak Finnish or play the piano. In general, it is by appeal to the abilities ascribed that 'good' candidates are distinguished from 'bad' candidates for score-entering facts: the good ones are facts that contribute (most obviously, as causes or physical bases) to the ability in question.

The rule of accommodation is a very general device that applies to many expressions besides 'can'. Why not, then, add an extra clause to the semantics of sentences '*x* can F' to specify that conversational score prejudices facts that are relevant to the ability to F? Lewis says no such thing, and that is no accidental omission. He *cannot* say any such thing because of the role that the analysis of expressions such as 'can' plays for Humean supervenience. Let me look at that role in a little more detail.

3.3.5 Humean Supervenience

Here is David Lewis's own much-quoted characterization of Humean supervenience:

> Humean supervenience is named in honor of the greater [sic] denier of necessary connections. It is the doctrine that all there is to the world is a vast mosaic of local matters of particular fact, just one little thing and then another. (But it is not part of the thesis that these local matters are

> mental.) We have geometry: a system of external relations of spatiotemporal distance between points. Maybe points of spacetime itself, maybe point-sized bits of matter or aether or fields, or both. And at those points we have local qualities: perfectly natural intrinsic properties which need nothing bigger than a point at which to be instantiated. For short: we have an arrangement of qualities. And that is all. There is no difference without difference in the arrangement of qualities. All else supervenes on that. (Lewis 1986b, ix f.)

It will not come as a surprise that I reject Humean supervenience. I believe that the supervenience base, if there is one, is not Humean: it must include potentialities.

(Note, incidentally, that imagery aside, the quote does not clearly say that potentialities are excluded from the supervenience base: being possibility-like, potentialities need not bring with them any necessary connections, and on a non-reductionist view it is hard to see why all potentialities should require more than a point to be instantiated. The reason for this is not, of course, that Lewis was open to potentialities at the supervenience base, but that his main target – and the main issue in debates about Humean supervenience – is the package of laws, causation, and dispositions according to the standard model. The exclusion of potentiality from the supervenience base is implicit in his talk of 'qualities', which I take it we are to understand as categorical properties, properties with no essential modal features, or 'quiddities'. Potentialities surely are not such properties.)

Lewis's semantic analyses of modal terms such as 'can' is best understood in the context of his project of defending Humean supervenience. In particular, they are aimed at reconciling common sense and ordinary language, both of which abound in modality, with Humean supervenience: to show that the Humean supervenience base, enriched by other possible worlds, is all we need to make the statements of ordinary language come out true or false exactly as common sense dictates. Since that is its goal, the analysis must not appeal to anything other than the Humean supervenience base. Lewis does not claim that analyses (such as Kratzer's) which do appeal to more than that supervenience base are false, and do not serve their own purposes, whatever they may be. From the point of view of Humean supervenience, they are simply incomplete; and from the point of view of defending Humean supervenience by reconciling it with common sense, they are useless.

We can think of the semantics as a mechanism that assigns semantic values to token linguistic expressions (say, propositions to utterances of sentences) based solely on the linguistic facts and the facts concerning the Humean supervenience base. Lewis's account of the 'conversational score' tells us something about how that mechanism works; the rule of accommodation, in particular, tells us that *ceteris paribus* the mechanism assigns true propositions wherever possible. To vindicate Humean supervenience, the mechanism should 'know' nothing about the world except for how things stand with the Humean supervenience base. In describing the mechanism, or giving the semantics of a language such as English, we should in principle be able to use a metalanguage whose terms refer to nothing but token linguistic expressions and the supervenience base.

We can see why Lewis did not adopt the view that I earlier attributed to Angelika Kratzer, a view on which an object's possessing a particular ability was part of the facts that are held fixed across the relevant worlds. To operate in accordance with that view, the semantic mechanism must 'know' about abilities; otherwise it could not hold them fixed across the relevant worlds. But then the analysis is of no use to the defender of Humean supervenience. Larynxes and training events, to be sure, aren't part of the Humean supervenience base either, but they are certainly one step closer to it. In particular, they are closer to it in being non-modal. The examples illustrate how the semantics of a modal expression is given without appeal to modal constituents of the world in which it is uttered. Since the supervenience base is non-modal, that illustration takes us one step closer to the final reductive project.

We can also see why Lewis did not provide any general guidance for the semantic mechanism to help it pick out larynxes rather than stomachs, etc.: if such guidance were phrased (as it would most naturally be) in terms of abilities, then the semantic mechanism would again need to 'know' about abilities. Abilities need not be constituents of the facts to be held fixed across worlds (as they are on the view I attributed to Kratzer), but they are still a vital input for the mechanism: without knowing about abilities, the mechanism would not know which facts to hold fixed.[3]

This, then, explains why Lewis relies on examples alone and does not make the natural appeal to abilities in explaining those examples; if he did, the analysis would no longer serve its purpose in the defense of Humean supervenience. Is there any other, Humeanly acceptable, way of providing what I asked for: a way for the semantic mechanism to distinguish the good from the bad candidates for relevant facts? It seems to me that there are two broad strategies available to Lewis. One is to accept the

[3]What if abilities could be reduced to the Humean supervenience base plus other worlds in some other way than the Lewisian? Kenny (1976) has argued that the possibility operator $\diamond$ of any normal modal logic is unsuited to express ability, not because of the kind of worries that I have formulated, but because it has the wrong logical structure: ability, according to Kenny, does not distribute over, nor is it closed under, disjunction. Possibility, even on the weakest normal modal logic, does both. Brown (1988) and Horty and Belnap (1995) have taken up Kenny's challenge by providing non-normal modal logics of ability. The basic idea of both their systems is that ability includes an aspect of necessity, or 'control', as well as of possibility. Brown introduces the notion of 'clusters' of worlds, which 'correspond to different logically possible actions.' He takes '*relevant* clusters to correspond to choices of actions of which I am actually *capable*.' (Brown 1988, 5) Then informally, 'I can bring it about that *p*' is true iff there is some relevant cluster of worlds in each of which *p* holds; i.e., iff there is some action of which I am capable and which ensures the truth of *p*. I do not wish to take a stand on Kenny's challenge. The important thing to note in the present context, however, is that Brown's semantics for the formal system, like Kratzer's semantics for the English 'can', is non-reductive: the notion of an action 'of which I am capable' is built into the semantics unanalysed. The same is true of Horty and Belnap's system.

challenge and provide an answer to it without appeal to abilities; the other is to reject the challenge altogether.

If Lewis were to take the first option, I think the most promising strategy would be to focus, not on the abilities as I have suggested, but on their exercises. In considering (C2), the relevant facts are those about David Lewis's larynx and nervous system because larynxes and nervous systems are normally involved in someone's speaking Finnish; stomachs aren't. In considering (¬C2), the relevant facts further include those about David Lewis's (lack of) training, because training-events normally precede events of people speaking Finnish. In considering (C1), Jenny's hands, nervous system and previous training are relevant because hands and nervous systems are normally involved in playing the piano, and training-events normally precede events of people playing the piano. And so on. The idea is that the relevant facts are those which are *normally* connected to the exercise of the ability I have appealed to; normality will be understood as some statistical (intra- or inter-world) measure. In general, the strategy is to replace my earlier talk about the bases and causes of an *ability* to F with talk about what is normally involved in F'ing, and what normally preceeds F'ing.

This strategy fails because abilities and their exercises are multiply realizable. We speak (Finnish or otherwise) using our larynxes but other beings might not. The failure becomes more obvious when we look at (i) cases where F'ing normally involves, or is preceded by, X, but a particular individual possesses the ability to F not in virtue of X but in virtue of some atypical base or cause, Y – say, someone can play the piano not with their hands but with their feet; or (ii) cases where there is no 'normal' base or cause, where F'ing is not only multiply realizable but is in fact performed by different individuals in so many different ways. We might say of someone 'He can deceive anyone' or 'He can adapt to new environments'; there is no normal way of deceiving or adapting, no typical training preceding them, or anything of that kind; what needs to enter the conversational score here is simply the *actual* cause or base of the ability in the given individual.

The second option is one that I suspect Lewis would have preferred all along: refuse the challenge. I have pointed out that our linguistic practices concerning 'can' are more restrictive than his rule of accommodation can account for, and have challenged him to provide a general explanation of these restrictions. To refuse this challenge is to say that, for all we know, context might fix on David Lewis's stomach in the truth-conditions for (C2), and on Jenny's hair-colour in the truth-conditions for (C1); but if semantics (like nature) is kind to us, then it won't.

I cannot refute that claim, but it seems to sacrifice a great deal. The account certainly becomes less informative; it thereby becomes less attractive too. The two versions of possible-worlds analyses that do make reference to abilities (the one that I have attributed to Kratzer, which admits abilities into the 'modal base' or relevant facts, and the alternative that merely allows the semantic mechanism to be guided by knowledge of abilities) certainly seem to do better as far as the explanatory value of the semantics goes.

Where does that leave us in the enterprise of trying to give a possible-worlds semantics of 'can'? The only reason for adopting the Lewisian version is its role in defending the tenability of Humean supervenience. Tenability, officially, is all that Lewis ever seeks or claims to establish for Humean supervenience: 'Really, what I uphold is not so much the truth of Humean supervenience as the *tenability* of it. If physics itself were to teach me that it is false, I wouldn't grieve.' (Lewis 1986b, xi) Now, the viability of Lewis's semantics may serve as a defense for the tenability of Humean supervenience. But the tenability of Humean supervenience in itself is no reason to adapt our semantics to its defense. It is certainly not good methodology to try to provide a semantics that could serve to defend all tenable theses in metaphysics, not least because there are many tenable theses that contradict each other.

If Humean supervenience were true, of course, that would be a different matter. But it is much more likely to be false.

As Maudlin (2007) notes, it is unclear why we should believe in Humean supervenience in the first place. Lewis's own stated motivation is 'to resist philosophical arguments that there are more things in heaven and earth than physics has dreamt of' (Lewis 1994, 474). But, as has frequently been pointed out, there are more things in physics itself, or to be precise: in quantum physics, than Humean supervenience has dreamt of; physics itself does teach us that Humean supervenience is false. At the heart of Humean supervenience lies the idea that the world is 'Separable': the total physical state of the world supervenes on the local, intrinsic states of each spacetime point and the spatio-temporal relation between them; 'the world as a whole is supposed to be decomposable into small bits laid out in space and time' (Maudlin 2007, 51). Separability is built into classical physics but rejected by part of our currently best physics, quantum theory. According to quantum theory, two electrons may be in an 'entangled state'. If they are, then the state of the system consisting of a pair of electrons e_1 and e_2 does not supervene on the intrinsic states of e_1 and e_2. On one construal, if e_1 and e_2 are in an entangled state, then neither of them 'has any local state, i.e. a state that can be specified without reference to the other member of the pair.' (Maudlin 2007, 61) Obviously, if there are no local states to begin with, then the 'global' state of the pair has no local states to supervene on. Another construal allows us to attribute intrinsic local states to each of e_1 and e_2, but a given combination of such local states of e_1 and e_2 plus a given spatiotemporal relation between them is compatible with more than one 'global' state, so again that global state does not supervene on the local states, and Separability fails. (See Maudlin 2007, 61, for this way of putting the point. Objections along the same lines are mounted, for instance, in Oppy 2000 and Karakostas 2009.[4])

There is also good reason to believe that the fundamental properties, those which make up the supervenience base, are also not 'qualities', that is, nonmodal quidditistic properties. Science, as Simon Blackburn and others have argued, 'finds only dispositional properties all the way down' (Blackburn 1990, 63; see also Molnar 1999 and Bird 2007).[5] In maintaining that there is more to the properties discovered by science, namely, a quiddistic nature, it is Humean supervenience that is guilty of supposing that 'there are more things in heaven and earth than physics has dreamt of'.

Apart from seeming unwarranted by the standards of physics, the supposition of quidditistic fundamental qualities leads to various problems. If Humean supervenience is true, then the fundamental properties are not the dispositions that physics *prima facie* takes them to be. Accordingly, the names of such a fundamental property (say, 'charge') do not refer to dispositional properties, but to the categorical properties that plays the role specified in the disposition's description (see Lewis 1970). This, however, leads to a rather unattractive proliferation of possibilities: on the assumption that charge and mass are categorical properties that play their dispositional roles only contingently, it should be possible that they completely swap their roles (see Black 2000, who adapts an argument from Chisholm 1967 concerning haecceitism). So there should

[4]It has even been suggested that quantum physics is best understood in terms of fundamental potentialities (see Karakostas 2007, who quotes Heisenberg to the same effect). However, I do not know enough about quantum physics to fully understand, let alone evaluate, that claim.

[5]In the same paper, Blackburn suggests that this finding should cause concern: 'To conceive of *all* the truths about a world as dispositional, is to suppose that a world is entirely described by what is true at *neighbouring* worlds. And since our argument was a priori, these truths in turn vanish into truths about yet other neighbouring worlds, and the result is that there is no truth anywhere.' (p. 64) Holton (1999) has shown that the worry is ungrounded if it is one of incoherence. The worry that remains is one of regress or circularity, if the fundamental properties are dispositions whose manifestations are in turn dispositions, whose manifestations ... and so forth. As Holton points out, it is not entirely clear what is so bad about circularity in this case. For a more detailed argument, see chapter 6 of Bird (2007).

be possible worlds just like the actual one in every detail except that mass plays the charge role and charge the mass role (and the same goes for any number of properties). Perhaps more worryingly, there is no reason why the same role should not be played by two distinct, indeed by any number of, categorical properties, in the actual world. But that would make it impossible for our theoretical terms such as 'charge' to refer to anything at all: if there is no one categorical property that plays the charge role, then there is no property for 'charge' to refer to (compare the definite description 'the word on this page'), and the best science that we could possibly achieve would merely deal in empty words. (The argument is from Alexander Bird 2007, 76-79.)

To conclude. It is hard to see why we should impoverish our semantics and deny ourselves a great deal of explanatory value for the sake of a speculative thesis that is most likely to be false. Kratzer's version of the semantics is superior to Lewis's as a semantic theory, and the independent motivation for Lewis's version has been found wanting: the view is designed to align philosophy with our best science, but it seems not in accord with either our best scientific theories or with scientific practice. My argument has been sketchy, and its conclusion is far from definitive. But once again, it seems that the burden of proof is on the opposite side: on those who hold that the speculative thesis of Humean supervenience should constrain our semantic theory.

3.3.6 A Provisional Conclusion

In section 3.3.1, I proposed that given its surface structure, 'can' is most plausibly construed as a predicate modifier that ascribes potentialities to objects. I suggested four reasons for the wide-spread resistance to this idea, and the view that 'can' should instead be treated as a sentence modifier expressing truth in at least one possible world. Those reasons were: (i) the possibility of providing unified treatment of circumstantial and epistemic modality; (ii) the elegant account afforded by the possible-worlds semantics of the context-dependence of 'can'; (iii) Humean supervenience; and (iv) the worry that there will not be enough potentialities for a comprehensive semantics of 'can', a worry which does not affect possible worlds. I addressed (i) immediately, and have been concerned with (ii) and (iii) in the last four sections. While my discussion has not been entirely conclusive, its upshot is this: (ii), the elegant account of the context-dependence of 'can', is best achieved without (iii), the attempt to defend Humean supervenience. The reductive project of Humean supervenience stands in the way of the best possible-worlds semantics of 'can'; that semantics makes reference to abilities in determining the modal base or relevant facts. (It may be conjectured that it will also have to make reference to other kinds of potentialities in other cases.) The proposal I made in section 3.3.1 has an explanation for this: abilities are indispensable to the best and most informative semantics of 'can' because ascribing abilities (and other potentialities) is precisely what 'can' does. Of course, that is not the only explanation available for the superiority Kratzer's over Lewis's version of the possible-worlds semantics for 'can'. But it makes the burden of proof on those who reject surface grammar in favour of a possible-worlds semantics just a little heavier.

My proposal is left with two challenges, arising from (ii) and (iv). First, it remains a fact that the possible-worlds semantics gives a very elegant treatment of the context-dependence of 'can'; has the proposed potentiality-based semantics anything comparable to offer? Second, the challenge posed by (iv) has not been addressed: are there enough potentialities to provide for all (true) uses of the circumstantial 'can'? Both challenges can be met, if at all, only once a general account of potentiality has been developed. The second thread of this chapter ends with the same desideratum as the first: the need for an account of potentiality in general. I will go on to provide such an account in the next chapter. The account should address the challenge that arises from (iv). I will not be able, in the scope of this thesis, to address the challenge of context-sensitivity (ii), though I will make some preliminary remarks on how it might be addressed at the end of chapter 4. I hope to have given some reason to take the proposal of a potentiality-based semantics of ‘can’ seriously; no more than that, but also no less.

Chapter 4

Potentiality

4.1 Generalizing

In chapter 3, I have provided some motivation for the development of a general and non-reductive account of potentiality – where ‘potentiality’ is understood as the common genus of dispositions, abilities, and whatever other properties of the same kind there are. Apart from its intrinsic interest to anyone with nonreductionist inclinations, I indicated two further aims which a general account of potentiality might serve. One is the development of a potentiality-based account of possibility; that will be my focus in chapters 5 and 6. The second is a semantics of ‘can’. That second aim is not my main target in this thesis, but it will be addressed in a sketchy and somewhat scattered manner along the way in this chapter. In particular, section 4.2 will introduce a conception of potentiality that is by stipulation weak and inclusive enough to serve in a semantics of ‘can’; and I will then argue that this conception should indeed be adopted. (Its use is not limited to the semantics of ‘can’; since the weak conception is plausibly closed under logical entailment, it will also enable me to formulate a more elegant logic of potentiality in chapter 6.)

In order not to lead to circularity when serving these two aims, the general account of potentiality should be developed without appeal to either the analogy between potentiality and possibility or the link between potentiality and 'can'. My method will instead be one of natural extension using the examples that helped introduce the notion of potentiality: dispositions, such as fragility, and abilities. I will argue that such extension takes us very far. The general strategy will be to provide examples from these two familiar classes of potentialities and then to argue for the inclusion of less familiar cases by arguing that any line between the familar and the unfamiliar case would be arbitrary.

To begin, consider fragility: there are some very fragile things, such as old parchments and long-stemmed champagne glasses, and others that are less so, but still fragile, such as an ordinary tumbler. Among the non-fragile things, there are some that come closer to being fragile, such as a robust plant-pot, and others that are far remote from being fragile, such as my desk or (even further) a rock. There seems to be a continuous spectrum that leads all the way from the champagne glass to the rock, and somewhere on that spectrum we can find the cut-off point – or perhaps we cannot, since it is a vague matter precisely where that cut-off point is. Even if we could precisely locate it, however, that cut-off point would not be likely to cut nature at its joints; it might have been a little more to one side of the spectrum or the other (and in another context we would probably say it is), thus excluding the tumbler or including the plant-pot.

I have said that there is a continuous spectrum; but a spectrum of what? The answer seems to be: of possibility-like properties that come in varying degrees of strength; in other words, of potentialities. Potentialities provide the non-reductionist with a way to understand the context-sensitivity of dispositional terms such as ‘fragile’: context determines the degree to which a given object must possess the potentiality to break in order for that object to count as fragile. But context can operate in this way only against the background of something that is not context-sensitive. Standardly, that background is taken to be possible worlds: which possible worlds there are (whatever they are) is not a matter of context; context merely determines which of them are relevant on a given occasion. On the non-reductionist view of potentiality, the background for contextual variation is provided by the potentialities that things possess and the degrees to which they possess them.

Perhaps that has not yet taken us beyond the realm of dispositions; the potentiality to be broken that comes in degrees and provides the background for the context-sensitivity of ‘fragile’ may, after all, be expressed by another dispositional term, ‘breakable’. (For the distinction between these two, see chapter 2.4.) But adjectives of the form ‘F-able’ provide limited expressive resources; they take only a single verb as a complement, and (if my conjecture is correct) only in the passive form. (Again, see chapter 2.4.) We can go further.

Consider abilities. I am hopeless at sports; I cannot even catch a ball if you throw it at me. That is to say, I have no ability to catch a ball if you throw it at me. Having an ability requires some kind of control, and I do not have that control. (Spelling out the kind of control required is complicated by cases such as our traumatized piano player in chapter 3.3.3; but my case is clearly not like this.) Still, I may be lucky in some cases, and catch the ball thrown at me. There is nothing about me that prevents me in principle from catching a ball thrown at me, in the way that lacking arms would prevent me from catching a ball, or in the way that a table is in principle prevented from catching a ball. Compared to my desk, I *can* catch a ball. (And compared to an ape, David Lewis *can* speak Finnish.) This is a fact about me, a property of me, just as its potentiality to break is a property of the plant-pot. It seems reasonable to say, then, that I have a potentiality to catch a ball when it is thrown at me, a potentiality which my desk lacks; but that my potentiality to catch a ball, unlike (perhaps) yours, is not an ability.

Potentialities, again, provide the non-reductionist with a way to understand the context-sensitivity of certain expressions, in this case 'can'. I truly said above that I *cannot* catch a ball. But of course, there is a perfectly good sense in which I *can* catch a ball, and it is evoked by the comparison with my desk. In the first context, 'can' was used to ascribe a potentiality that qualifies as an ability; in the second, it is used to ascribe a potentiality whether or not that potentiality qualifies as an ability. What makes a potentiality qualify as an ability? I do not have an answer to this question. (I wish I did!) Answering it may require doing a substantial amount of work in the philosophy of action. But we clearly have some intuitive grasp of the distinction even without being able to give a general account of it.

It would seem, then, that there are more potentialities than there are dispositions and abilities; but just how many more are there? That question is of course ill-phrased. To reformulate it, let me first say something about the individuation conditions for potentialities.

In general, a potentiality is individuated by its manifestation and nothing else. (I have argued this point at length for dispositions in chapter 2.) It is a potentiality to ...; and that is all that it is. A potentiality to F may be possessed by different individuals, or by the same individual at different times, to different degrees; and it may or may not come with certain features that make it qualify as an ability. I have professed ignorance on the question of just what those features are, but for the time being I submit that whatever they are, they are additional features of one and the same potentiality.

If a potentiality is individuated by its manifestation, and its manifestation alone, then the obvious reformulation of my ill-phrased question is this: what manifestations can a potentiality have? The answer is: in the widest possible sense of 'property', every property is the manifestation of some potentiality.

This widest possible sense includes complex properties: if there is a property of being F and a property of being G, then there are also properties of being F-and-G and of being F-or-G, for instance. Why should there by potentialities for such manifestations? Here are some intuitive examples.

Many potentialities have complex manifestations. Take, as a simple example, the ability to walk. Walking is an intricate combination of movements: lifting one foot while leaving the other on the ground, moving the lifted foot forward and putting it back on the ground while lifting the other, and so on. In having an ability to walk, I already have an ability for a conjunctive manifestation: the ability to lift one foot and move it forward and leave the other foot on the ground, etc. If things never had potentialities for conjunctive manifestations, then I could not have an ability to walk. I take that to be a reductio.

What about disjunctive potentialities? Here the intuitive example is determinate and determinable properties.

'Breaking' is a determinable with such determinates as: breaking into two pieces, breaking into three pieces, and so forth. Having one determinate, say breaking into seven pieces, is a *way* of having the determinable property, breaking. In fact, it is the only way: there is no other way to possess a determinable property than by possessing one of its determinates. Everything that breaks must break into some number of pieces.

Conversely, possession of any one determinate ensures possession of the determinable: by breaking into seven pieces, an object must break.

In these respects, determinables and determinates behave exactly as a disjunctive property and its disjunct properties do: being F is a way of being F-or-G; in fact, being F and being G are the only ways of being F-or-G (nothing can be F-or-G without being F or being G); and finally, having one of the disjunct properties already ensures having the disjunctive property. The determinates of a given determinable exclude each other, while the disjuncts of one disjunctive property do not, but that seems to be the only notable difference in the formal structures. Having a determinable property at least entails having the disjunctive property that has all its determinates as disjuncts.

Given that fragile objects have the potentiality to break, and that breaking is a determinable property, it is therefore extremely difficult to see how a fragile object should fail to have the potentiality for a disjunctive manifestation: the potentiality to break into two pieces or to break into three pieces or .... And given that there are fragile objects, we should conclude that there are objects with potentialities for disjunctive manifestations.

This type of reasoning carries over to more arbitrary disjunctive manifestations, and also to tautological ones. Do I have a potentiality to walk or not to walk? If I do, then that is a potentiality which I am constantly manifesting. That in itself should not be a problem: an animal's ability to breathe is manifested throughout the animal's life; fire constantly exercises its disposition to release heat. But how do I exercise the potentiality to walk or not to walk? Very simply: by walking, or by not walking, as the case may be. Walking and not walking are both *ways* of having the tautological property of walking-or-not-walking, just as breaking into two pieces and breaking into seven pieces are ways of breaking. Similarly, I possess the potentiality to walk or not to walk by possessing the potentiality to walk, or by possessing the potentiality not to walk, or both.

We can go further still. There are not only potentialities for complex properties, there are also potentialities for 'such-that' properties. (A 'such-that' property is a property that is possessed either by everything or by nothing, depending on whether a particular proposition is true. Semi-formally, it is a property that can be expressed with a lambda operator expression, $\lambda x.p$, where p contains no free occurrences of x. An object possesses a 'such-that' property usually without making any contribution to its possession of that property.) I begin with a complex property. I have an ability to walk in the rain; I do not have an ability to walk in a tornado. In other words, I have an ability to walk *while* it is raining, or an ability to walk *and be such that* it is raining. In order for me to exercise that ability it must be the case both that I am a certain way (namely, walking) and that things stand a certain way outside me (namely, that it is raining). Once we have admitted 'such-that' properties as parts of complex manifestations, there is no reason for not allowing them to form a potentiality's manifestation on their own: by having a potentiality to walk, I have a potentiality to be such that I am walking, and indeed to be such that someone is walking. The manifestations of the former two potentialities (to walk, and to be such that I am walking) by me will always consist in the very same events or states, namely, my walking. Nevertheless, the manifestation properties themselves are to be distinguished: their possession by anyone but me will take very different forms.

By stretching our notions of the familar kinds of potentialities only a little, we have now already reached a rather liberal conception of potentiality. It is liberal in two ways.

First, it is liberal concerning which things have which potentialities. Rocks have potentialities to break, and someone as hopeless at sports as I am has a potentiality to catch a ball. To be sure, there are differences between rocks and champagne glasses in terms of the potentiality to break, and differences between me and a professional handball player in terms of the potentiality to catch a ball. But those, I submit, are best construed as differences *between potentialities* – differences in their degree, or (more vaguely) differences in the control that their possessors have over their exercise – and not as differences between having and lacking a potentiality.

Second, the conception we have arrived at is liberal concerning which properties can be the manifestations of potentialities: they include complex properties of the conjunctive or disjunctive variety, as well as 'such-that' properties, and complex properties involving 'such-that' properties.

Taking these two strands of liberalism together, it seems plausible that for every property F that an object *a* possesses, *a*'s possessing F is a manifestation of a potentiality of *a*'s. For why should it not be? It might be because, while *a* did possess a potentiality to have F, its possession of F did not come about in the right way. Given the first strand of liberalism, it is difficult to see how that should be the case. (If you catch a ball accidentally, that may not count as an exercise of your ability to catch a ball, or of your potentiality insofar as it is an ability; but it will still be an exercise of that potentiality, just as much as my catching a ball is an exercise of my potentiality which does not qualify as an ability.) Or it might be because F is not the kind of property that things (of *a*'s kind) might ever have a potentiality to possess; but given my second liberal strand, it is not obvious that there are any such properties left. Or are there?

4.2 Two Notions of Potentiality

Take two apples, a_1 and a_2, and suppose that at a certain time t, a_2 changes its colour to a bright red, while a_1 does not. a_2 has acquired a property: the property of having a certain shade of red. a_1, too, has acquired a property: the property of being such that a_2 is red. So how many potentialities have been manifested?

I believe that the intuitive answer is: one. a_2 has manifested a potentiality to turn red. a_1 has not manifested any potentiality by a_2 turning red: that event had nothing to do with a_1. One way of taking account of this is to simply forbid potentialities for 'such-that' properties: nothing, and a fortiori not a_1, has a potentiality to be such that a_2 is red. But we have seen that those properties are needed to account for some intuitive examples (for instance, my ability to walk in the rain). Moreover, once we have allowed for there to be such a property as the property of being such that a_2 is red, we must allow that a_2 has acquired that property by turning red. It would then be arbitrary not to allow that a_2 has manifested its potentiality to be such that a_2, itself, is red. To do justice to the intuition to which I pointed, we do not need to forbid potentialities for 'such-that' properties; we need to forbid merely that objects (e.g., a_1) have potentialities for 'such-that' properties that do not concern themselves (e.g., the property of being such that a_2 is red). A restriction is to be posed not on what properties things, in general, have potentialities for, but on which things have potentialities for which properties. a_2 may, but a_1 may not, have a potentiality to be such that a_2 is red.

The intuitive idea behind this reply, I take it, is this. An object's manifesting or exercising a potentiality should be something that the object itself does, though in the weakest possible sense of 'does'. No matter how weakly we understand 'do', however, it is plain to see that a_1 would do nothing at all in manifesting a potentiality to be such that a_2 is red; the doing is all a_2's. An object must exercise its potentialities itself, not have other objects exercise them, as it were, for it; and we should attribute to an object only such potentialities as it could exercise itself. Exercising a potentiality may require some contributions from objects other than the potentiality's possessor (witness my potentiality to walk in the rain – in order for me to exercise this potentiality, the weather must comply), but it will always require some contribution from the potentiality's possessor itself; or so the thought goes. Intuitively, a_1 has not contributed to, and could never contribute to, being such that a_2 is red. Hence a_1 is in principle excluded from having a potentiality to be such that a_2 is red, since it is in principle excluded from exercising such a potentiality. We certainly do not ascribe any dispositions or abilities for such properties; if potentialities are introduced by extending our concepts of abilities and dispositions, it is hard to see how they should ever 'reach' any such potentiality as a potentiality possessed by a_1 to be such that a_2 is red.

However, there is an alternative train of reasoning in response to my earlier question: how many potentialities were manifested in a_2's turning red and a_1's acquiring the property of being such that a_2 is red? The alternative answer, unsurprisingly, is: two. This answer goes with the idea that 'can' is used to ascribe potentialities. For, while it is quite unusual to say that a_1 can be such that a_2 is red, it still seems true; and it is certainly false to say that a_1 *cannot* be such that a_2 is red. Hence, if (as I suggested in chapter 3.3.1) 'can' is to be construed as ascribing potentiality, we must ascribe, rather than deny, to a_1 a potentiality to be such that a_2 is red. In fact, it would seem that if usage of 'can' were to be our guide to potentiality, things would have potentialities for whatever is compatible with their existence. For not only can a_1 be such that a_2 is red; a_1 can be such that all apples are red, that all cucumbers are green, and it can be the only apple in the world (try negating any of these sentences!). This explosion of potentialities stops short only of what is incompatible with a_1's existence: a_1 cannot be such that there are no apples, or such that it, a_1, does not exist.

For convenience, let me label the conception that motivated the first answer the 'strong' conception of potentiality, and the conception that motivated the second answer the 'weak' conception. Which of the two answers to my initial question should we adopt? The fact that the second, weak, conception goes with usage of 'can' may not, of course, count as evidence for the correctness of that conception; that 'can' always ascribes potentialities is a thesis still to be established. It merely makes the weak conception desirable for the purposes of such a semantics. There is a second reason to find that conception desirable.

If we go with first, the less inclusive suggestion and the strong conception, then potentiality is certain to create hyperintensional contexts. Given the existence of both a_1 and a_2, being such that a_1 is red or not red, and being such that a_2 is red or not red are logically equivalent: whatever has one of these properties must have the other one too, as a matter of logic, since both are tautologous. Now, a_1 is red or not red; it makes its contribution to that state of affairs by being not-red (let us suppose); its being red or not red should, then, count as a manifestation of its potentiality to be red or not red. So a_1 has a potentiality to be such that it, a_1, is red or not red; yet it has no potentiality to be, because it has no way of contributing to being, such that a_2 is red or not red. With a_2, matters are directly reversed. Hence both a_1 and a_2 will have one and lack the other of a pair of potentialities whose manifestations are logically equivalent. On the more inclusive 'weak' understanding, however, there is no reason to deny to a_1 the potentiality to be such that a_2 is red or not red, and in general no prima facie reason to doubt that potentiality is congruent, i.e., closed under logical equivalence. Congruence is not quite the same as intensionality, i.e., closure under *necessary* equivalence; but given congruence, it may be hoped that potentiality on the weak conception would even turn out to be intensional (as opposed to hyperintensional). For the purposes of logic, congruence is already a great advantage. (More on this below in section 4.7.)

Other formal aspects of potentiality will also differ depending on whether the strong or the weak conception is adopted: on the weak conception, whenever an object has a given property, that property is a manifestation of a potentiality of that same object; if x is F, it follows that x *can* be F. (Try saying: 'x is F, but x cannot be F'!) On the strong conception, that is not so: a_1's being red is a manifestation of its potentiality to be red, but a_1's being such that a_2 is red is no manifestation of any potentiality of a_1's. (More on this entailment below: section 4.9.)

None of the considerations so far have provided *evidence* that the weak conception is the metaphysically more adequate one. However, I believe that we should indeed adopt that weak conception. The reason, in a nutshell, is that there is no sharp and non-arbitrary distinction to be drawn between what is admitted and what is excluded by the stronger, less inclusive conception. That is a problem if we believe that the stronger, and not the weaker, conception articulates the most inclusive sense we should give to the term 'potentiality', that it provides the boundaries to the potentialities that things have. For if the boundary drawn by the stronger conception is a metaphysically substantial one between being and not-being, it had better be a sharp and non-arbitrary one. If, on the other hand, we adopted the weaker conception of potentiality, we might still make sense of the somewhat fuzzy boundary between potentialities that fulfil some further condition (the 'strong' potentialities) and those that do not. Distinctions among things that exist, and properties that things have, need not be sharp; but distinctions between what exists and what does not, and between what properties things have and which they do not have, do need to be sharp.

So: *if* there is no sharp and non-arbitrary delineation between what qualifies as a potentiality on the strong conception and what does not, then we should adopt the weaker conception. To see that the antecedent of this conditional is true, let us consider how the distinction between the strong and the weak conception is to be articulated.

Note, first, that my strong conception is still very, very weak – it is not meant to capture a causal element, nor is it intended to describe 'active' powers as opposed to passive ones. The reason for this is not that I am not interested in causal elements, or that I am opposed to the active / passive distinction. There are potentialities to cause something, and in fact many of the potentialities we are ordinarily interested in may be potentialities to cause something or be caused to do something. The active / passive distinction is a puzzling one, but probably one that should play a role in understanding the special kind of potentiality that we call ability. However, in this chapter I am concerned to stretch the notion of potentiality to its limits; to find the most general and inclusive conception of potentiality that can still count as a conception of potentiality. The question then is, how far can we stretch that notion? And there are two candidate answers, which I labeled the weak and the strong conception. The weak conception requires only that an object be *compatible* with the manifestations of its potentialities; the strong conception requires that it *contribute* to these manifestations. (We might think of them, and I will occasionally refer to them, as two senses of the term 'potentiality'. My question is not, of course, what the real sense of that word is – I have deliberately chosen a term of art with few linguistic intuitions behind it. Rather, the question is which sense we should give to the word.)

I have been using the somewhat vague term 'contribute' to capture that which must be present in a potentiality's manifestation on the strong conception, but not on the weak one. On the strong conception, an object *x* has a potentiality to be F just in case

x has a potentiality to contribute to being F. We might try to capture the distinction as one between the kinds of properties that a given object may have a potentiality to possess: being such that a_2 is red is a Cambridge property relative to a_1; a Cambridge property of x is a property that x possesses purely in virtue of things standing one way or another with objects entirely disjoint from x. We might think, then, that x contributes to its being F just in case being F is not a Cambridge property of x; and hence that x has a potentiality, on the strong conception, to be F iff x has a potentiality to possess the property of being F as a non-Cambridge property.[1]

Of course, it is not entirely clear where the line is to be drawn between the (mere) Cambridge properties and the non-Cambridge properties; or between a property to which an object contributes, and one to which it does not. For exactly the same reasons, it is not clear where the line is to be drawn between the potentialities that qualify in the stronger sense and the putative potentialities that do not. Consider the following examples.

I have argued that things have potentialities for 'tautologous' manifestations: a_1, for instance, has a potentiality to be red or not red, a potentiality which it exercises (let us suppose) by being not-red. Being red-or-not-red is not a Cambridge property of a_1; it is a property to which a_1 makes its contribution by being not-red. I have also argued that things may have potentialities for complex manifestations part of which concerns objects wholly disjoint from them: a_1, for instance, has a potentiality to be red simultaneously with a_2, i.e., to be red while a_2 is red. If a_1 were to manifest that

[1] It is important that 'as a non-Cambridge property' is in the scope of the potentiality operator. Otherwise we might rephrase the condition as: x has a potentiality, on the strong conception, to be F just in case x has a potentiality to be F *and* being F is not a mere Cambridge property (relative to x). The problem with this formulation is that the same property may be possessed by the same object as either a Cambridge property or a non-Cambridge property. a_1 possesses the property of *being such that something is red* as a Cambridge property if it is not itself red, and as a non-Cambridge property if it, a_1, is itself red. As a consequence, the less inclusive conception must say that a_1 has the potentiality to be such that something is red if a_1 itself has a potentiality to be red, but not otherwise. There is no general answer to the question: is *being such that something is red* a Cambridge property relative to a_1?

potentiality and possess the property of being red-while-a_2-is-red, a_1 would possess that property not as a mere Cambridge property; it would contribute to its possession by being red. Of course, by being red, a_1 would also contribute to being red-or-not-red (a property to which it is now contributing by being not-red). And if a_1 contributed, by being red, to being red-or-not-red as well as to being red-while-a_2-was-red, then how should a_1 fail to contribute to being red or not red while a_2 is red (the property that is expressed by $\lambda x.((Rx \vee \neg Rx) \wedge Ra_2)$, R standing for 'is red')? Hence, should we not ascribe to a_1 a potentiality to be red or not red while a_2 is red – even on the strong conception of potentiality? And finally, is there really a deep distinction between that property and the property of being such that a_2 is red ($\lambda x.Ra_2$)?

Here is another route to the same conclusion. Given that objects have potentialities for certain tautologous manifestations (such as being red or not red), it seems plausible that a_1 has the potentiality to be self-identical; being self-identical surely is not a Cambridge property. And if a_1 has the potentiality to be red while a_2 is red, then certainly a_1 has the potentiality to be self-identical while a_2 is red (the property we might express by $\lambda x.(x = x \wedge Ra_2)$). Is that property a Cambridge property for a_1, or does a_1 contribute to possessing it? Its contribution to being self-identical while a_2 is red is minimal indeed; it is the same as its contribution to being self-identical: simply existing is all that a_1 needs to do to contribute to possessing that property. Now, *that* contribution is made by a_1 even to its possessing the property of being such that a_2 is red ($\lambda x.Ra_2$); a_2 is certainly making the greater contribution here, but it does not quite do all the work, for a_1 has to exist in order to possess the property of being such that a_2 is red. Existing is the minimal contribution that an object makes to every 'such-that' property.

Where, then, is the line to be drawn – at which point should we (according to the stronger conception) stop ascribing potentialities to a_1? Is the line between *being red while* a_2 *is red* and *being red or not red while* a_2 *is red*? Or between *being red or not red* and *being self-identical*? Or perhaps between *being self-identical while* a_2 *is red* and *being such that* a_2 *is red*? None of these lines seems very natural, a 'joint of nature'. And if we cannot draw a line at which we should stop ascribing potentialities, then it would seem that we should go all the way and ascribe to a_1 the potentiality to be such that a_2 is red. In other words, we should adopt the weaker conception of potentiality.

For what it is worth, I did not initially like this conclusion very much. But I can offer some consolation to myself and the equally concerned reader. The 'extra' potentialities that are recognized by the weak but not the stronger conception of potentiality come at no extra cost: they arise naturally from potentialities that are perfectly acceptable even on the strong conception – or more precisely, on any version of it (corresponding to the various ways of drawing the boundary between weak and strong potentiality, suggested in the previous paragraph). To see how they do, we need to take a closer look at certain features of potentiality that are acceptable even to those who favour a stronger conception. These features will enable us to start with a conception of potentiality on which a thing may have a potentiality only to be a certain (non-Cambridge) way, and get all the way to things having potentialities whose manifestations concern objects wholly distinct from themselves. For those less worried about the weaker conception of potentiality, I hope those features are interesting in their own right. I will return to the issue of the weaker and the stronger conception in section 4.6. In the meantime, any unqualified use of the term 'potentiality' is to be understood in the stronger sense; the fact that the stronger sense has not been clearly delineated (or that there are various versions of 'the' stronger conception) will cause no trouble, for I am going to appeal only to examples that are uncontroversially acceptable to the proponent of that strong sense (and avoid any of the problematic borderline cases).

4.3 Joint Potentialities

Individual things have potentialities; so do pluralities of things. I have the ability to see; you and I have the ability to see each other. Each person in a crowd has some potentiality to panic and run; together, the people in a crowd have a disposition to stampede. The glass has a potentiality to break, and it has that potentiality to a degree high enough to qualify as fragile in most contexts. Wrap the glass in styrofoam; *it* is still fragile, it has the potentiality to break to the same relatively high degree as it did before. What about the glass *together with* the styrofoam? They have (among others) the following potentialities: a potentiality to break, a potentiality to be such that the glass breaks, and a potentiality to be such that the styrofoam breaks, but not one of these has an equally high degree as the glass's potentiality to break.

We must be careful to draw distinctions here.

First, when I speak of the glass and styrofoam together, I genuinely mean to treat them as a plurality, not as one composite object. If you believe in unrestricted mereological composition, you will want to say that the glass and the styrofoam together make up another object, which has them as proper parts. That composite object then has its own potentialities that are in some way constituted by those of its parts and relations between the parts. The same is true of more mundane composite objects such as glasses and people: my potentialities are partly constituted in some way or other by those of my bones, muscles, organs, etc.

I do not object to unrestricted mereological composition, but I do not want to presuppose its truth. So rather than speaking of a composite object of which the glass is a *part*, I wish to speak of *objects* that the glass is *one of*. The potentialities that a plurality of objects possess together are very much like those potentialities that would be possessed by the composite object that the objects would make up if they did make up one object – indeed I conjecture that they are the same, once we replace plurals by singulars, and part-of relations by one-of relations. I will use the analogy between the potentialities of pluralities and of composite objects below. Both are to be distinguished from the potentialities of the individual objects taken in separation from each other.

This leads me to the second set of clarifications. The glass and styrofoam together share one potentiality with the glass on its own: the potentiality to break. Of course, the state of one thing (the glass) possessing that potentiality is quite different from the state of other things (the glass and styrofoam together) possessing the same potentiality. Equally, the manifestation of that potentiality as possessed by one thing typically takes a different form than the manifestation of the potentiality by other things: if the styrofoam breaks, that will be a manifestation of the plurality's, but not the glass's, potentiality to break.[2]

Besides possessing the same potentiality that would be manifested in different circumstances by each of them, the glass on the one hand and the glass-and-styrofoam on the other also possess different potentialities that would be manifested in the exact same events: the glass's potentiality to break and the glass and styrofoam's potentiality to be such that the glass breaks. The possession of different properties by different things may nonetheless make for the same state of affairs; for instance, it may be thought that

[2] I bracket the question of whether there really is one property, breaking, that can be possessed both by individual objects and pluralities.

a's being to the left of *b* and *b*'s being to the right of *a* are the same state of affairs, differing merely in mode of presentation. (I do not wish to take a stand on the question whether it indeed is the same state; I merely wish to provide an illustration.) Similarly, it might be thought that the glass's having the potentiality to break and the glass and styrofoam together having the potentiality to be such that the glass breaks amount to the same state of affairs. But whatever we may want to say about the identity conditions for states of affairs (and again, I do not wish to endorse any particular view on this), the case just mentioned certainly is not one of identity. The manifestation of the potentiality, as possessed by either object(s), would surely be the same event. But the glass on the one hand, and the glass and styrofoam on the other, possess their respective potentialities to rather different degrees. The glass possesses its potentiality to break to such a high degree that we are prepared to call it 'fragile'. The glass and styrofoam together, on the other hand, possess the potentiality for the glass to break to a rather low degree. The reason for this difference in degree is, of course, that the glass and styrofoam together consist, precisely, of the glass *and* the styrofoam; and while the glass is disposed to break, the styrofoam has various potentialities that we may summarily say amount to a disposition to protect the glass from breaking (more about which in a moment).

In the previous section, I have distinguished between a stronger and a weaker conception of potentiality, and announced that I would concentrate on uncontroversial examples of the former until further notice. On the stronger conception, an object can only have potentialities for *itself* to be some way or other. In ascribing to the plurality of glass and styrofoam a potentiality to be such that the glass breaks, am I not tacitly relying on the weaker conception?

I am not: after all, the glass is *one of* of the glass and styrofoam. It is not at all surprising that a plurality such as the glass and the styrofoam together should possess potentialities of its own, and that the manifestations of some of these potentialities should concern only some of the things that make up the plurality. Here the analogy with composite objects is helpful; just as composite objects have potentialities whose manifestations concern only some of their parts, so pluralities of objects have potentialities whose manifestations concern only some of those objects. I have a potentiality to have long hair, a potentiality to think of Oxford, and a potentiality to scratch my nose. The manifestations of these potentialities will concern my hair, my brain, and my hand and nose, respectively; not one of them concerns, say, my right foot. Yet we do not ascribe these potentialities to me minus my right foot (and other equally unconcerned body parts). We ascribe them to me. In fact, if we did ascribe them merely to the parts of me that are concerned in their manifestations, we should at the very least think that they were possessed to a rather different degree. My hair, on its own, has the potentiality to be long to a rather lower degree than the degree to which I have the potentiality for my hair to be long; for my hair, on its own, might be detached from my body which provides it with the biological prerequisities for its growth. The case is very much parallel to that of the glass and styrofoam, except that in this instance the complex system (my body) has the potentiality for its part (my hair) to be a certain way to a higher degree than the part itself has the potentiality to be that way. That, again, should not come as a surprise; difference in degree can go both ways. (And in both cases we can find examples which go the other way around: thus the glass and styrofoam together have the potentiality for the glass to withstand certain breaking forces to a higher degree than the glass on its own has the potentiality to withstand those forces;

my hair on its own has the potentiality to remain exactly as long as it is to a higher degree than I have the potentiality for my hair to remain exactly as long as it is.)

In a composite object such as a human being, as well as in a plurality of objects such as the glass and the styrofoam, the potentialities of the whole or the plurality are in some way contributed to by the intrinsic potentialities of each of the objects constituting it. These potentialities may 'add' to one another to yield a higher degree of the potentiality for the whole object or the plurality to be such that a part or one of them is some way; or they may 'detract' from each other, yielding a lower degree for the corresponding potentiality of the whole or plurality. Thus when glass and styrofoam are put together in the right way, the styforoam's potentiality to absorb certain forces detracts from the glass's potentiality to break, thus yielding a lower degree for the jointly possessed potentiality of the glass and styrofoam to be such that the glass breaks; that same potentiality of the styrofoam adds to the glass's potentiality to withstand certain forces, thus yielding a higher degree for their jointly possessed potentiality to be such that the glass withstands those forces. In a case described earlier, in chapter 3.3.3, our traumatized piano player's ability to play the piano was detracted from by her disposition to fall into paralysis when trying to play the piano. (Identifying the parts of her which possess these two potentialities would be a matter of neuroscience.) The potentialities of a complex object or a plurality of objects are much like resultant forces, the potentialities of their parts like the component forces.[3]

Exactly how do the 'component' potentialities of the individual objects contribute to the 'resultant' potentialities jointly possessed by these objects? I wish I had an answer to that question, but I will have to leave it for another time. For the time

[3]The comparison with forces is intended to be no more than an analogy; I do not know enough about forces to make any general claims concerning their relation to potentialities. For a more ambitious use of the analogy, see Mumford and Anjum (2010), who represent the 'component' potentialities as vectors much like component forces.

being, I would merely like to make the claim that the potentialities of the individual objects do in some way contribute to those which they possess together, and that the potentialities jointly possessed by any objects are determined, at least partly, by those of the individual objects.

I have insisted on treating the glass and the styrofoam not as one composite object, but as a genuine plurality. However, it may be suspected that my argument has rested on our tacitly taking them to be something very much like a composite object, or at least a (mildly) complex system of objects; for what makes it natural to ascribe to them joint potentialities is, primarily, their spatial contiguity. What if we unpack the glass, and dissolve that contiguity? Can they nonetheless still have potentialities together? Well, to begin with, they have a potentiality to be in spatial contiguity. They also have a potentiality to be exactly 500 miles apart. Their joint potentiality to be such that the glass breaks has a higher degree than it had when they were in close contiguity; the styrofoam's potentiality to absorb certain forces is now no longer detracting from the glass's potentiality to break, or at least not as much as it used to. Is the glass's and styrofoam's joint potentiality for the glass to break now equal to the glass's own potentiality to break? Or does the glass's and styrofoam's joint potentiality for the glass to be packed in the styrofoam still detract somewhat from that potentiality? These are tricky questions, which I will not be able to answer here. All I want to show for the moment is that these questions make sense; and for them to make sense, it must be possible to ascribe the potentialities in question to the glass and styrofoam jointly even when they no longer build a spatially contiguous system, even if they are at opposite ends of the universe.

Once we have recognized this, examples of potentialities that are possessed jointly by objects with no current spatial contiguity abound: the objects on my desk jointly have a potentiality to be in disarray, the Eiffel tower and I have a potentiality to stand in the *seeing* relation, all the trees on earth jointly have a potentiality to blossom simultaneously, humankind – that is, all human beings – jointly have the potentiality to make the planet uninhabitable. There is no reason to stop short of the totality of all objects whatsoever: all the objects that exist jointly have a potentiality to co-exist with one another (as well as several other potentialities).

I submit, then, that we say the following:

- *Any objects can jointly have potentialities.* It does not matter how many they are, or whether the objects are organically connected (as in the case of the parts making up a human being) or scattered randomly throughout (a portion of) the universe (as in the case of the trees on earth). I will call a potentiality possessed by a plurality of objects a *joint* potentiality.[4]
- A joint potentiality's manifestation can take various forms: it may be a relation between the objects jointly possessing it, such as *seeing*; it may be a genuine plural property of the objects jointly possessing the potentiality, such as *being in disarray*; or it may be a property that concerns only one or some of the objects, e.g., *being such that the glass breaks*.

4.4 Intrinsic and Extrinsic Potentiality

Earlier on in the contemporary debate on dispositions, it was generally thought that a disposition had to be an intrinsic property of an object. (See, e.g., Johnston 1992, Lewis 1997, Molnar 1999.) Jennifer McKitrick (2003) has successfully challenged

[4]Joint potentialities have much in common with Popperian propensities: see Popper (1959). If Popper is right that propensities are what objective chance consists in, then potentiality may form the basis not only of possibility, as I will argue in the next chapter, but also of probability. I find this prospect attractive but will not further explore it here. In section 4.10, I briefly address the question of how the degrees of potentialities relate to probabilities.

that assumption and argued that dispositions, like other properties, may be intrinsic or extrinsic. The same line of thought applies to potentiality in general. I believe, however, that having the notion of *joint* potentialities at our disposal, we can say something more informative about the relation between intrinsic and extrinsic potentialities.

McKitrick does not rely on any particular definition of intrinsicality or extrinsicality, but on the following intuitive characterizations (appeal to perfect duplicates, of course, is a reference to Langton and Lewis 1998 and related work by David Lewis):

> Intuitively, a property is intrinsic if anything that has it has it regardless of what is going on outside of itself. Perfect duplicates necessarily share intrinsic properties. Extrinsic properties, by contrast, are simply those that are not intrinsic. If a property is extrinsic, it is possible that a thing's having that property depends on what is going on outside of the thing. Perfect duplicates can differ with respect to their extrinsic properties. (McKitrick 2003, 158)

She then argues for the existence of extrinsic dispositions by examples. I will here look at two of her examples, which I believe to constitute a representative sample of the list. One is a key's power to open a particular door, another one is the disposition of vulnerability.

By the intuitive characterization that McKitrick has given, both examples are extrinsic properties, and I find no reason to quarrel with that characterization. Whether a particular key has the power to open a particular door depends not merely on the key, but on something 'outside of' the key, namely, on how things stand with the door. If the door has its lock changed or ceases to exist, then the key will lose its power to open it; an intrinsic duplicate of the key at a world without the door, or one where the door that has a different lock, will not have the power to open this door. A person's vulnerability depends on her surroundings: I may be vulnerable walking down a dark alleyway at night, but I am not (or: my intrinsic duplicate is not) strolling along High Street in the daylight. A city may cease to be vulnerable by erecting a wall that is wholly outside of its territory and hence no part of it, or: a vulnerable city may have a non-vulnerable intrinsic duplicate which has a wall just outside its (the duplicate's) territory. It would seem uncontroversial, then, that these dispositions are extrinsic.

George Molnar, in defending the claim that all dispositions (or 'powers', in his terms) are intrinsic, had earlier developed a counter-argument to any such putative counterexample (Molnar 2003, 102-110). He held that there were extrinsic dispositional *predicates*, but denied that there had to be any extrinsic dispositions or powers to correspond to them. What makes it true that a given object satisfies an extrinsic dispositional predicate is the fact that the object itself *and other objects* in its surroundings had certain intrinsic powers, which combined in certain ways. While his examples differed slightly from McKitrick's, his strategy is equally applicable to her examples. We can imagine him replying to McKitrick as follows:

"The key does not really possess a power to open this particular door (if it did, then that power would admittedly be extrinsic). Rather, the key only possesses the power to open a door with a lock of such-and-such shape. That potentiality is not at all extrinsic: its possession by the key does not require that there exists any door with such a lock, or any door at all. In addition, the door has the intrinsic property of having a lock of such-and-such shape, and thereby the intrinsic power to be opened by a key of a certain shape (a shape which is, in turn, possessed by the key). Their intrinsic properties, and in particular their intrinsic powers, combine to make it true that 'the key has the power to open this door'. No such property as a power to that effect is required."[5]

[5] See Molnar (2003, 105) for Molnar's actual treatment of the example, which I am not reproducing here because it contains terminology that is due to its context in the book and would be of no help here.

So far my imagined speech by George Molnar. I agree with much of what he says: the key and the door both have intrinsic properties, among them powers or potentialities, and those intrinsic properties combine in such a way as to make it true that the key has the power to open this door. In fact, this process of intrinsic potentialities combining with each other is precisely what I described in the previous section as giving rise to joint potentialities. The key and the door, I suggest, together have a joint potentiality for the key to open the door. That joint potentiality is intrinsic too, but it is intrinsic to a pair of objects rather than to one object; it is an intrinsic relation, rather than an intrinsic (one-place) property. In saying that there is a genuine potentiality, jointly possessed by the key and the door, I may already be disagreeing with Molnar. I am certainly disagreeing with him when I go one step further and say: in virtue of standing in this intrinsic relation to the door, the key has the extrinsic potentiality to open the door. My disagreement with Molnar, however, is not really one about powers, or potentialities: it is one about ontology. I am perfectly happy to say that all it takes for the key to have the extrinsic potentiality to open the door is that both key and door possess certain intrinsic potentialities; but I do not conclude that there is no such thing as the extrinsic potentiality possessed by the key. The disagreement is about parsimony, and about whether we should look to reduce things (or properties) 'away'. This is not a disagreement I will try to settle here.

Now for the second part of the imagined speech by George Molnar:

"As for vulnerability, here we have an implicitly extrinsic predicate. In our earlier case, the predicate made it clear which objects' intrinsic powers were supposed to make it true that the key had 'the power to open this door'. In this case, however, some of the objects whose intrinsic powers are relevant are contextually determined: the street and whatever is on and around it in the one case, and the objects in the more or less immediate surroundings of the city in the other case. Nonetheless, what makes it true in any particular case that a given object is vulnerable is the intrinsic powers of that object, plus the intrinsic properties (powers and non-powers) of a contextually selected range of other objects."

Again, I agree with everything except the reductive spirit. For x to be vulnerable is for x to have, with a contextually selected range of other objects, a joint potentiality to be such that x is hurt or harmed. If an object stands in that relation to, i.e. has such a joint potentiality together with, the contextually salient objects, then the object possesses the extrinsic potentiality that, in the given context, is denoted by 'vulnerability'.

Note, however, that the example of vulnerability is complicated by focussing, as I do, on potentialities in general and not just on dispositions. The city surrounded by a wall still has some potentiality to be harmed, as do I walking along High Street in the daylight. The surrounding objects still make a difference: they determine the *degree* of that potentiality, and thereby also whether it counts as vulnerability (which appears to have a proportion-like modal force similar to that of 'fragile'). It is unlikely that any surroundings could diminish an object's potentiality to be harmed so much that the object, relative to those surroundings, did not possess an extrinsic potentiality to be harmed to any degree. What makes the potentiality to be harmed extrinsic, if it is extrinsic, is that its degree, not its possession, depends on how things stand outside of the potentiality's possessor.

The example is further complicated by the fact that we can make sense of a potentiality to be harmed that is not extrinsic, not even in this reduced sense. Just as the glass remained fragile (i.e., retained the same high degree of its potentiality to break) while packed in styrofoam, so it may be said that my own (intrinsic) potentiality to be harmed remains the same wherever and whenever I take my walks. Do I, then, have two distinct potentialities to be harmed, one intrinsic (and varying in degree only if I undergo intrinsic change) and one extrinsic (and varying in degree depending on my surroundings)? If so, then the glass has an intrinsic potentiality to break, as well as an extrinsic one, the latter being grounded in its joint potentiality with the styrofoam. And there is no reason to stop here: the glass may have more than one extrinsic potentiality to break, grounded in the the various joint potentialities that it has with various collections of objects. That would give us reason to relativize what I said at the outset of this chapter: a potentiality is not only individuated by its manifestation, but also by the objects that its possession and degree depends on. Alternatively, we might try to distinguish these different potentialities by their manifestations. Perhaps vulnerability is not, after all, just the disposition to be harmed or hurt; perhaps it is, rather, the disposition to be harmed *in one's current circumstances*. The glass's extrinsic disposition to break would then not be a disposition to break, fullstop, but a disposition to break *while packed in the styrofoam*. I leave open the question of how exactly to deal with these cases. For the purposes of this chapter, the simpler cases – extrinsic potentialities such as the key's power to open a particular door – are all that I will need to appeal to. (And I will understand expressions of potentiality so as not to bring with them any implicit dependence on any other objects.)

I propose, then, the following picture. Objects possess intrinsic potentialities, and these intrinsic potentialities combine to yield intrinsic relations between objects: joint potentialities. In virtue of standing in those relations, the objects individually possess extrinsic potentialities.[6]

That should not come as a surprise: quite generally, objects that stand in relations to other objects are thereby endowed with extrinsic properties. By standing in the sitting-

[6] I am open to the idea that there are primitive joint potentialities, which are not grounded in the intrinsic potentialities of the objects that jointly possess them. Such primitive joint potentialities would give rise to extrinsic potentialities equally ungrounded in the object's intrinsic potentialities.

on relation to my chair, I possess the extrinsic properties of sitting on this chair, sitting on a chair, and sitting on something.

Let me close this section with a few general notes on extrinsic potentialities.

First, potentialities for extrinsic properties are not automatically extrinsic properties: the potentiality to sit on something is not an extrinsic property, although the property of sitting on something is extrinsic. Any potentiality to stand in a particular relation to a particular object, however, must be extrinsic: a_1's having the potentiality to stand in R to a_2 depends, at the very least, on a_2's existence, and must be based in a joint potentiality possessed by a_1 and a_2.

This last observation puts an (independently plausible) limit on the potentialities that things possess to stand in relations to other objects: for an object x to have a potentiality to stand in relation R to an object y, y must exist in the first place. Questions about the existence of past and future objects will then have consequences for the potentialities of things existing at a given time. Do I have a potentiality to stand in the reading-a-book-by relation to Plato? Only if Plato, a past individual, can still be considered an existent in some weak sense. (The same goes for the question whether I do stand in that relation to Plato when I read the *Phaedo*.) Did my mother, 30 years ago, have a potentiality to give birth to *me*? Only if I, then a future individual, could already be considered an existent in some sense. Even if that should not be the case, my mother would still have had a potentiality to give birth to a daughter who was F, where F is a complete description of my non-haecceitistic features; and we could describe that potentiality in retrospect, helpfully though inaccurately, as a potentiality to give birth to me. Unlike the potentiality to give birth to me, however, that potentiality would not be extrinsic, or at least not extrinsic in depending on my existence. Again, this is not a special feature of potentiality: if objects do not exist timelessly in some sense, then nothing can stand in any relation, potentiality or otherwise, to any past or future object; if objects do exist timelessly in some sense, then things can stand in relations, including the relation of jointly possessing a potentiality, to past or future objects. The same, of course, goes for what we might call (perhaps misleadingly) contingently non-existent objects. Do I have a potentiality to meet my younger sister who was never actually born, if we read that description referentially? For necessitists such as Tim Williamson (Williamson 1998, Williamson 2002, Williamson forthcoming), the answer may be yes: if there could have been something x that was my younger sister, then there is something x that could have been my younger sister, and I may well have a potentiality to meet that x. For contingentists, this line of reasoning does not hold: there is no x that could have been my younger sister, hence there is no x that could be my younger sister and that I could have a potentiality to meet. I may have a potentiality to meet someone of a certain description, where that description includes being a sister of mine and being younger than me. (I.e., I may have a potentiality to meet my younger sister, where the description 'my younger sister' is used attributively.) Again, that potentiality would not be extrinsic, or at any rate not because it depended on the existence of something that possibly was my younger sister. (For more on necessitism and contingentism, see chapters 4.7 and 5.4.4.)

My next note arises from the observation that potentialities in general, and extrinsic potentialities in particular, come in degrees. How is the degree of an extrinsic potentiality determined? I submit that it is the same as the degree of the joint potentiality that gives rise to it. Thus I possess the potentiality to see you to the same degree to which you and I together possess the potentiality to be such that I see you; I am vulnerable to the very same degree to which I and my (contextually selected) relevant surroundings together have the potentiality to be such that I am harmed; etc. Assigning degrees to potentialities beyond the standard examples, in particular, of dispositions, is a difficult matter, and one that I am unable to discuss at length here (see section 4.10). But for the time being, my proposal seems to me to be the most reasonable one.

Finally, note that, while an individual object's extrinsic potentialities are thus (I suggest) grounded in joint potentialities, there are also extrinsic joint potentialities, grounded in further (intrinsic) joint potentialities. Just as a_1 may have an extrinsic potentiality to stand in R to a_2 in virtue of having a joint potentiality with a_2 to stand in R, so a_1 and a_2 may have a joint potentiality to stand in R' (a three-place relation) to a_3, grounded in a joint potentiality that a_1 and a_2 possess, with a_3, to stand in R'.

4.5 Iterated Potentiality

Things have potentialities to possess properties. Potentialities themselves are properties. So prima facie, things should have potentialities to have potentialities. Moreover, we have intuitive examples of such potentialities to have potentialities from the paradigm examples of dispositions and abilities. Here are some.

I do not have an ability to play the violin. Nor does my desk. Unlike my desk, however, I possess the ability to learn to play the violin – the ability, that is, to acquire the ability to play the violin. That latter ability distinguishes me from my desk, insofar as playing the violin is concerned. Let me call it an *iterated ability* to play the violin; or more specifically, a twice iterated ability (one iteration being what we would otherwise call a non-iterated potentiality).

We can go further. A violin teacher possesses a particular skill: the ability to teach students how to play the violin. The manifestation of this ability consists in another individual, a student, acquiring an ability: the ability to play the piano.

There are further intuitive examples that do not involve abilities. Here is one. Whether or not colours *are* secondary qualities, i.e., dispositions to appear in a certain way to a normal observer, coloured objects certainly *have* these dispositions. Now take an apple about to ripen and turn a bright shade of red. The apple has a potentiality to become red; it thereby has a potentiality to acquire the disposition to look red to a normal observer. Or take a colouring agent, used to turn white textiles red. When that potentiality is manifested, a piece of textile becomes red and thereby acquires the disposition to look red to normal observers.

The violin teacher and the textile dye both seem to have an iterated potentiality (a potentiality to produce a potentiality to ...) the manifestation of which involves something other than the possessing objects, the teacher and the dyeing agent, themselves. Having seen this intuitive evidence that there is such a thing as iterated potentiality (and iterated potentiality for a manifestation that does not concern the potentiality's possessor), let me now turn to spelling out in some detail what these iterated potentialities involve.

We can begin by defining an n-step potentiality, as follows. An object x has a one-step potentiality for p just in case x has a potentiality to be such that p. x has an $n + 1$-step potentiality for p just in case x has a (one-step) potentiality to be such that something has (or some things have) an n-step potentiality to be such that p. x has an iterated potentiality for p just in case for some n, x has an n-step potentiality for p.

Some comments are in order.

First, the definition of n-step potentiality is inductive and can be applied to any (natural) number n. There is no finite limit to the number of times that a potentiality can be iterated.

Second, I have quantified over numbers in defining iterated potentiality. I am not, however, thereby committed to the existence of numbers (though I may be, for other reasons). The definition could be given as an infinite disjunction of n-step potentialities; quantifying over *n* is merely a matter of convenience and brevity. (For a suggestion of how to spell this out more formally, see chapter 6.5.1.)

Third, I have included potentiality simpliciter as once-iterated potentiality; but my definition explicitly includes only potentialities *to be such that p*. For the sake of having a catch-all phrase, I will adopt the following convention: when an object, *x*, has a potentiality to be F, I shall say that *x* has an iterated potentiality for *x* (itself) to be F. This is innocuous, for if *x* has a potentiality to be F, *x* thereby has a potentiality to be such that *x* is F, and vice versa. (Mutual entailment may not amount to identity either of the potentialities or of the state of affairs of *x* possessing them, but it is good enough for my purposes.) That potentiality can then be expressed, in accordance with my definition, as a one-step potentiality for *x* to be F. (Note that we can apply this little trick even when 'F' is already of the form 'such that *p*': thus *x*'s having a one-step potentiality for *p* entails *x*'s having a one-step potentiality for *x* to be such that *p*.)

Fourth, iterated potentialities are potentialities 'for *p*'; their manifestation is not a property. The reason is straightforward: the manifestation of the iterated potentiality may concern something other than the iterated potentiality's possessor itself. The above-mentioned dyeing agent does not have an iterated potentiality to appear red to normal observers; it has an iterated potentiality for a textile to appear red to normal observers. The iterated potentiality's manifestation is not, then, *appearing red to normal observers*; it is *that a textile appear red to normal observers*. It is not even the property of *being such that the textile appear red to normal observers*; for that property would still have to be possessed by the dyeing agent when the iterated potentiality is manifested. But the dyeing agent itself may have ceased to exist by the time a dyed textile appears red to anyone; in which case *it* cannot have the property of being such that the textile appears red to normal observers, yet its iterated potentiality for the textile to appear red is still being manifested.

Iterated potentiality thereby provides us with a veritable extension of the 'reach' of an object's potentialities. Even with the strong conception of (one-step) potentiality, an object *x* may have an iterated potentiality for it to be the case that *p*, where *p* is entirely about objects distinct from *x*. To alleviate any suspicion that I am thereby introducing a weak conception of potentiality through the back-door, let me provide a schematic picture of how this extension works.

Suppose *a* has an iterated potentiality for *b* to be F, and suppose that *b* is an object wholly distinct from *a*. *a* cannot simply have a potentiality, in the strong sense, to be such that *b* has a potentiality to be F; this is one of the marks of this strong sense of potentiality. But *a* may very well have a potentiality (in the strong sense) to stand in a certain relation to *b*; that is perfectly compatible with the constraints on strong potentiality. Now that relation may be the relation of having a joint potentiality together. And as we have seen, the manifestation of a joint potentiality need not concern all of the objects that jointly possess it. Hence the joint potentiality that *a* has a potentiality to have together with *b* may be one whose manifestation concerns only *b* and not *a*; in particular, it may be a potentiality to be such that *b* is F. So here we have the chain linking *a*, by iterating potentialities, to *b*'s being F: *a* has a potentiality to have, with *b*, a joint potentiality to be such that *b* is F. We can go further, of course: *a* may have a potentiality to have, with *b*, a joint potentiality to be such that *b* has a potentiality to be F; or, *a* may have a potentiality to have, with *b*, a joint potentiality to be such that *b* has a joint potentiality, with *c*, to be such that *c* is F; and so on. Introduce existential quantification in the right places, and you have an iterated potentiality according to my definition above: *a* has a potentiality to be such that some things (namely, *a* and *b*) have a potentiality to be such that *b* is F: a two-step potentiality. Or, *a* has a potentiality to be such that some things (namely, *a* and *b*) have a potentiality to be such that something (namely, *b*) has a potentiality to be such that *b* is F: a three-step potentiality. And so on.

To illustrate this schema, let me apply it to one of my earlier examples.

The textile dye has a potentiality to stand in a certain relation to a piece of textile: the relation of possessing a joint potentiality, which typically comes with the dyeing process, and whose manifestation is: being such that the piece of textile has a potentiality to appear red to normal observers. Hence the textile dye has a potentiality to be such that some things (the dye itself, together with a piece of textile) have a potentiality to be such that something (the piece of textile) has a potentiality to be such that a piece of textile (namely, itself) looks red to normal observers. Hence the dyeing agent has a three-step potentiality for a piece of textile to look red to normal observers.

This concludes my survey of iterated potentiality. In the following, I will use the term 'iterated potentiality for *p*' to include n-step potentialities, for any n including 1. (Potentialities simpliciter, or one-step potentialities, *to be F* are covered by that locution given the trick that I outlined under my third observation.) 'Potentiality' itself will continue to be used as a predicate modifier referring to one-step potentiality.

4.6 Two Notions of Potentiality Revisited

In section 4.2, I distinguished between a weaker and a stronger conception of potentiality. The weak, more inclusive conception was introduced as going with the least demanding use of the modal verb ‘can’ and with the minimal requirement that a potentiality’s manifestation be at least compatible with the existence of its bearer; the strong, less inclusive conception was introduced as involving a requirement to the effect that an object must contribute to its potentialities’ manifestation, and consequently may not be ascribed potentialities the manifestations of which do not ‘involve’ that object in any way. The example I used to highlight the two conceptions’ disagreement was the question whether one apple, a_1, should be ascribed a potentiality to be such that another apple, a_2, is red. For the weak conception, such potentialities are acceptable; for the strong conception, they are not. I then argued that there is no sharp and non-arbitrary line to be drawn between the weak and the strong conception, and that therefore we should adopt the weak conception, and I promised to show how the ‘extra’ potentialities that are admitted by the weak conception but excluded by the strong one arise, at no extra ontological cost, from potentialities that are unproblematic even on (any version of) the strong conception.[7]

In the last three sections (4.3–4.5), I have taken care to appeal only to such potentialities as are unproblematic on the strong conception. I believe that I have thereby provided the materials to fulfil my promise. Extrinsic potentialities, understood as arising from joint potentialities, and iterated potentialities both provide ways of extending the strong conception of potentiality so as to ‘reach’ manifestations that no longer concern the original possessor of the potentiality in question – manifestations, that is, which are problematic on the strong conception of potentiality but not the weaker conception. This is not to say that we can *ascribe* those potentialities on the strong conception, but we can understand where they come from. This, I hope, should make the weaker

[7] In the following, I will speak of ‘the strong conception’ as if there were exactly one; this smoothes the formulation and will not do any harm.

conception more palatable; we have seen independent reasons to adopt that conception in section 4.2.

Let us begin with extrinsicality. An object, a_1, may have an extrinsic potentiality to be such that p because a_1 together with another object, a_2, jointly has a potentiality to be such that p. Joint potentialities may concern only some of the objects that jointly possess them: a_1 and a_2 may jointly have the potentiality to be such that a_2 is red. Does that joint potentiality endow a_1 with an extrinsic potentiality to be such that a_2 is red? Not on the strong conception. Here, then, is one way of articulating the difference between the two conceptions: on the weak conception, but not on the strong conception, every joint potentiality gives rise to an extrinsic potentiality of each of the objects that jointly possess it. That way of articulating the difference also provides a way of seeing that the disagreement is not so fundamental: the weak and the strong conception agree on the joint potentiality posessed by a_1 and a_2 to be such that a_2 is red. The weak conception is simply more liberal in allowing that joint potentiality to give rise to an extrinsic potentiality of a_1's. Generalizing from the particular case, we may conjecture that for every such 'controversial' potentiality admitted by the weak conception and not the strong conception, there is a (joint) potentiality admitted by both with the exact same manifestation property. Since the cases of disagreement concern only potentialities for 'such-that' properties, properties that are possessed by everything if possessed by anything, the uncontroversial potentialities would be manifested in exactly the same circumstances as the controversial ones. (a_1 and a_2's uncontroversial joint potentiality to be such that a_2 is red is manifested if and only if a_1's controversial potentiality to be such that a_2 is red is manifested.)

Iteration provides another useful instrument. We have seen earlier that an iterated potentiality is a potentiality 'for (it to be the case that) p', where p may or may not concern the object that has the iterated potentiality, and we have seen how this is achieved by a series of potentialities linking one object through potentialities (in the stronger sense) to others. In a way, this linking goes further than even the weak conception would allow: an iterated potentiality may be manifested without its initial possessor even existing. Existence, however, was the one minimal contribution that an object would have to make to the manifestations of its potentialities even on the weakest conception of potentiality. Iterations, then, extend the 'reach' of things' potentialities even on the weak conception of potentiality. However, iteration is also useful in showing how the controversial 'weak' potentialities correspond to some uncontroversial 'strong' potentialities. Suppose that nothing now is a sentient robot, and nothing has a potentiality to be one. Nonetheless (let us suppose) there could be sentient robots, and I could be such that there were sentient robots; my existence is fully compatible with the existence of sentient robots. There is nothing with which I have the relevant joint potentiality, for nothing has a potentiality to be a sentient robot, and hence nothing (on the strong conception) has a potentiality to be such that there are sentient robots. Nonetheless, some things may have an iterated potentiality for there to be sentient robots. Someone may, for instance, have the potentiality to build something which would then have a potentiality to be a sentient robot. Jointly with them, I may have an iterated joint potentiality *for me* to be such that there are sentient robots. (The 'for me' part is important: it carries the requirement that I must exist to exercise this potentiality.)

While there is no proof in the foregoing, there is reason for optimism. I propose the following generalization:

(W1) An object, *x*, has a potentiality to be F on the weak conception just in case some things, of which *x* is one, have an iterated potentiality for *x* to be F on the strong conception.

(W1) has the welcome consequence that having a strong potentiality to be F entails having a weak potentiality to be F (the potentialities admitted by the strong conception are a subset of those admitted by the weak conception). We merely need to include *x* itself as a trivial case of 'things of which *x* is one', and (once-iterated) potentiality simpliciter, as suggested above, as a case of iterated potentiality. Obviously, (W1) also holds when '*x*' is replaced by a plural term (and other singular expressions in (W1) are replaced by suitable plural ones).

It is easy to see how my proposed schema works for cases such as that of my two apples: whenever *x* has a potentiality on the weak conception, but not the strong one, to be such that *y* is F, *x* and *y* will jointly have a potentiality, on both the strong and the weak conception, to be such that *y* is F. (The same holds when '*x*' and '*y*' are replaced by plurals.)

However, those examples are not the only controversial potentialities. On the weak conception of potentiality, our apple a_1 not only has a potentiality to be such that a_2 is red; it also has a potentiality to be such that all cucumbers are green, that there is a red cucumber, or that there are no cucumbers at all. (All of these are compatible with the existence of a_1, and it is not true that the apple, a_1, *cannot* be such that cucumbers are green, etc.) In general, we need to consider weak potentialities to be such that *p*, where *p* is a quantified sentence rather than a simple predication. Let me discuss the two obvious cases, those where *p* is existentially or universally quantified; other quantifiers ('most', 'more than three', etc.) may be expected to behave in similar ways.

Suppose, first, that a_1 has a weak potentiality to be such that *something* is F. (To make things simpler, let F not itself be a 'such-that' property.) There are various ways in which this might be true. a_1 might simply have a potentiality to be F. That potentiality would qualify on the strong conception, and so, accordingly, would a_1's potentiality to be such that something is F. Or a_1 might merely have an iterated potentiality for itself to be such that something was F: a piano teacher, for instance, has an iterated potentiality for herself to be such that something (her student) plays the piano. Further, a_1 might be one of several objects, $a_1, ..., a_n$, which jointly have an iterated potentiality for a_1 to be such that something is F. That in turn might be true because one of them, say a_2, had a potentiality to be F, which was preserved by the plurality. Such was the case of the two apples: a_2 has a potentiality to be red, a_1 and a_2 together have a joint potentiality to be such that a_2 is red, hence to be such that something is red. Of course, a_2's potentiality may be multiply iterated: a_2 may be a textile dye that has an iterated potentiality for something to be red, which is preserved by the plurality including a_1, hence a_1 and the textile dye together have a joint iterated potentiality for a_1 to be such that something is red.

In all of these cases, a_1 has a weak potentiality to be such that something is F because something, a_1 itself or something else, has an iterated potentiality for something to be F on the strong conception, and that potentiality is preserved in a joint potentiality with a_1. What if one of these conditions failed? It might be that nothing had an iterated potentiality for something to be F. Could it not still be possible that something was F, and that a_1 existed at the same time? Then a_1 would be compatible with being, and hence should have a weak potentiality to be, such that something is F. In chapter 5, I will defend the thesis that such cases do not arise: for every possibility that p, there is something that has (or some things that have) an iterated potentiality for p to be the case. What, then, if the second condition failed: if something had an iterated potentiality for p, but that iterated potentiality was not 'preserved' into a joint potentiality with a_1? Why should that be the case? Perhaps a_1's very existence is incompatible with anything's being F. But then we should not ascribe to a_1 even a weak potentiality to be such that something is F. So there is good reason to believe that, when (W1)'s right-hand side does not hold, its left-hand side does not hold either.

Suppose, second, that a_1 has a weak potentiality to be such that *everything* is F. There are fewer ways in which this might be true. a_1 itself may have an (iterated) potentiality (for itself) to be F. Suppose this potentiality were manifested; and suppose further that the rest of the world complied, so that all other things were also F. Then a_1 would be such that everything was F, and a_1 would have *contributed* to being such that everything was F, by being F. Thus a_1 would have manifested, hence must have possessed, a potentiality to be such that everything is F *even on the strong conception*. What if a_1 has no potentiality, even iterated, (for itself) to be F? Then a_1 could not contribute to being such that everything was F, hence could not have a strong potentiality to be such that everything was F. But nor would it be compatible with being such that everything was F: for if a_1 has no potentiality, even iterated, (for itself) to be F, then for everything to be F, a_1 would have to cease to exist. Hence it appears that an object will have a weak potentiality to be such that everything is F just in case that same object has a strong potentiality to be such that everything is F. Again, (W1) is confirmed.

In fact, the considerations adduced so far suggest that the following principle (which will be useful in chapter 5) holds:

(W2) Something has, or some things have, an iterated potentiality for p on the weak conception *if and only if* something has, or some things have, an iterated potentiality for p on the strong conception.

The right-to-left reading of (W2) follows from (W1). The left-to-right reading does not. But it is supported by generalizing the considerations of the last few paragraphs. In all the examples I have discussed, the relevant iterated potentiality is jointly possessed by objects (of which x is one) *because* (some of) those objects, first of all, have an iterated potentiality on the strong conception for p; if the proposition that p does not involve x in the way required by the strong conception, the objects that jointly have the iterated potentiality are precisely those that are involved in p, or have an iterated potentiality for p (still on the strong conception). Those things, then, have an iterated potentiality in the strong sense for p, not only for x to be such that p. The same goes for every step in an iterated potentiality. It seems safe to conclude that (W2) holds in both directions.

I propose the following. The notion of potentiality can be stretched from its intuitive beginnings, dispositions and abilities, all the way to potentiality in the weak sense that I have outlined. The distinction between potentiality in the weak sense and potentiality in the strong sense is not a sharp one, though it can be applied without problems to all but a few problematic cases. However, the controversial potentialities, those which qualify as an object's potentialities by the standards of the weak but not the strong conception, are based in joint, iterated potentialities that do qualify (as joint and iterated potentialities) by the standards of the strong conception.

If the Molnarian hypothesis proposed in section 4.4 is correct, then we can obtain all the potentialities, in the weak sense, from intrinsic strong potentialities: those intrinsic potentialities combine to yield joint potentialities, joint potentialities endow objects with extrinsic potentialities, among which are the controversial potentialities. (If there are primitive joint potentialities, then they too will make their contribution.) If that is true, then a subset of the strong potentialities are the ontological 'base' in which the entire spectrum of potentialities is grounded: the totality of the intrinsic (and primitively joint, if such there are) strong potentialities, on this hypothesis, determines which potentialities there are altogether.

The controversial potentialities are in many ways rather boring potentialities. Objects possess them merely in virtue of having iterated joint potentialities with other objects, to which they do not contribute much; and if we wanted to know what potentialities (i.e., potentialities for which manifestations) the world contained, we would get the same result whether or not we count the controversial potentialities. (If my hypothesis about the degrees of extrinsic potentialities at the end of chapter 4.4 was correct, then we would get the same result not only if we were to count the potentialities, but even if we were to measure their degrees.) My proposal, then, is that no harm is done by admitting them, while some good is done: we avoid drawing arbitrary boundaries.

With the weak conception of potentiality established as the best candidate, we can now have a closer look at some of its formal features; I have said earlier that those formal features are one advantage that the weak conception has over the stronger conception. An understanding of those features will enable us to answer a question that I left open in chapter 3.2: the question whether potentiality is the dual of essence. It will turn out that it is not. Since the weak conception of potentiality has been introduced as the conception of potentiality that goes best with a potentiality semantics of ‘can’, and since that conception has now been defended on independent grounds, we may in the following tentatively appeal to intuitions concerning ‘can’ when spelling out the weak conception, without being led into circularity.

4.7 Closure

4.7.1 Closure under Logical Implication and Logical Equivalence

Potentiality, in the weak sense that I have proposed to adopt, is closed under logical implication:

(C1) If x has a potentiality to be F, and x's being F logically implies x's being G, then x has a potentiality to be G.

If an object has a potentiality to be F and being F, as a matter of logic, ensures being G, then the object should have a potentiality to be G. If it did not, then it could not be G; since by being F, it would have to be G, it could not be F either, contrary to the initial assumption.

Note that if we read 'potentiality' in (C1) in the strong sense, then this little reductio fails. Here is a counterexample.

a_1 has a potentiality to be red while a_2 is also red. (In other words, a_1 has a potentiality to stand to a_2 in the following relation: being red simultaneously.) Being red while a_2 is red entails being such that a_2 is red; a_1 has a potentiality in the strong sense to be red while a_2 is red, but not to be such that a_2 is red; hence (C1) fails for potentiality in the strong sense.

(To put the point semi-formally: '$\lambda x.$(Red(x) while Red(a_2))' behaves like '$\lambda x.$ (Red(x) $\wedge$ Red(a_2))': an object's satisfying that predicate logically implies the object's also satisfying both '$\lambda x.$Red(x)' and '$\lambda x.$(Red(a_2))'. a_1 has a strong potentiality to be $\lambda x.$(Red(x) while Red(a_2)), but not to be $\lambda x.$(Red(a_2)); hence (C1) fails for potentiality in the strong sense.)

No worry of this kind may be raised if we read 'potentiality' in the weak sense. The counterexample to (C1) for potentiality in the strong sense traded on the fact that some properties are not suitable for a potentiality in the strong sense, possessed by a particular object: being such that a_2 is red, for instance, is not suitable for a potentiality in the strong sense possessed by a_1. No such restrictions apply to potentiality in the weak sense, so no similar counterexample can be construed for it. There would, then, be no reason not to rely on the little argument given for (C1).

Moreover, if we treat the weak sense of potentiality as derived from the strong sense, we can also explain how the counterexample is avoided. a_1's potentiality to be red while a_2 is red is an extrinsic potentiality, just as a key's potentiality to open a particular door. Extrinsic potentialities, I have suggested, are grounded in joint potentialities: in this case, unsurprisingly, the joint potentiality of a_1 and a_2 to stand in the relation of being red together or simultaneously. Jointly, of course, a_1 and a_2 have both the potentiality (read in the strong sense) to be such that a_1 is red, and the potentiality (strong again) to be such that a_2 is red. But that is all it takes for a_1 to have a weak potentiality to be such that a_2 is red, according to (W1) in section 4.6. Hence we should not be surprised that the inference in accordance with (C1) goes through for potentiality in the weak sense, even it if does not for potentiality in the strong sense.

Should we then conclude that (C1) holds, without restriction, for potentiality in the weak sense? There is one worry that may be raised about such an unrestricted closure condition. Here it is.

To be sure, things can have potentialities, in this weak sense, for properties that involve objects wholly distinct from them, and indeed for properties that involve *only* things wholly distinct from them. But might there not be a limit as to which objects may be involved in a property, if a given thing, say a_1, has a potentiality to have that property? It may be thought that there is such a limit if there could be incompatible objects – or, to put the idea less paradoxically, if it is the case that there could be an object *x* and there could be an object *y* such that it could not be that there were both *x* and y. Take a given sperm *s* and two eggs, *e* and *e**. s could combine with *e* to form a human being, call her *h*; or *s* could combine with *e** to form another human being, call her *h**; but *s* cannot do both. If *s* combines with *e*, then *h* comes into existence but *h** does not; and mutatis mutandis for *e**. (The example is from Williamson forthcoming, who also provides further examples.) Suppose *h* comes into existence and has a potentiality, say, to dance. '*x* dances' logically implies '*x* dances and is such that *h** dances or *h** does not dance'. But given the incompossibility of *h* and *h**, how could *h* have any potentialities concerning *h**? That potentiality could never manifest; for it to manifest, it would have to be true that *h** is dancing or not dancing, hence that *h** exists; but if *h** existed, then *h* could not exist, and therefore could not manifest its potentiality.

The worry is mistaken. Since *h* exists, it is not the case that *h** exists. Or to be more precise: whether *h** exists depends on whether necessitism is true, where necessitism is the view that 'what there is is a wholly necessary matter' (Williamson forthcoming, 9). It is contingent, on any view, whether *s* combines with *e* to form *h*, or with *e** to form *h**. On the necessitist view, that just means that it is contingent whether *h* or *h** (or neither of them) becomes concrete and a bearer of non-trivial non-modal properties. On the opposed view, contingentism, 'what there is is partly a contingent matter' (Williamson forthcoming, 9). Contingentism does not entail but is most plausibly combined with the view that as a matter of contingent fact, there is no such thing as *h**, even in the weakest sense of 'there is' and 'thing'.

What we say about the worrisome entailments depends on whether we are contingentists or necessitists. It also depends on how we conceive of the tautological conjunct's logical form. To use a convenient bit of formalism, we may think (1) that an object's satisfying the predicate $\lambda x.\text{Dances}(x)$ logically implies the object's satisfying the predicate $\lambda x.(\text{Dances}(x) \land (\text{Dances}(h^*) \lor \neg \text{Dances}(h^*))$. Or we may have in mind (2) that the predicate $\lambda x.\text{Dances}(x)$ logically implies the predicate $\lambda x.(\text{Dances}(x) \land \lambda y.(\text{Dances}(y) \lor \neg \text{Dances}(y))(h^*))$. On the first understanding, dancing entails dancing and being such that it is either the case or not the case that h^* is dancing; on the second, dancing entails dancing and being such that h^* has the tautologous property of dancing-or-not-dancing.

On the contingentist view, 'h^*' is an empty name. There are different ways of dealing with empty names. One is to say that sentences containing them have no truth-value at all; on that view, neither of the inferences in (1) or (2) will go through, hence (C1) does not apply to the case. Another option is to say that when h^* is an empty name and Φ any predicate, $\Phi(h^*)$ will always be false but $\neg\Phi(h^*)$ true. In that case the inference in (1) goes through but is harmless; the tautological conjunct will always be true simply because its negated disjunct $\neg$ Dances(h^*) is true. *h*'s manifesting the entailed potentiality does not require any more than *h*'s dancing and there not being any such thing as h^*, hence *h* may safely be attributed the potentiality that is entailed in accordance with (C1). The inference in (2) will not go through because the second conjunct of the predicate is in fact no tautology; it is of the form $\Phi(h^*)$ and therefore false.

On the necessitist view, both inferences (1) and (2) will go through, but both will be harmless. The incompossibility of *h* and h^*, on the necessitist view, is not an incompossibility of their existence in the most general sense of the term. It is an incompossibility of their being concrete, or of their bearing properties other than logical ones (including the tautologous property of dancing-or-not-dancing). (1): *h* may be dancing and such that it is the case or it is not the case that h^* is dancing, by dancing and being such that it is not the case that *h** is dancing. (2): *h* may, further, be dancing and such that *h** is dancing-or-not-dancing; all it takes for this to be the case is that *h* be dancing and *h** not be dancing (which is guaranteed by *h** not having any non-modal properties apart from logical ones). By having the potentiality to dance, *h* will have the two further potentialities entailed in accordance with (C1), but both are harmless, since both can be manifested simply by *h* dancing and *h** failing to be concrete.

I conclude that potentiality in the weak sense that I have proposed and that goes with the use of 'can' is, after all, closed under logical implication. It follows directly that potentiality in this weak sense is closed under logical equivalence, i.e. mutual logical implication: if an object's being F is logically equivalent to the object's being G, then that object's being F logically implies its being G, and hence by (C1), the object's having a potentiality to be F entails having a potentiality to be G.

(C2) If *x* has a potentiality to be F and *x*'s being F is logically equivalent to *x*'s being G, then *x* has a potentiality to be G.

Note that, again, the principle does not hold if we read 'potentiality' in a stronger sense. For a_1's being red or not red is logically equivalent to a_1's being such that a_2 is red or not red (given that a_2 exists). But on the strong conception, a_1 may have a potentiality for the tautologous manifestation *being red or not red*, yet it may not have a potentiality whose manifestation concerns only objects wholly disjoint from itself, such as *being such that* a_2 *is red or not red.* As I have noted earlier, potentiality on the stronger conception is hyperintensional, while potentiality on the weak conception is at least congruent (and perhaps intensional).

4.7.2 Consequences of Closure

From the closure of (weak) potentiality under logical implication, two interesting consequences follow immediately.

First, what an object has a potentiality for is closed under disjunction. That is, if *a* has a potentiality to be F or a potentiality to be G, then *a* has a potentiality to be F-or-G. For, if the antecedent holds, then at least one of its disjuncts must be true. Whichever it is, the manifestation of the attributed potentiality logically implies being F-or-G; hence *a* has a potentiality to be F-or-G. This holds for both weak and strong potentiality. The only reservation we had for strong potentiality's closure under logical implication was that a potentality's manifestation must not be entirely about objects that are completely disjoint from the potentiality's possessor. If that requirement is met by the property of being F, then it will also be met by the property of being F-or-G.

Second, what an object has a (weak) potentiality for distributes over conjunction: If *a* has a potentiality to be F-and-G, then *a* has a potentiality to be F, and *a* has a potentiality to be G. The reason, of course, is that being F-and-G logically implies both being F and being G. For strong potentiality, this principle does not hold unrestrictedly: on the strong conception, a_1 may have a potentiality to be red and such that a_2 is red, but a_1 will not have a potentiality (on the strong conception) to be such that a_2 is red.

What about the opposite directions: is potentiality closed under conjunction, and/or does it distribute over disjunction? (Neither of these two questions has an answer that follows from (C1).)

The answer to the first question is easily and straightforwardly, no. It is not the case that if *a* has a potentiality to be F and a potentiality to be G, then *a* has a potentiality to be F-and-G. A simple counterexample suffices to show this: replace G with not-F. Things have many potentialities whose manifestations contradict each other, without thereby having potentialities to have a contradictory property: I have a potentiality to walk and a potentiality not to walk, a potentiality to speak and a potentiality to be silent. But certainly I do not have (nor does anyone or anything else have) a potentiality to both walk and not walk, or to speak and be silent.

The second question is more interesting, and should be answered in the affirmative. To get a feel for disjunctive manifestations, we may again think of a determinable property and its determinates, e.g., being red (the determinable) and being any particular shade of red: red_1, red_2, etc. (the determinates). Take, again, our apple's, a_1's, potentiality to be red. That potentiality is intimately linked to, and at least entails, a potentiality with the following (disjunctive) manifestation property: to be red_1 or to be red_2 or ... Does it follow that a_1 has a potentiality to be red_1 or a potentiality to be red_2 or ... ? I say, it does: if a_1 manifests its potentiality to be red, it will inevitably do so by being some particular shade of red or other. Whichever shade it is – call it red_n –, a_1 then has the property of being red_n, and it has that property through a manifestation of a potentiality of its own. There is, then, no reason to deny to a_1 the potentiality to be red_n. But red_n must, by hypothesis, be one of the determinates red_1, red_2, ... Hence there must be at least one disjunct in the long disjunctive property which is entailed by being red that a_1 has a potentiality to have: a_1 must have a potentiality to be red_1 or a potentiality to be red_2 or ....

Now for the generalization. The crucial feature of the determinates / determinables relation is that to possess the determinable property, an object must possess exactly one of the determinate properties; in fact, what was crucial to my argument was merely that to possess the determinable, an object must possess *at least* one of the determinates. But that is precisely what is true of a disjunctive property. If a has a potentiality to be F-or-G, that potentiality could be manifested in no other way than by a's being F, or *a*'s being G. Hence *a* must have at least one of the corresponding potentialities: a potentiality to be F, or a potentiality to be G.

(Note that the line of reasoning also holds when we read 'potentiality' in the strong sense.)

All of these results hold for any properties, including 'such-that' properties. Since those properties will be of importance when the informal results of this section are put to a more formal use in chapter 6, we should have a quick look at them. Having a potentiality to be such that p or such that q entails and is entailed by having a potentiality to be such that p or a potentiality to be such that q. The first, disjunctive, potentiality may be a reformulated as a potentiality to be such that p or q. The reformulation does not make a difference to the property expressed but allows us to use the connective 'or' as it (or rather, the corresponding connective $\vee$) is used in classical logic: as connecting sentences rather than predicates. (The disjunctive property can then be expressed by $\lambda x.(p \vee q)$.) For future reference, here is the result:

(D) An object x has a potentiality to be such that p or q if and only if x has a potentiality to be such that p or a potentiality to be such that q.

The same reformulation can be given for conjunctive properties.

It will be noted that the formal features of potentiality discovered so far – closure under logical implication, distribution over disjunction – are those of possibility, as studied in modal logic. This fact will be used in chapter 6. First, however, I turn to the relation between potentiality and essence.

4.8 Potentiality and Finean Essence

Having studied some of the formal features of potentiality, I can now answer a question that I considered at the end of section 4.6 (and earlier on, already in chapter 3.2), and show that potentiality, even in the very weak sense that I have been examining, is not the dual of Finean essence. To do so, I must provide a brief sketch of Fine's formal treatment of essence (given in Fine 1995a and Fine 2000), which will then enable me to show that potentiality, even in the weak sense, is not straightforwardly interdefinable with it.

4.8.1 Fine on the Logic of Essence

The logic of essence, unsurprisingly, is structurally very similar to the logic of necessity. It is not, however, identical with it. (We cannot, that is, take the axioms and rules that are given for systems of normal modal logic and simply replace the necessity operator $\Box$ throughout with an essence operator $\Box_F$.) Fine notes one important difference at the outset of Fine (1995a, 242): like propositions that are necessarily true,

> the propositions true in virtue of the nature of given objects are taken to be closed under logical implication; any logical consequence of such propositions is also to be such a proposition. ... However, this closure condition is subject to a certain constraint. For we do not allow the logical consequences in question to involve objects which do not pertain to the nature of the given objects. Let us suppose, for example, that the empty set does not pertain to the nature of Socrates, then we do not allow the proposition that the empty set is self-identical to be true in virtue of the

> nature of Socrates even though this proposition is a consequence of any set of propositions whatsoever.

Despite the first sentence of the quote, then, what stands in the scope of the essence operator $\Box_F$ – read: it is true in virtue of the nature of the Fs that ... – is not closed under logical implication. Or to be more precise: it is closed under logical implication subject to one constraint: the implication must not 'involve objects which do not pertain to the nature of the given objects'.

That introductory remark can be made more precise using Fine's notion of *objectual content*. The objectual content of a sentence or predicate, intuitively, is the set of objects that the sentence or predicate is 'about'. A precise definition can be given for a formal language of the usual kind:

> Let M be a model and E a sentence or closed predicate whose constants are $a_1, ...a_m$ [...] Then the *objectual content* $[E]^M$ of E (in M) is taken to be $\{v(a_1), ...v(a_m)\}$ [...]. (Fine 2000, 548)[8]

In addition to the notion of objectual content, we need the notion of *dependence*. This notion is closely linked to that of essence; indeed, so much so that it could be defined in terms of it: x depends on y iff y is among the objectual content of some proposition that is true in virtue of the nature of x. Or, more formally speaking, let $x \leqslant y$ stand for 'x depends on y'. Let v be the valuation function that goes with a model M. Then $v(x \leqslant y)$ is true just in case there is a sentence A such that $v(\Box_x A)$ is true and $y \in [A]^M$. The Fs depend on y just in case some of the Fs depend(s) on y. — In

[8]In accordance with my own notation in chapter 6, I have replaced Fine's ϕ (for the valuation function) by v. Note that Fine himself does not limit the definition to constants. He uses so-called 'rigid predicates'. A rigid predicate is a predicate that expresses a rigid property; by 'a rigid property is meant a property of being identical to x_1 or x_2 or ..., for certain specific objects $x_1, x_2, ...$. For example, the property of being identical to Socrates or Plato ($\lambda x(x = \text{Socrates} \vee x = \text{Plato})$) and the property of being identical to a natural number ($\lambda x(x = 0 \vee x = 1 \vee ...)$) are rigid.' (Fine 1995a, 244) Then the objectual content $[E]^M$ of a formula E, where $a_1...a_m$ are E's constants and $P_1...P_n$ are E's rigid predicates, is $\{v(a_1), ...v(a_m)\} \cup \{v(P_1), ...v(P_n)\}$.

his treatment of the logic of essence, Fine does not define dependence in this way, but stipulates the connection as an axiom. For my purposes, that is of no consequence.

Now we can reformulate the restriction on closure in the scope of $\Box_F$.

If closure were unrestricted, then we would have the following principle: If it is a theorem that $A \rightarrow B$, then it is a theorem that $\Box_F A \rightarrow \Box_F B$. Fine poses a restriction on this principle. To a first approximation, the restriction goes as follows: Suppose $\Box_F A$ is true (in a given model M), and $A \rightarrow B$ is a theorem. Then it follows that $\Box_F B$ *only if* B's objectual content (relative to the model M) does not *exceed* the Fs and what they depend on. We can refine this a little: if $\Box_F A$ is true (in M), then the objectual content of A cannot exceed the Fs and what they depend on; that follows directly from the link between essence and dependence. All we need to worry about, in that case, is that part of B's objectual content that exceeds A's objectual content. If $\Box_F A$ is not true, then we don't need to worry about B's objectual content at all, because it is then trivially true that $\Box_F A \rightarrow \Box_F B$. So we get the following principle (cF stands for the 'closure of the Fs', the set of all objects that the Fs depend on, including the Fs themselves, $|A|$ is the object-language expression for the objectual content of A, all other symbols should be self-explanatory):

> [Theorem 4]
>
> (ii) If $\vdash A \rightarrow B$, then $\vdash |B| - |A| \subseteq cF \rightarrow (\Box_F A \rightarrow \Box_F B)$.
>
> (Fine 1995a, 253)

So much for the precise nature of Fine's restriction on closure. Its rationale should be intuitively clear too. Given the systematic connection between essence and dependence, $\Box_F B$ entails that the Fs depend on the objectual content of B. But that dependence cannot hold *just because* of the logical implication; the Fs must, as it were, already depend on the objects in B's objectual content; that is, there must already be some truths that have those objects in their objectual content and are true in virtue of the nature of the Fs. Logical implication by itself does not engender dependence relations; there must be some essential connection between the Fs and the objects that have something true of them in virtue of the nature of the Fs.

4.8.2 Essence and Weak Potentiality

In Fine (1995a), Kit Fine defines an operator that is the dual of his essence operator, as follows:

$$\Diamond_F A \equiv \neg\Box_F\neg A$$

Let me call the converse of Finean essence, i.e., whatever is expressed by $\Diamond_F$, 'E-potentiality' (E, of course, standing for essence). If weak potentiality (as I understand it) is the dual of Finean essence, then it is identical with E-potentiality. In this section, I will show that this is not the case.

Of course there is one difference between weak potentiality as I construe it, and E-potentiality: the former operates on properties (and is best expressed by a predicate operator), the latter operates on propositions (and is therefore expressed by a sentence operator). Given the 'such that' constructions that I have allowed in the scope of a potentiality operator, this difference is easy to get around in most instances. Another difference that may safely be ignored is the fact E-potentiality is ascribed to groups of objects using a predicate (the Fs), while I have been ascribing potentiality, weak and strong, using singular terms. Fine allows $\Box_a A$ as an abbreviation for $\Box_{\lambda x.x=a} A$, and similarly for $\Diamond_a A$.

Leaving all that aside, then, is weak potentiality (to be such that A) expressed by Fine's operator $\Diamond_F A$? No: and the reason is that weak potentiality is, while E-potentiality is not, closed under logical implication *without restrictions*.

Remember Fine's closure condition: If it is true in virtue of the nature of the Fs that A, and A logically implies B, then it is true in virtue of the nature of the Fs that B – unless B's objectual content exceeds the Fs and what they depend on. That restriction, of course, affects the closure condition for E-potentiality too. Here are the two closure conditions; I repeat the condition for $\Box_F$ for comparison (as before, cF is the 'closure of the Fs', the set of all objects that the Fs depend on, including the Fs themselves, $|A|$ is the object-language expression for the objectual content of A, the other symbols should be self-explanatory):

[Theorem 4]

(ii) If $\vdash A \to B$, then $\vdash |B| - |A| \subseteq cF \to (\Box_F A \to \Box_F B)$.

(iii) If $\vdash A \to B$, then $\vdash |A| - |B| \subseteq cF \to (\Diamond_F A \to \Diamond_F B)$.

(Fine 1995a, 253)

In Fine's words, these pose the following restrictions: 'for necessity [meaning: essence] to be preserved by a provable implication there must be no "new content" in the consequent whereas for possibility [meaning: E-potentiality] to be preserved there must be no new content in the antecedent.' (Fine 1995a, 254) I have sketched Fine's motivation for (ii) above, so let me now concentrate on (iii). The following is a rational reconstruction.

Take an object a and a property F such that a is essentially not-F: $\Box_a \neg Fa$ is true, and so is $\neg\Box_a Fa$. Take, further, an object b such that a does not depend on b. Note that $Fa \wedge (Gb \vee \neg Gb)$ logically implies Fa:

$$\vdash (Fa \wedge (Gb \vee \neg Gb)) \to Fa$$

Note, further, that $(Fa \wedge (Gb \vee \neg Gb))$ has objectual content that, by stipulation, exceeds a and what a depends on. The same is true of its negation, $\neg(Fa \wedge (Gb \vee$

$\neg Gb$)). Because of that excess objectual content, neither the original proposition nor its negation is true in virtue of the nature of a. We have:

$$\neg\Box_a\neg(Fa \wedge (Gb \vee \neg Gb),$$

which is by definition equivalent to

$$\Diamond_a(Fa \wedge (Gb \vee \neg Gb),$$

However, we also have

$$\Box_a\neg Fa,$$

and hence

$$\neg\Diamond_a Fa.$$

This then, is a case of logical implication between two propositions without logical implication between their corresponding E-potentiality propositions. If E-potentiality was (what I call) weak potentiality, this should not happen; or so I have claimed. But perhaps I was mistaken? Let me look at the example in a little more detail.

We assumed that being F was excluded by a's very nature; a, then, *cannot* be F, i.e., a does not have a weak potentiality to be F. Correspondingly, a does not have an E-potentiality to be F. However, a does have an E-potentiality to be F and such that b is G or b is not G. To fill in a concrete example, I may have an E-potentiality to be a stone and be such that you are a philosopher or you are not a philosopher. Evidently, we do not want to ascribe to me such a potentiality, in whatever weak sense of the word, and certainly not in the weak sense I have proposed: I *cannot* be a stone and such that you are a philosopher or not a philosopher, because I cannot be a stone. The fact that the second conjunct of this conjunctive property introduces an individual, you, on whom I do not depend, does not endow me with even a weak potentiality to have the conjunctive property. It does, however, endow me with an E-potentiality – which merely goes to show that we should not think of E-potentiality as potentiality at all but merely as the dual of Finean essence (Fine does no more than that). Potentiality, even in the weak sense, is not the dual of Finean essence.

4.9 Potentiality and Actuality

I now turn to a final feature of potentiality that is of some formal (but not only formal) interest. I continue to focus on the weak conception of potentiality.

It is part of the conception of potentiality that I am proposing that whenever an object has a property, having that property is exercising a potentiality of that object. So for every property that an object possesses, we must ascribe to that object a potentiality at some time or other. (If we are speaking of strong potentiality, we may have to restrict the scope of 'properties'.) To a first approximation, we may put this principle as follows:

(T1) x is F $\rightarrow \exists t$: at t, x has a potentiality to be F.

As it stands, this principle is too weak. After all, potentiality is *forward-looking*: a potentiality cannot be exercised before it is possessed. In other words, t in the consequent has to be relativized to a time-index implicit in the antecedent, to make sure we are not speaking of potentialities that are possessed later on in a thing's career. A potentiality's exercise must not lie in the past; it usually lies in the future. What about the present? Depending on how we interpret the 'forward-looking' nature of potentialities, there are two ways of answering that question. Let $<$ stand for the 'earlier than' relation between times, and $\leq$ for 'earlier than or simultaneous with'. Then we have:

(T2) x is F at $t \rightarrow \exists t' : t' < t$ and at t', x has a potentiality to be F.

(T3) x is F at $t \rightarrow \exists t' : t' \leq t$ and at t', x has a potentiality to be F.

In this section, I will first provide an argument for preferring (T3) over (T2). I will then argue that given that argument, we should indeed adopt an even stronger principle:

(T4) x is F at $t \rightarrow$ at t, x has a potentiality to be F.

Note that (T2) and (T4) each entail (T3); the consequent of (T3) is equivalent to a disjunction of the consequents of (T2) and (T4). While (T2) and (T4) are compatible, the reason for adopting (T2) instead of the weaker (T3) would clearly be the desire to exclude what (T3) explicitly allows, and (T4) requires: that a potentiality is possessed while it is exercised.

Since (T4) is the ultimate aim of my argument in this section, let me provide some reason for finding it attractive in the first place; it derives from the weak conception of potentiality and its connection with the modal verb 'can'. It would seem that 'x is F' entails 'x can be F'; or in other words, 'x is F but x cannot be F' is inconsistent. (Try saying 'the apple is red but it cannot be red'!) In other words, the following principle, with no relativization to times, should hold:

(T) If an object x has a property F, then x has a potentiality to have the property F.

But for (T) to hold, (T4) must be true: there must not be a time at which the antecedent is true but the consequent is not.

Given a choice between (T2) and (T3), why should we favour (T2)? The reason would be the same that induced us to pass on from (T1) to (T2) and (T3): potentialities are forward-looking; they concern the future, not the past. A disagreement between (T2) and (T3) may be understood as the question whether the present should be classified with the future or the past. The idea behind a preference of (T2) over (T3) is a strict interpretation of the idea that potentiality is 'forward-looking': an exercise of a potentiality, on that interpretation, must not be simultaneous with the possession of that very potentiality. Potentiality concerns the future, not the past *or* the present. It is worth noting that with (T3), I propose only a minimal revision of this: potentialities concern the future *or* the present, but not the past. While future and past are long, perhaps infinitely long, stretches of time, the present is but a single moment, the temporal equivalent of a point. It is only this point-sized bit (not even a 'stretch') of time that proponents of (T2) and (T3) disagree about. I agree that potentiality is forward-looking; with (T3), we merely include the outer limit of what lies in the 'forward' direction, as a limiting case of what lies 'forward'.

Here is what I take to be a reductio ad absurdum, not so much of (T2) itself, but of that motivation, the strict interpretation of 'forward-looking', which may be summarized thus: no potentiality is possessed and exercised at the same time; the exercise of a potentiality must always lie in the future relative to its possession.

Take an object a at t_n. Suppose that at t_n, a is red, and a remains red at least until t_{n+1}. Since a is red at t_{n+1}, we may assume that at t_n, a has a potentiality to be red. Given the strict interpretation of that potentiality's forward-looking nature, we would have to say that at t_n, a has both the property of being red, and the potentiality to be red, but without the former being a manifestation of the latter. But a's being red is the manifestation of a potentiality of a's, whether we are using 'potentiality' in the weak or the strong sense. Of which potentiality, then, is it a manifestation? Answer: of a's potentiality, at t_{n-1}, to be red. So a possesses the potentiality to be red both at t_{n-1} and at t_n, but a's being red at t_n is a manifestation of its potentiality to be red at t_{n-1}, without being a manifestation of its potentiality at t_n. This is odd. The same potentiality, possessed by the same thing, cannot both be manifested in a given event and not be manifested in that same event.

(For clarification, it may be useful to contrast the potentiality to be red with the potentiality to become red. The potentiality to become red is manifested in an object's becoming red; once the object is red, it no longer manifests the potentiality to become red. Becoming red is a process with an end-point; once the end-point is reached, the potentiality is no longer manifesting. The potentiality to be red, on the other hand, is a potentiality that is manifested in a state, the state of the object's being red. That state may come to an end but it may go on indefinitely. As long as it continues to hold, the object continues to exercise its potentiality to be red.)

The only way out of this, as far as I can see, is to say that there is not one potentiality but two: one possessed at t_{n-1} and no longer possessed, but manifested, at t_n, another possessed at t_n (and probably also at t_{n-1}) but not manifested at either t_{n-1} or t_n. We had underspecified these potentialities and made them look as if they were the same. In fact, one is a potentiality to be red *after* t_{n-1}; the other is a potentiality to be red *after* t_n. If we want to avoid a's simultaneously having and manifesting any potentialities, we must deny that a has such a potentiality as the potentiality to be red, fullstop, and claim instead that there are only such potentialities as the potentiality to be red after t_n, possessed only up to t_n, the potentiality to be red after t_{n+1}, possessed only up to t_{n+1}, etc.

That's a rather strong claim: there is no such thing as the potentiality to be red! There are various ways of construing potentiality according to this claim. One, and I think the least attractive, option is to say that there are some properties – for instance, being red – for which there are no potentialities; there are potentialities only for time-indexed properties such as *being red after* t_n. Another option is to deny that there is such a property as *being red* in the first place; there are only the time-indexed properties, hence it should not come as a surprise that there are no potentialities for anything other than the time-indexed properties. Finally, a third option is to deny that a potentiality is individuated only by its manifestation property; rather, we need a two-place operator for potentiality (but quite unlike the two-place operator for dispositions that I have argued against in chapter 2!): a potentiality is a potentiality to ... after time ... .

Note that whichever of these options we would take, the result is that there is no one potentiality which a might retain until the end of its existence, and of which its being red at t_0, at t_n, at t_{n+1}, and at various other times, are all manifestations.

The choice, then, is this: either we accept that a potentiality can be manifested and possessed simultaneously, thus perhaps giving up to a small degree on the idea that potentiality is forward-looking. Or else we cling to that idea in its strictest interpretation, and give up another intuition: that a thing's being red at various points in its life are the various manifestations of one and the same potentiality. I suggest we go for the first option, and adopt (T3).

But what about (T4)? (T4) says not only that sometimes an object's having a property is the exercise of a simultaneously possessed property; it says that this is *always* so. (At least it is implicitly intended to say that – strictly, of course, it only says that when an object possesses a property it also simultaneously possesses the potentiality for that property.) Why should that be the case?

Consider again the argument for (T3). It showed that if an object possesses a certain property F at time t_n and continues to possess F at t_{n+1}, then at t_n, the object has the potentiality to be F, and its being F at t_n must be an exercise of that potentiality.

Now consider the following proposition. Whether a given state or event e (an object possessing a property F) is an exercise of a given potentiality, P, is a matter that depends only on *e* and P; it is an *intrinsic relation* between them. Whether the object existed before the time at which *e* is taking place, whether it will continue to exist after that time, and whether it will continue to possess the property F, is irrelevant to that question. What the question depends on is merely: whether F, the property that the object possesses in *e*, is the property that P is the potentiality to possess.

I cannot provide any further argument for this proposition, but I submit that it is overwhelmingly plausible given the conception of potentiality I have been developing in this chapter. If we put this proposition together with the argument for (T3), we get the following result. Since an object's being red is sometimes a manifestation of that object's simultaneously possessed potentiality to be red – sometimes: namely, when it continues to be red afterwards – it must always be. For, insofar as the potentiality and the state or event of the object's being red are concerned, there is no relevant difference between cases where the object is red and continues to be so, and cases where the object is red but will cease to be so the moment after.

I conclude, then, that we should adopt (T4): whenever an object has a property, F, the object's possessing that property is an exercise of the object's simultaneously possessed potentiality to be F. Given (T4), we can drop relativization to times and embrace the general principle (T): having a property entails having the potentiality to have that property.

4.10 Summing Up

In this chapter, I examined various features of potentiality: potentialities may be possessed by any thing individually, or jointly by any number of objects; potentialities may be intrinsic or extrinsic; and they may be iterated. I further suggested that these features allow us to extend the notion of potentiality, from the initial and intuitive examples of dispositions and abilities, all the way to what I have called the 'weak' conception of potentiality. I investigated some formal features of the weak conception: potentiality on that conception is closed under logical implication, it distributes over disjunction, and an object's having any property whatsoever entails its having a potentiality to have that property.

Potentiality on the weak conception is wide enough, and has the right kind of formal features, to be used in a semantics of 'can'. At the end of chapter 3, I noted that two challenges remained for my proposal of a semantic theory that construes 'can' as a predicate modifier ascribing potentialities to objects. One (arising from consideration (iv) in chapter 3.3.1) was to provide an account of potentiality that is wide enough to accommodate all true sentences containing 'can', a task that abilities and dispositions alone could not accomplish. I believe that this challenge is met with the weak conception of potentiality. A second challenge (arising from consideration (ii) in chapter 3.3.1) was to match the elegance of the standard, possible worlds-based, semantic theory in accounting for the context-sensitivity of 'can'. It is beyond the scope of the present investigation to meet that challenge, but let me sketch very briefly why I believe that there is reason to hope that it can be met.

Possible-worlds semantics capture the context-sensitivity of 'can' a with contextually varying restrictions on the 'relevant' or 'accessible' worlds: 'x can F' is true iff x F's in some world, but we do not consider all worlds in all contexts. There is reason to hope that a potentiality-based semantics of 'can' would be able to do the same: 'x can F' ascribes to x a potentiality to F, but different kinds of potentiality are relevant in different contexts. In many contexts, we use 'can' only to ascribe potentialities that qualify abilities: witness 'I can play the piano, but I cannot play the violin'. We also tend to ignore – for good practical reasons – potentialities that have a very low degree: witness 'the vase can break, but the desk cannot'. Finally, we use 'can' sometimes to ascribe intrinsic potentialities, and sometimes to ascribe extrinsic ones: contrast 'the vase is wrapped so well that it cannot break' (extrinsic) and 'the vase can break easily so I wrapped it just to be sure' (intrinsic).

Before I move on, I would like to draw the reader's attention to some loose ends. One such loose end is an account of abilities. 'Ability' is the most natural noun to go with 'can', but it seems to be used more narrowly, with a view to actions and intentions. ('He can fall down the stairs' seems fine, but 'He has an ability to fall down the stairs' does not.) Whether that is a matter of the semantics of 'ability', requiring a metaphysical account of abilities, or perhaps a pragmatic aspect of 'ability', is not entirely clear, and I have no answer to the question what, if anything, distinguishes abilities from potentialities in general.

Another loose end is how best to construe the degrees of potentialities. I have made use of the fact that potentialities come in degrees. But not only does it become increasingly difficult to assign comparative degrees once we leave the realm of everyday dispositions; it is also still an open question how the structure of those degrees is best described. A model that naturally suggests itself is that of probabilities. However, that model comes with some undesirable features. One is the failure of non-zero probabilities to distribute over infinite disjunctions. My paradigm example of potentialities for disjunctive manifestations has been the case of determinables and determinates. Many determinables have an infinite number of determinates, but we should certainly not want to say that an object has a potentiality, say, to be red without having a potentiality to be any particular shade of red, even though the probability of its having any one of the infinitely many shades is zero. Another worry relates to the question of what amounts to the possession of a potentiality to the maximal degree. One suggestive answer is: a potentiality is possessed to the highest possible degree when manifested. Another is: a potentiality to F is possessed to the maximal degree by x iff x has no potentiality not to F; in other words, if x must F. The second answer is the one that would be prescribed by the probabilistic model: on that model, the degrees of the potentiality to F and the potentiality not to F must add up to the maximal degree. But it takes potentiality closer towards necessity than we might wish.[9] I leave the question of how to construe degrees of potentiality open. Much as I would like to have an answer to that question, I will now turn to an issue that does not require such an answer: an account of metaphysical possibility in terms of potentiality.

[9] On the possible worlds view of possibility, there are two ways of understanding degrees of possibility: closeness and proportion. It might be useful to formulate a corresponding distinction for potentiality. But I have not yet been able to do that.

Chapter 5

Possibility

5.1 The Basic Idea

In chapter 3, I have briefly motivated the idea (parallel to Fine 1994's arguments concerning essence and necessity) that potentiality should not be reduced to possibility, but understood on its own terms, and that this might be explained by potentiality being the more basic notion from which that of possibility ought to be derived. In chapter 4, I have then provided some steps toward the required understanding of potentiality on its own terms, and argued that the notion of potentiality is a very flexible one. I now return to the second and more tentative part of the proposal: can we understand possibility as derived from potentiality?

The intuitive idea of a potentiality-based account of possibility is this. When we think about possibility, we think about potentiality in abstraction from its bearer; a possibility is a potentiality *somewhere or other* in the world, no matter where; possessed by something (or some things) or other, no matter what. Of course, the potentialities may be iterated: in possibility claims, we abstract not only from a potentiality's bearer, but also from the number of steps that it takes the potentiality to reach its manifestation.

Progressing from this intuitive idea to a more precise formulation, and using 'iterated potentiality' to include once-iterated potentialities (i.e., potentiality simpliciter) in accordance with chapter 4.5, I propose the following:

Possibility It is possible that *p* if and only if something has (or some things jointly have) an iterated potentiality for it to be the case that *p*.

Obviously, it is necessary that *p* if and only if it is not possible that not *p*; i.e., if and only if nothing has (or no things have) an iterated potentiality for it to be the case that not-*p*.

I have distinguished between two senses of 'potentiality' in the previous chapter. The definition of possibility may be read using either of them. I have argued that the boundary of the stronger sense is somewhat unclear and that we should therefore accept the weaker sense. However, I have also argued that something has or some things have a potentiality for *p* on the weak conception if and only if something has or some things have a potentiality for *p* on the strong conception (however exactly that conception is delineated; see (W2) in section 4.6.) So the two ways of reading the right-hand-side of my possibility definition are equivalent.

To prove its rank as a serious contender, my proposed account of possibility has to meet two requirements. First, it has to be formally adequate: possibility defined in this way must conform to the formal structure of possibility as studied in modal logic. Second, it has to be materially or extensionally adequate: it has to make the right possibility claims come out true (or false, as the case may be). My attempt to meet the first requirement will be given in the next chapter, where I provide a formal treatment of potentiality based on the metaphysics given in chapter 4, and derive from it, by the formal equivalent of my above clause for possibility, the logic of possibility. In doing so, I will use the weaker sense of potentiality, which is easier to formalize. This chapter is about the second requirement. I am going to argue that a potentiality-based account can meet the requirement of extensional adequacy, and that it has various options in meeting that requirement. That is to say, I am not going to put forward a particular version of the potentiality-based account. Rather, I am going to point out that there are various versions available, all of which preserve the spirit of the proposal made above. I will make life harder for myself, and my answers to the various challenges more intuitive, by restricting myself to clear examples of potentiality on the stronger conception. My argument against that conception was based on its lack of a sharp boundary delineating it from the weak conception. But I will not appeal to any of the borderline cases; all the potentialities that I will be appealing to should be uncontroversially admitted by the strong conception of potentiality as well as the weak conception.

No matter how exactly the account is developed, it will be a realist, actualist and modalist account: it will be an account that rejects the existence of worlds other than ours, and that takes modality in some form (in my case: potentiality) to be an objective and primitive feature of our world. It differs from other modalist accounts in the kind of modality that it takes to be a primitive feature of the world: not necessity, not possibility, but potentiality. Modality, as a primitive feature of the world, is primarily a feature of the objects of our world. Of course, like any other actualist account, mine can admit that there is some interest in talk about possible worlds, only we should not take the label 'worlds' too seriously. In fact, we should not take the label 'possible' too seriously either. The so-called possible worlds of modal logic are a useful device to produce a formal model, but they are neither worlds, nor do they have any more intimate connection to possibility than they do to any of the other phenomena that they can be used to model: knowledge, obligation, vagueness, and a wide variety of others.

With these general remarks in mind, let me return to the requirement of extensional adequacy. This seems relatively easy to get in the cases of simple predicative possibilities: it is possible that the apple is red because the apple itself has a potentiality to be red; it is possible that something is red because something, the apple, has a potentiality to be red. It is possible that my granddaughter will be a painter, even though there is no individual yet who is my granddaughter, because I have a three-times iterated potentiality for it to be the case that my granddaughter is a painter: a potentiality to have a child who has the potentiality to have a daughter who has the potentiality to be a painter.

But what about the possibility that everything is red, or that nothing is? And what about the possibility that the laws of nature had been different? I will discuss these and a variety of other potentially problematic examples in the remainder of this chapter. My discussion will be primarily exploratory: trying to understand which options are available to a potentiality-based theory of possibility. The positive conclusion will not be an endorsement of any particular such option. Rather, it will be that there are a great many options, several of which are promising. A potentiality-based account of modality is a fruitful research programme, though it is one that I cannot claim to have carried out here.

I begin with an issue that is slightly different from the ones to come. It is the question of how to account for the context-sensitivity of modal claims in natural language (section 5.2). After that, I will turn to worries that are more metaphysically motivated, and outline ways to respond to them (sections 5.3-5.7).

5.2 Restricted Modality

The possible-worlds approach to modality comes with a range of theoretical advantages. One advantage is its systematic relation to possible-worlds treatments of other areas, such as epistemic modality, semantic content, and of course counterfactual conditionals. Another is the pure elegance of reducing the operators 'possibly' and 'necessarily' to the well-understood quantifiers 'some' and 'all'. I can match only half of that elegance: possibility and necessity, on the view I propose, are defined with existential and universal quantifiers, but their interdefinability is not quite as elegantly rendered as it is on the possible-worlds approach. The reduction to the quantifiers has a further benefit apart from mere elegance, however. In everyday talk, we tend to quantify with restrictions: if, looking into the fridge, I say 'there is no beer', I am not thereby claiming that there is no beer to be found anywhere in the world. My quantifiers are restricted to the contents of the fridge. (The example, of course, is Lewis's: Lewis 1986a, 136f.) Similarly, if I say that it's impossible for me to be in New York an hour from now, I am not claiming that there is no world where I am in New York an hour from now (or, where my counterpart is in the counterpart city of New York an hour after the time that corresponds to now) – rather, I am restricting my quantifiers to worlds that are like this one in terms of history up to now (in all of which I am in Oxford now) and in terms of the laws of nature. We can easily accept a sense in which it *is* possible for me to be in New York an hour from now – if we had some more advanced, science-fiction technology, I might travel to New York from Oxford in less than an hour. In saying this, I am widening the scope of my quantifiers to include worlds that are much like this one but contain more advanced technology than ours. Of course, even with such technology, I could not travel faster than the speed of light. Or could I? That last question invites us to drop the restriction to worlds that share our laws of nature. In everyday language, we almost invariably use some restricted sense of possibility, and it is a matter of context just how restricted it is. When we do metaphysics, we try to drop these restrictions and talk in a context that poses no restrictions at all, about what is possible in the most general sense. Similarly, in everyday language, we almost invariably restrict our quantifiers to some relevant subset of the objects there are, and just which subset it is that we quantify over is a matter of context. When we do metaphysics, we try to drop these restrictions and talk in a context that poses no restrictions on our quantifiers, about what exists in the most general sense. Understanding modal talk as quantification over possible worlds allows us to understand the first phenomenon, the context-dependence of modal talk, as merely a case of the second phenomenon, the context-dependence of our quantifiers' range. The account I am proposing should provide the resources to explain the context-dependence of modal talk without appealing to quantification over possible worlds.

I believe that it can. A first point to note is that given the metaphysics of potentiality and the evidence provided in chapter 3.3.1, there is reason to believe that many putative possibility statements *are* ascriptions of potentialities. Any sentence of the form 'x can F', on the potentiality view, may be construed as ascribing to x a potentiality to F. However, there are possibility statements (of the metaphysical, not the epistemic) variety that do not specify a subject. While I have advised caution about those statements (chapter 3.3.1), I will here sketch how contextual variation can be accounted for, on the potentiality view of possibility, in those statements that do express circumstantial or metaphysical possibility, rather than ascribing potentialities. On my view, 'It is possible (for it to be the case) that p' (or 'It could have been that p', etc.) is true just in case something has an iterated potentiality for it to be the case that p. There are two places in this formulation that may be subject to contextual restriction: one is the existential quantifier ‘something’, the other is the potentiality.

First, the existential quantifier: notwithstanding their very different metaphysics, my proposed account of possibility shares a crucial structural feature with any possible-worlds semantics: the existential quantifier. To be sure, mine quantifies over things in this world, while Lewis’s quantifies over worlds. But both are existential quantifiers; both provide an opportunity to restrict the domain to which those quantifiers are applied. In fact, since my existential quantifier quantifies over the things of this world, there are precedents for the mechanism by which it is restricted (not just the formal mechanism, but the details of what goes into it): whatever we say in general about quantifier domain restriction in ordinary speech contexts may be transferred without alteration to the restriction of the quantifier inherent in possibility statements. That makes life easier in one way: a potentiality-based account does not have to say anything new about the way in which the existential quantifier in its clause for possibility is restricted. It also makes life more difficult: there are independent ‘guidelines’ for the kinds of restriction that are plausible and the kinds that aren’t. I take this to be a virtue of the account: it allows an independent test of how plausible some of its explanations are, a test that is not available for possible-worlds treatments of modality.

Obviously, I cannot give a comprehensive account here of the kinds of such restrictions that there are; but I will point out a few typical examples. Lewis’s beer example illustrates some of them. First, we tend to quantify over things in our vicinity: the things in the fridge that I am looking into, or the flat in which the party is taking place, or perhaps the flat and the near-by shops (if I’ve tried them). Second, we tend to quantify over things of a certain relevant kind (not necessarily a natural one): if I say that there is nothing in the fridge, I ignore everything that is not food or drink (after all, there's a rack). Third, we tend to ignore even things of the right kind and distance if they are inaccessible for us to interact with in some salient way. Thus in saying that there is no more beer I may include items in the flat and in the nearby shop that I have just been to, but ignore items in my neighbour's fridge (my neighbour is out, and I don't have his key). This third aspect may even be the source of the first two: we are more likely to interact with things in our vicinity, and certain salient kinds of interaction require certain kinds of objects for us to interact with.

Second, restrictions on the potentiality ascribed: in the very short sketch of a potentiality-based semantics of 'can' that I have given at the end of the previous chapter, we have already seen that and how context can pose restrictions on the kind of potentiality that is relevant. Here are two ways in which a given context may plausibly restrict the relevant potentialities. (i) Potentialities come in degrees, and we tend to neglect those that have a degree below a certain threshold. Thus we call things fragile that have the potentiality to break to a certain minimal degree, and mostly ignore the fact that desks and rocks are also 'breakable', or can break. Potentialities with low degrees escape our notice when we talk about potentialities explicitly; it should not come as a surprise that they escape our notice too when we quantify over things that have them. There is good reason for this: the lower the degree of a potentiality, the less imminent (*ceteris paribus*) is its manifestation. (ii) The proposed account of possibility is phrased in terms of iterated potentiality. The recognition of iterated potentialities required some argument from my part. Iterated potentialities escape our notice when we talk about potentialities explicitly; it should not come as a surprise that they escape our notice too when we quantify over things that have them. Again, at least in practical contexts there is a good reason for our neglecting those potentialities: their manifestations are not, in general, particularly imminent.

With these observations in mind, we can formulate provisional truth-conditions for possibility statements in natural language that respect their context-sensitivity. An utterance of the form 'it is possible for it to be the case that *p*' (or: 'it can be that *p*', etc.) is true just in case there is or are some *salient* object(s) that (jointly) have a *salient* iterated potentiality for it to be the case that *p*. Which objects and which potentialities are salient is for the utterance's context to determine.[1]

Now let me apply my provisional truth-conditions to a few examples.

The restriction of our quantifiers to things that we can expect to interact with helps us understand a kind of case that would otherwise be puzzling. Suppose that I am moving to a new house, and discuss how to handle the furniture. I say 'the vase is quite fragile, but it cannot break (it is virtually impossible for it to break) – it's been packed so well in styrofoam.' Here I have attributed a potentiality and gone on to deny the possibility of its being manifested; applying the possibility clause above, I have said that the vase is fragile (has the potentiality to break to some non-minimal degree) and yet nothing (salient) has a (salient) iterated potentiality for the vase to break.

Contextual salience may shift within the course of one sentence; even the range of the quantifiers may (cf. Lewis 1986a, 6, who cites the example: 'Nowadays there are rulers more dangerous than any ancient Roman'). In the second part of the quoted sentence, the vase is no longer salient. There is a simple reason for this: being packed in styrofoam, the vase on its own is not one of the objects I can reasonably expect to interact with in the relevant ways – lifting it, dropping it, putting it in the van. What I do interact with is the vase-cum-styrofoam, and this has the potentiality for the vase to

[1]Note that the verbs on both sides of the biconditional may be read as tensed or time-indexed. On that reading, 'it is (now) possible for it to be the case that *p*' is true just in case *now*, at the time of utterance, something has an iterated potentiality for *p*. Potentialities may be lost over time: nothing now has a potentiality to avoid World War I, though some individuals had such a potentiality in 1914. This goes well with the idea that the past is in some sense necessary: it is not (now) possible for World War I never to take place, though it *was* possible in July 1914.

break to a negligible degree – where, of course, context determines what is negligible and what isn't. The quantifier that (I say) is implicit in the possibility statement does not differ from other quantifiers: when asked how many things are still left to be put into the van, I do not count vase and styrofoam as two; I count them as one, because the relevant interaction (putting in the van) can be expected to involve the vase-cum-styrofoam, not the vase on its own.

What of the examples I started with? It is not possible – so I say in an everyday context – for me to be in New York one hour from now. No salient things in the world, including myself, New York, and aeroplanes, have a salient (joint) potentiality for me to be transported to New York from Oxford in one hour. Well: what about the potentialities that engineers, or the potential parents and teachers of engineers, etc., possess for science-fiction technology to be developed which would have the potentiality to transport me from Oxford to New York in one hour? The counterfactual supposition 'if we had science-fiction technology' invites us to widen the domain of relevant potentialities: in particular, it invites us to consider such iterated potentialities as the ones I have pointed to. Thus the domain is expanded, and it appears possible after all that I should travel to New York in one hour. How much further can we widen that scope, and will we finally have to include the possibility of my travelling faster than light? This questions leads us already into the worry that will be discussed in the remainder of this chapter: the worry that there might not be enough potentialities around to provide grounds for all the possibilities to be found in metaphysics; that possibility, on the potentiality-based account, is construed to narrowly. It is to this worry that I now turn.

5.3 Objections and Strategies

There is one obvious form of objection to the view I have proposed: counterexamples of the form 'It is possible that *p*, but nothing has a potentiality (no matter how many times iterated) for it to be the case that *p*'. These counterexamples are best understood as a challenge: a challenge to point out where in the world something has an iterated potentiality for it to be the case that *p*.

Here are some examples.

Alien properties. David Lewis has argued (see Lewis 1986a, 159ff.), and has convinced many, that there could have been properties that are completely alien to our world. For David Lewis, a property is alien to a world iff it is never instantiated in that world. Of course my proposal allows for the possibility of alien properties in this sense: some potentialities may never be manifested, and hence their manifestation property never instantiated; so some properties that are never instantiated are nonetheless possibly instantiated. However, there is no guarantee that the potentialities of actually existing objects 'reach' all the alien properties there might have been; there may be uninstantiated potentialities for some alien properties but not others. Or so it may be argued.[2]

Alien objects. It is contingent which objects there are; there might have been different objects from the ones there actually are. In fact, it might have been that none of the actually existing objects had existed, and that the world had consisted of a completely different set of objects. Now, the actually existing objects may have potentialities for there to be a whole range of other objects: people have unmanifested potentialities to have a child of such-and-such description, to build a watch of such-and-such descrip-

[2]Lewis uses alien properties to argue against ersatzist views on possible worlds. More recently, Divers and Melia (2002) have argued that Lewisian modal realism itself has difficulties in dealing with certain intuitions concerning alien properties without appealing to modal concepts. So whatever the potentiality view's standing is with regard to alien properties, it will not face problems alone.

tion, etc.; things jointly have potentialities to constitute (to jointly be all the parts of) infinitely many objects that they never actually constitute; and so on. But, once again, that may not seem sufficiently far-reaching: could there not have been objects that were not produced by any actually existing objects (or by anything that actually existing objects have a potentiality to constitute or produce, or ...), nor constituted by any of the actually existing objects (or anything that actually existing objects have a potentiality to produce, or ...)? Moreover, which actually existing object(s) should possess a potentiality, even iterated, for it to be the case that none of the actually existing objects ever existed?

Laws of nature. The world is governed by certain laws, the most fundamental of which are probably the laws of (real) physics. But it might have been governed by a different set of laws; the laws of nature are contingent. Yet how could any of the things that are actually governed by the actual laws have an (iterated) potentiality for those laws to be different?

Those are typical objections to the potentiality-based account of possibility. They take the form: surely it is possible that *p*; potentialities, perhaps, can take us some way towards *p*, but there is no guarantee that they will take us all the way; the intuition about *p* is too fundamental for the concept of possibility to be left to (epistemic) chance concerning which potentialities there are; so the potentiality view should be rejected.

I find some of these counterexamples genuinely worrying (and others less so). Before looking at them in more detail, it is important to be aware of the options that are available to the potentiality view in responding to them. The potentiality view has a variety of strategies at its disposal, and of course different strategies may be applied to different cases. Here is a general outline of these options.

Strategy 1: The Piecemeal Strategy. This strategy takes the counterexamples one by one and tries to offer potentialities of objects that take us 'all the way' to the alleged possibility that *p*. We have seen in the previous chapter that the notion of a potentiality is more flexible than we might at first have thought; it has been stretched far indeed from its humble beginnings with dispositions and abilities. That should be reason for some (guarded) optimism: with a little ingenuity, we should find the potentiality that is responsible for its being possible that *p*.

Strategy 2: Biting the bullet. The piecemeal strategy, by its very nature, may work in some cases but not in others. Where it does not work, we may consider simply biting the bullet and saying: yes, we thought it was possible that *p*, but we were wrong. There are precedents for this kind of claim: before Kripke, did people not think it was possible that water ≠ H_2O? There are precedents even for some particular such claims: the view that the laws of nature are metaphysically necessary, for instance, has been increasingly popular in the last decade or so. In biting that bullet, the potentiality view will at least not have to stand alone. However, in denying what appeared (at least to the opponent) to be a reliable intuition regarding the possibility of *p*, this strategy incurs an obligation: to explain how we (or, at any rate, the opponent) came by that intuition.

Sydney Shoemaker, in defending one restriction to what is metaphysically possible (viz., only what is nomologically possible), says that 'the main theoretical obstacle to [recognizing this restriction] has been the deeply rooted conviction that necessary truths should be knowable a priori', a conviction that metaphysicians after Kripke should long have abandoned (Shoemaker 1998, 74). Of course, we cannot know a priori what the laws of nature are, and we cannot exclude variant laws a priori – but then, we could not know a priori that water is H_2O, and hence could not exclude a priori that it was XYZ. In general, the problematic possibilities (problematic, that is, for the potentiality view, which does not accord to them the status of possibilities) are not metaphysical possibilities but only, in some way, epistemic ones.

One way of substantiating this strategy is to borrow from linguistics. I earlier (chapter 3.3.1) pointed out that sentence-modifying modal expressions ('It is possible that', 'possibly', etc.) usually express epistemic modality (compatibility with our knowledge of the world) rather than circumstantial modality (ways the world could *really* be), while circumstantial modality is typically expressed by predicate-modifying expressions such as 'can'. (Cf. Kratzer 1981, DeRose 1991) Linguistically speaking, metaphysical possibility should be a generalization of circumstantial possibility: like circumstantial possibility, it concerns 'real' possibility; it differs in not relativizing that possibility to how things stand actually and contingently. So the modal intuitions that we rely on in doing metaphysics, insofar as they are based on linguistic intuitions, should be based on expressions of circumstantial, not epistemic, possibility. They are, however, very often phrased in the form 'it is possible that p' (and other sentence-modifying expressions) – expressions of epistemic possibility *par excellence*. In adopting the strategy of biting the bullet, then, the defender of a potentiality view may claim that the apparent intuitions regarding the metaphysical possibility of a proposition that p are based on the opponent's misunderstanding her own linguistic intuitions: it is possible that p in some generalized epistemic sense, but not metaphysically. (It may be replied that these expressions are artifacts of philosophers, and that we can stipulate them to mean whatever we please. Of course we can: but if we do, then we have no right to appeal to our linguistic 'intuitions' about the truth of sentences containing them. We have linguistic intuitions as speakers of a natural language; if we stipulate what our terms mean, then we neither need to, nor can we, appeal to our intuitions about them.)

Of course strategies 1 and 2 may be used together, accounting piecemeal for some possibilities while rejecting others, and giving some principled reason for drawing the line right where it is drawn. This mixed strategy may be called a strategy of *Divide and Conquer*. It will be applied throughout sections 5.5 to 5.7. First, however, I will look at an alternative: a catch-all solution.

Strategy 3: The Catch-all Solution. The strategy of Biting the Bullet is as piecemeal as the Piecemeal Strategy itself: we can bite the bullet in some cases but not others, and when we do, we are committed to some explanation of the misled intuition in the particular case. The strategy I am now going to suggest is different: the idea here is to solve all problems in one go. This strategy takes seriously the opponent's worries about epistemic chances above: the opponent wants a guarantee that potentiality is always going to go far enough? She can have it. (Or, perhaps, she can have it for the majority of important cases. Some bullet-biting may be unavoidable in the end.) The catch-all solution is to come up with an object or set of objects that plausibly possess all the potentialities required for the problematic possibilities. One such object is the world itself. Another option is to distribute potentialities more evenly and allow for there to be many more objects than we might have thought: this option is to adopt Timothy Williamson's arguments to the effect that there are merely possible objects – whenever it is possible that there be an object which is F, then there is an object which is possibly F. Ascribing potentialities to Williamson's merely possible objects is another catch-all solution. Because this kind of solution is general in spirit, I will give it a closer look on its own in the next section, before examining in some more detail the most prominent counterexamples in the spirit of the piecemeal strategy.

5.4 The Catch-All Solution

5.4.1 Ways the World Might Be

Suppose that nothing in the world has or ever had a potentiality to be a unicorn, or even an iterated potentiality for there to be unicorns. Suppose further that we wanted to say both (i) that it is possible that there existed a unicorn; and (ii) that it is possible that there existed nothing but unicorns. (Of course, if there were unicorns, there would also have to be parts of unicorns, etc.; I leave these implicit for the sake of simplicity. You might also be convinced by Kripke 1981 that it is impossible for there to be unicorns; in which case, feel free to replace the example with any other that you prefer. The points I want to make are structural and do not depend on the nature of the illustration.) On my account, it must then be that (i) something (or some things) has an iterated potentiality for there to be a unicorn; and (ii) something (or some things) has a potentiality for it to be the case that everything was a unicorn. What could that be?

There is one very simple answer: the world itself has a potentiality (i) to contain a unicorn, and (ii) to contain nothing but unicorns. Similarly, suppose L_1 is the set of all the actual laws of nature; and suppose further that we want to say that it is possible for the world to have been governed instead by a completely different set of laws, L_2. Then something must have an iterated potentiality for the world to be governed by the laws in L_2 instead of those in L_1. What could that be? Again, there is a very simple anwer: the world itself has a potentiality to be governed by the laws in L_2 instead of those in L_1.

That simple answer, however, needs to meet two challenges. One is to clarify what exactly is meant by 'the world'. Another is to make sure that the potentiality account does not, by appealing to potentialities of the world, collapse into a particular version of (moderate) realism about possible worlds, namely, Stalnaker's identification of possible worlds with 'ways the world might be'.

The proposal at issue is that we ascribe potentialities to the world itself. Some such potentialities may be potentialities to contain unicorns, or to contain exactly one unicorn, or to be governed by the laws in L_2. Once that much is accepted, why not ascribe potentialities to the world for properties that are maximal states of the world as a whole, an entire 'way the world might be'? And once that is conceded, why not accept an identification that Robert Stalnaker has suggested, between such 'ways the world might be' and possible worlds? Finally, once I have thus helped myself to possible worlds, why not do away with all the potentialities of particular objects that and simply do possible-worlds semantics like everyone else? Let me begin with a look at Stalnaker's view.

Stalnaker's identification of possible worlds with 'ways the world might be', or rather his interpretation of that identification, is a reaction to an argument that David Lewis gives for modal realism. Here is Lewis, trying to convince us that possible world are not such strange and unfamiliar entities after all:

> It is uncontroversially true that things might have been otherwise than they are. I believe, and so do you, that things could have been different in countless ways. But what does this mean? Ordinary language permits the paraphrase: there are many ways things could have been besides the way that they actually are. On the face of it, this sentence is an existential quantification. It says that there exist many entities of a certain description, to wit 'ways things could have been'. I believe permissible paraphrases of what I believe; taking the paraphrase at its face value, I therefore believe

> in the existence of entities which might be called ‘ways things could have been’. I prefer to call them ‘possible worlds’. (Lewis 1973, 84)

Stalnaker responds:

> The argument Lewis gives for ... identifying possible worlds with ways things might have been, seems even to be incompatible with his explanation [elsewhere] of possible worlds as more things of the same kind as I and all my surroundings. If possible worlds are ways things might have been, then the actual world ought to be *the way things are* rather than *I and my surroundings*. *The way things are* is a property or a state of the world, not the world itself. The statement that the world is the way it is is true in a sense, but not when read as an identity statement (compare: ‘the way the world is is the world’). This is important, since if properties can exist uninstantiated, then *the way the world is* could exist even if a world that is that way did not. One could accept [the] thesis ... that there really are many ways things could have been ... while denying that there exists anything else that is like the actual world. (Stalnaker 2003, 28, first published as Stalnaker 1976)

Stalnaker’s suggestion that possible worlds, rather than being concrete things like this world, should be identified with uninstantiated *properties* of this world, is clearly and explicitly non-reductionist. After all, those uninstantiated properties of the world that should be counted as possible worlds are those ways that the world *might* or *could* have been. Stalnaker is not interested in a reduction of modality: ‘modal notions are basic notions, like truth and existence, which can be eliminated only at the cost of distorting them. One clarifies such notions, not by reducing them to something else, but by developing one's theories in terms of them.' (Stalnaker 2003, 7) The concept of a possible world, for Stalnaker, is a functional one, like the concept of an individual in formal semantics: a possible world is whatever plays a certain role in rational activities such as distinguishing how things are from how things might have been. (Cp. Stalnaker 2003, 8, 38)

I agree with a great deal of what Stalnaker says. I can accept the concept of a possible world as long as it is understood to be a functional one; though I may want to restrict the role it can play a little more narrowly than Stalnaker. I agree that modal notions are basic; but I stress that there are more modal notions than possibility, necessity and the counterfactual, and that developing our theories in terms of modal notions may involve imposing an explanatory hierarchy on them, and I claim that the fundamental element in that hierarchy is potentiality. Finally, I agree that uninstantiated properties of the world are the best candidates to play the functional role of possible worlds, and that singling out those among the uninstantiated properties of the world that do play this role involves already applying our modal concepts; however, I think the modal concept to be applied is that of a potentiality. I would take it to be an advantage if I could provide something to play the role of possible worlds, and thereby explain the success of possible-worlds-based semantics (in logic and linguistics). But can I make sense of the world possessing potentialities of its own, and can I make sense of it without thereby making the definition I have given in terms of potentialities becoming superfluous?

5.4.2 'The World'

This, then, takes me back to the original challenges. To begin with: what is meant by 'the world'? This is a question that Stalnaker, if he spelled out his suggestion, would have to answer as well as I do.

First, by ‘the world’ we might simply mean the totality of what exists, the mereological sum of all the things ‘in’ the world. But that understanding will not deliver the results that both Stalnaker and I need. For how could the world, understood in this sense, have contained objects other than the ones it does contain? (It does not matter whether we understand this ‘can’ as expressing possibility, as I suppose Stalnaker would, or (joint) potentiality, as I would. No objects have a potentiality to be any objects other than they are; no mereological sum has a potentiality to be another sum, i.e. a sum of other objects, than it is.)

Second, we might mean by ‘the world’ something that is constituted by the things that exist but could have been constituted by other things; a ‘whole’ that has all things that exist as its ‘parts’ but is not just the (mereological!) sum of its parts. The analogy is with complex objects such as ourselves: I am constituted by my bones, organs, etc., or further down by the molecules that make up my body, but I might have existed while being constituted by parts other than these. The analogy is very limited, however. A complex object such as a human being can exist with parts different from the ones that it actually has, but there are limits as to what it can have as parts instead – in a transplantation, my heart may be exchanged by another heart or even (with some technological advance) by an artificial heart consisting of different material, but it cannot be exchanged with, say, the Eiffel tower. Certain structural features must be maintained in exchanging a complex object’s parts. Does this apply analogically to the world itself? If so, what kind of ‘structure’ should we ascribe to the world such that this structure would be maintained if the world consisted only of a lonely apple, or of a multitude of unicorns, or perhaps even (though I for one would be willing to give up on this option) of immaterial souls?

Bigelow et al. (1992) have proposed that we treat the world as 'one of a kind', that is, an object which belongs to a natural kind and thereby has certain essential properties. Their concern is to provide a grounding in natural kinds for laws of nature that seem to concern the very structure of the world: for instance, the conservation laws, the principles of relativity, and the symmetry principles (Bigelow et al. 1992, 371). Bigelow, Ellis and Lierse themselves are open to its being possible that instead of the world that we inhabit, there might have been a world of a different natural kind; indeed, that is the only way for them to make sense of the possibility that the laws had been different. However, I am not concerned to ground laws of nature in the essential nature of the world, but to ground possibilities in its potentialities, and if certain laws of nature were essential to the world, then the world could have no potentiality to be without those laws, and there is nothing that might have a potentiality for there to be another kind of world altogether. Hence, if we understood the structural features of the world to be those laws that are essential to it (on this view), then potentialities of the world might take care of the possibilities of alien objects or properties, but not of radically different laws of nature.

Joseph Almog has suggested a rather different view of the 'structure of the world': it is described by logic (Almog 1989). Structural facts or 'pre-facts', on Almog's view, are those that are 'permutation-resistant': they are facts that would still remain facts if the objects or properties that constitute them were 'permuted' with any others. On this view, 'Quine exists' is not only a truth of logic, but a structural fact about the world (substitute any other object for Quine, and the fact remains a fact – the example is Almog's, and the intended sense of 'exists' is a timeless one). Of course, Quine's existence is contingent, and in fact Almog explicitly divorces logic from modality (Almog 1989, 203: 'the falsehood of a proposition in a counterfactual situation is no foothold against its truth in virtue of the structural traits of this, very actual, world'). Again, this view does not appear very promising for my present purposes: after all, I moved to the second sense of 'the world' precisely to find a sense in which the world might have existed without containing, or consisting of, the very objects that it does.

The second sense of 'the world', then, if it is to be appealed to in defending the potentiality-based account, would have to be spelled out in some new way different from either Bigelow et al.'s or Almog's, or else its use as a catch-all solution may be rather limited.

Third, we might think of 'the world' not as the whole made up of all existing things as its parts, but rather as the space in which those things exist; not the totality of things in the world, but that in which those things all are. Things would be 'in' the world, on that third understanding, not in the way in which (to borrow an example from Stalnaker for a different purpose) raisins are in a pudding, but rather in the way in which water is in a glass, or better still, things are in space. What kind of thing is 'the world' in this perhaps unfamiliar sense? If you are a substantivalist about spacetime, you might think of spacetime as the world. If you can make sense of the idea of a structure of the world, it might be the structure in distinction from that which is structured by it (the form in distinction from the matter that has it) – but that understanding, of course, would run into the same difficulties as the attempt to treat the world as a structured entity (form and matter). Neither Bigelow et al. nor Almog provide a candidate for being the 'structure of the world' that would suffice for the catch-all solution to work. That, of course, in no way proves that no such candidate *can* be found. However exactly it is spelled out, the world in this sense would be a different thing from the things it contains; though it is a 'thing' only by stretching that term a little. The term must, of course, be thus stretched in order for it to be plausible that we ascribe potentialities to 'the world'.

I have not defended any of these meanings we can give to 'the world'. I admit that I know of no defense for the claim that there is any such thing as 'the world' in the second or third sense, and of no extant conception of the world in either of these senses that would quite suffice for the catch-all solution envisaged. (It is trivial that there is such a thing as the world in the first sense if unrestricted mereological composition is true; and it is likely that there is not if unrestricted mereological composition is not true. But that first sense is of no use for my present purposes.)

I began with a comparison between my catch-all solution and Stalnaker's view of possible worlds. It is worth noting that Stalnaker's suggestion is in the exact same situation as my proposed catch-all solution: for him too the interpretation of 'the world' as the mereological sum of all there is will not do. There is no sense in which that sum or totality could have instantiated the property of being made up of entirely different objects; that is not a way that this totality might have been. If the challenge can be met of spelling out in a satisfying manner what Stalnaker must mean by 'the world', then that same challenge is met for the sake of a potentiality-based view's catch-all solution.

5.4.3 Potentialities of the World and Possible Worlds

I can now return to the second challenge: why does a potentiality account which allows for there to be potentialities of the world not collapse into a possible-worlds based account of modality?

In this section, I will argue that even given a commitment to potentialities of the world and the possibility of identifying those potentialities with possible worlds, my theory is nonetheless better off speaking of the potentialities of individual objects, rather than just defining possibility through the potentialities of the world (and thereby very nearly collapsing into Stalnaker's suggested theory). The argument differs slightly depending on which sense of 'the world' is adopted.

First, the world-as-a-whole (the second of my three senses distinguished above). If we follow the analogy with a complex object, the potentialities of the whole – the world – are constituted at least in part by the potentialities of its parts – the things in the world. True, the whole might have had (has a potentiality to have) different parts; nonetheless, the parts that it does have contribute to its potentialities. Thus it is true that the world-as-a-whole has a potentiality to be such that I am sitting. However, the world has that potentiality in virtue of *my* having the potentiality to be sitting, not vice versa. The world as a whole also has a potentiality to be in various total states, some of which include my sitting while others do not. Let S be such a total state which includes my sitting. (S is a 'possible world', if you like.) How is it that the world has a potentiality to be in S? That potentiality is constituted in part by the potentialities of the things in the world, such as my potentiality to be sitting, just as my potentiality to be sitting is in turn constituted in part by the potentialities of my legs and back, or my bones and muscles, or my molecules, etc.

Of course, not all the potentialities of the world will be like this. If the world has a potentiality to contain nothing but unicorns, this will not be grounded in any particular thing's potentiality (other than the world itself). Note here one disanalogy with the case of a complex object. I (a complex object) can have an (extrinsic) potentiality to have some particular other object(s) as my part(s). I have a potentiality to eat and digest this apple, making its molecules become part of me. That potentiality, like any potentiality concerning another object, is extrinsic and grounded in a joint potentiality that I possess together with the molecules of the apple. The world-as-a-whole does not have extrinsic potentialities to contain any particular objects other than the ones it does contain, for there are no other objects for it to have a joint potentiality with. It can only have potentialities to contain other objects of a certain kind, just as I can have a potentiality to contain other objects of a certain kind even if nothing of that kind exists, e.g. an artificial heart.

Second: the world-as-container. You, I and the apple in front of me are not *parts* of the world in this sense; we are merely in it. The world has the property of containing me, but that property is not, as it was for the previous sense, an intrinsic property (just as my property of having this particular heart is intrinsic). It is an extrinsic property grounded in a relation between two distinct objects: the world on the one hand, and me on the other. The 'total states of the world' or ways things (as a whole) could have been are, in this sense, to be construed slightly differently: there is the way *things* could have been – a total state of things in the world, call it S, and there is the state of the world which is, strictly speaking, not S but the property of containing (all and only) things in S.

The world, in this sense as well as the previous sense, has a potentiality to be such that I am sitting, and a potentiality for S to be the total state of the things in it, where S is a particular state that includes my sitting. Now, if the world-as-container has a potentiality (as it should) to contain me sitting, then that potentiality is an extrinsic one: it involves an object disjoint from the world itself, me. It is a potentiality that is grounded in a joint potentiality which I have with the world: to stand in the contains-sitting relation. On its own, the world-as-container could have no such potentiality; its intrinsic potentialities can now not only involve no particular non-existent object (that we have seen to be the case with the world-as-a-whole too), but also no particular object existing in it. It can have an intrinsic potentiality to contain unicorns, or to contain unicorns and nothing else, or to contain someone who is sitting and is F (where F is a complete list of my non-haecceitistic properties); but it can have no intrinsic potentiality to contain *me*.

Whichever of the two suggested senses of 'the world' we adopt, then, talk of individual objects' potentialities will not be idle. In both cases, the potentialities of the individual objects in the world will provide the metaphysical grounds for an important class of the potentialities of the world: those potentialities that concern any particular object at all. Unlike Stalnaker, I am looking for a substantial metaphysical account of where in the world modality is to be found. On the present version of the view I am proposing, modality *is* to be found in the potentialities of the world itself, and to the extent that we accept Stalnaker's identification of possible worlds, it is to be found in possible worlds. But that is not, metaphysically, the most enlightening thing to be said about it. For the source of much of the modality to be found in potentialities of the world is the potentialities of the particular objects in the world. That is why a potentiality-based account of modality, even if it accepts potentialities of the world (-as-a-whole, or -as-container), should nonetheless appeal to the potentialities of particular objects in the world wherever that is an option.

5.4.4 An Alternative: Mere Possibilia

I have discussed in some detail one catch-all solution: the ascription of potentialities to the world. There is an alternative catch-all solution. It is to allow that there are 'mere possibilia': objects that exist but do not have any properties other than logical or modal ones. The view to which I am referring is, of course, Timothy Williamson's necessitism (which I have already discussed briefly in chapters 4.4 and 4.7). Williamson has argued that everything necessarily exists (see Williamson 1998, Williamson 1999, Williamson

2002, Williamson forthcoming). On this view, it is not true that you and I (contingent beings that we are) could have failed to exist. Rather, we could have failed to be *concrete*. Further, if *we* could have been non-concrete, merely possible objects, then nothing prevents there from being actually non-concrete, merely possible objects: I do not have a younger sister, but I might have had one; according to Williamson, this is to say that there is something which is possibly my younger sister but fails to be concrete. Nothing *is* my younger sister; but something is, or more plausible many things are, my possible younger sisters. Whenever it is possible that something had been F, there is something which is possibly F, in accordance with the Barcan formula ($\Diamond\exists x Fx \rightarrow \exists x \Diamond Fx$):

> On the envisaged view, two very different states are possible for one object. It is capable of being an embodied person, knowing, feeling and acting in space and time. It is also capable of being a merely possible person, disembodied, spatiotemporally unlocated, knowing nothing, feeling nothing and doing nothing. Is so radical a difference in properties consistent with the identity of the object? But the two sets of properties are not wholly disparate. The person actualizes the potential to have properties characteristic of a person. The merely possible person has the unactualized potential to have such properties. What they share is the potential. Why should that not suffice? (Williamson 2002, 21)

Note Williamson's use of the term 'potential'. If there are such objects as his mere possibilia, then those objects may have potentialities as well as the concrete objects that we are more usually confronted with. We could then say that where it appeared that no thing (save, perhaps, the world) had a potentiality to be some way (e.g., such that everything was a unicorn), what is in fact true is that no *concrete* thing had a potentiality to be that way. But if there is more than the concrete things (plus the more familiar abstract objects), then there are more things that have potentialities. If it is possible that something is a unicorn and no concrete object has a potentiality to be a unicorn, then some non-concrete object has a potentiality to be a unicorn (and thereby concrete); if it is possible that everything is a unicorn and no concrete object has a potentiality to be a unicorn, then there there are some non-concrete objects with the potentiality to be a unicorn, and everything (concrete and non-concrete objects together) jointly has a potentiality for those to be the only concrete objects while all actualizing their potentiality to be unicorns.

But would not appeal to such mere possibilia introduce a further modal primitive, and thereby be a grave threat to the ambition of my project: to establish potentiality as the one fundamental source of modality in the world? It would not. Mere possibilia are not, despite their name, in any way 'modal' objects. They are called merely possible objects or mere possibilia because they possess all non-trivial, non-modal properties merely possibly: they are not persons, not thinking, not concrete; they are merely possible persons, possibly thinking, and possibly concrete. But those possibilities, on my account, are just unexercised potentialities. No modal primitive is needed in addition to potentiality. By widening the realm of what there is, and ascribing potentialities to everything in that widened domain, we can widen the space of potentiality-based possibility without introducing a new modal primitive.

However, there is one related problem with appeal to mere possibilia, though it is a problem of a purely dialectical nature. The main challenge to the potentiality view of possibility is to produce 'enough' potentialities to account for all the possibilities that we intuitively think there are. In responding to that challenge, the potentiality theorist will do better if she can point to potentialities that we have independent reason to ascribe to objects – independent, that is, of the possibility intuition in question. It is a difficult question whether and to what extent we can separate our intuitions concerning some (familiar) objects' potentialities from general possibility claims. But using mere possibilia as a catch-all solution leaves no room at all for such separation. Here is an example.

Take the challenge of showing that the laws of nature could have been different (given the potentiality view). Appeal to mere possibilia has one easy solution for this challenge: there are some mere possibilia, that is, actually non-concrete objects, which have potentialities to exist under laws of nature different from the actual laws. Why should we believe in the (non-concrete) existence of such objects? Williamson can say: if it is possible that there are things governed by different laws, then there are things (namely, mere possibilia) that are possibly governed by different laws, or that have a potentiality to be governed by different laws. If it were not possible that there were things governed by different laws, then there would not be any such objects (not even mere possibilia). There is no way of arguing for or against the existence of mere possibilia with a potentiality to exist under different laws of nature, except by direct appeal to the intuition that the laws could have been different, hence that there could have been things existing under different laws.

As a strategy in defending the potentiality-based account of possibility, then, appeal to mere possibilia runs the risk of having it too easy and thereby seeming ad hoc: the mere possibilia we stipulate to exist will be exactly those that the theory needs, and there is no independent touchstone for the plausibility of that stipulation. Let me emphasise again that this is not a refutation of the view, merely a dialectical burden: mere possibilia may seem to make life too easy for the potentiality theorist. Appeal to the world, in any of the senses that I have discussed, is not making life equally easy: we have some grasp on what we mean by 'the world' (though that grasp leads to a variety of different meanings), and that grasp is not already (or at least, not only) guided by anything we know or believe about possibility. For this very reason, it may turn out that there is no suitable sense of 'the world' which does the job we need it to do for a catch-all solution. Appeal to the world is riskier than appeal to mere possibilia; but that, I have suggested, may make it theoretically more virtuous.

Having examined in some detail the third of my three available strategies from section 5.3, I will now go on to look at some of the apparently problematic cases in detail, to see how they might be treated by the other two strategies, the piecemeal strategy and the strategy of biting the bullet.

5.5 Aliens

5.5.1 Alien Objects

The world contains a limited number of objects, instantiating a limited number of properties. Which objects there are, and which properties they have, seems to be mostly a contingent matter. There could have been an entirely different set of objects; and there could have been an entirely different set of properties instantiated. (Or rather, almost entirely different: there may be necessary existents, e.g. the number two, and necessarily instantiated properties, such as the property of being self-identical.) Or so the intuition goes. What can the potentiality account say to accommodate (or defuse) these intuitions?

Let me begin with the idea that there could have been different objects. This has two sides: there could have been some objects that don't actually exist, and some of the objects which do actually exist might not have existed. Taking both sides together and to the extreme, it might have been that none of the actual existents existed (leaving aside necessary existents, if such there are) and that an entirely different set of objects existed instead.

There could have been objects that don't actually exist: for the potentiality account, this is clearly true on one reading – the *de dicto* reading – and clearly false on another – the *de re* reading. (The latter is probably more difficult to hear, and would be better expressed by 'some things that don't actually exist could have existed.') It is not true that for some object which does not actually exist, that object might have existed; there are no objects that don't exist in actuality. Nor is this a peculiarity of the potentiality-based account: it follows from any actualist account of modality. (This includes the Williamsonian theory of mere possibilia: they, too, exist in actuality. The *de re* reading may then be true if slightly reinterpreted: there are object that are not actually concrete but could have been.) However, it is clearly true that there might been objects other than the ones that there actually are: my parents had the potentiality to have another child, the machines in a car factory have the potentiality to make another car (in addition to the ones they do make), an apple seed has a potentiality to become an apple tree (which is plausibly another object) which bears apples of such-and-such description that never exist if the seed is thrown away with the apple. So we have the possibility of alien objects by the actual objects' potentiality to *produce* them. We have, furthermore, the possibility of a great many more alien objects because objects in the world jointly have a potentialiy to *constitute* them: the handle of one knife and the blade of another have a joint potentiality to make up one knife even if they are never put together; the molecules in my body have joint potentialities to make up a great many things of various descriptions that are very different from me, and most of which will never come to exist; the atoms in the world have joint potentialities to make up different molecules, which would then have potentialities to constitute different objects; etc. Going all the way down to the fundamental particles, whatever those are (and if there are any – if Schaffer 2003 is right, we can just keep going down), it will be possible that there are such-and-such objects just in case those particles (or, if Schaffer is right, something somewhere down the infinite hierarchy of levels), or some of them, have a joint potentiality to constitute an object of such-and-such a nature.

That certainly yields a wide range of possible alien objects. It is nonetheless a limited range. It may be said that, intuitively, there could have been objects other than the ones that could have been made up of *our* particles; there could have been different kinds of particles, or things not made up of any particles at all – immaterial souls and their like. In response, the potentiality theorist can employ any one of the three strategies I have outlined above. She can bite the bullet and say that this restriction may indeed come as a surprise, but so do most interesting discoveries; the limit that it sets to the possibility of other things existing is narrower than we might have thought, but it is certainly not arbitrary. She can use the catch-all solution of appealing to the world in one of the senses outlined in the previous section: the world-as-a-whole may be thought to have potentialities to be made up of radically different things, perhaps, though this may be one of the cases where the analogy with the complex objects familiar to us breaks down. The world-as-container may be thought to have a potentiality to 'contain' radically different objects; here at least the analogy is not in our way. But, thirdly, she may try to use the piecemeal strategy and find some way of accounting for the possibilities at issue in some things other than 'the world'. There are (at least) two ways of doing this.

First: suppose that there are universals, and suppose that some of them exist despite never being instantiated. Suppose further that universals, and perhaps abstract objects in general, have potentialities. (I will say more about this supposition below, in section 5.7.) Then we can ascribe to those universals the potentiality to be instantiated (if universals possess any potentialities, then it seems reasonable they all possess this one), thus yielding the possibility for there to be the alien objects that instantiate the universal. Similarly, we can ascribe to those potentialities that are actually instantiated a potentiality to be uninstantiated, thus yielding the possibility that there existed no such objects as those that actually exist.

A second alternative might be more interesting. We can ascribe to the objects that actually exist the potentiality to exist *while* something instantiates an alien property; in which case we would have extended the realm of possible alien objects to those that some of the actual objects could co-exist with. Suppose, for instance, that there were two kinds of fundamental particles, the Q_1s and the Q_2s. The Q_1s have certain potentialities to interact with Q_2s. More precisely: Take any two particulars, a and b, where a is Q_1 and b is Q_2. a has the potentiality to interact with b in certain ways; this is an extrinsic potentiality of a which depends on the existence of b. a also has the potentiality to interact with any Q_2 in a certain way; this is an intrinsic potentiality of a which does not depend on the existence of any particular Q_2. In fact, it seems to be a potentiality which a might possess (though not exercise) in the absence of anything that is actually Q_2. And if this is so, then it seems only reasonable that a should also have potentialities to interact with things of a certain uninstantiated kind, Q_3, even though there really are no Q_3s. If a has a potentiality to interact with Q_3s, then a must a fortiori have the potentiality to be such that there are Q_3s for it to interact with.[3] In fact, a's potentialities extend even further: a may have a potentiality to interact with, and hence

[3]In case that suspicions are raised that I am here tacitly appealing to the weaker conception of potentiality, let me spell out this iterated potentiality according to the schema in chapter 4.5: a, which is a Q_1, has a potentiality to be such that some things of which it is one – namely, a and something which is Q_3 – have a (joint) potentiality to be such that something which is one of them – namely, the hypothetical Q_3 particle – is such that there are Q_3s (namely, itself).

to be such that there are, entities of a kind Q_3, where Q_3s are in turn such as to have a potentiality to interact with, and hence to be such that there are, entities of a kind Q_4, which may in turn ... a, which is of the kind Q_1, co-existing only with entities of kind Q_2, may have iterated potentialities for there to be entities of any kind Q_n that can be reached via a chain of potential interactions.

We now also have the material to answer the second aspect of the question here discussed: could it not have been that none of the actual objects existed, and an entirely distinct set of things had existed instead? The manifestation of an object's (more than once) iterated potentiality, unlike the manifestation of an object's (only once-iterated) potentiality, may involve the object's nonexistence. (Or so I argued in chapter 4.5.) Thus our particle of type Q_1, a, may have a potentiality to be such that there are objects of (the actually uninstantiated) kind Q_3, which may in turn have a potentiality to be such that there are objects of (the actually uninstantiated) kind Q_4 *while there are no objects of kind* Q_1. Hence the Q_1s themselves, or rather, any one of them such as a, can have a more-than-once iterated potentiality for there not to be any Q_1s. (If, as I have assumed implicitly, being Q_1 is an essential kind property, then that potentiality amounts to a potentiality for there not to be any of the actual Q_1s.)

Whether any of this is actually the case is, from the philosophers' part, mere speculation. It is a feature of the potentiality-based account that truths about what is possible and what is not are hostage to the way things actually are, about which we can find out (at best) by empirical inquiry. To a certain degree, that situation is accepted by every (orthodox) metaphysician post-Kripke. The potentiality-based view merely intensifies that dependence of the possible on the actual – but it does so by making the actual deeply modal.

5.5.2 Alien Properties

What I have been discussing in the previous section might as well have been labelled 'alien properties': after all, it was the possibility of actually uninstantiated kind properties being instantiated that was really at issue. When we move from kinds to properties, the arguments can be run in exact parallel. But they face a challenge that I have not yet discussed.

The first of my proposals – accepting that universals exist, and that they have potentialities – rested on the idea that the existence of a property is independent of its being instantiated. The second proposal – ascribing potentialities for alien kinds of objects to the real objects, in virtue of their potentialities to interact with those objects if they existed – did not explicitly make any such commitment. But the worry may be raised that it secretly relies on it too. I have argued earlier that an object cannot have a potentiality to stand in any relation to anything that does not exist: if the door did not exist, the key would not have a potentiality to open this particular door, and if there are no objects instantiating Q_3, then *a* has no potentiality to interact with any particular Q_3. Now the same line of thought may be applied to the properties that figure in a potentiality's manifestation: if there was no such property as being Q_3, then *a* could not have a potentiality to interact with, and hence to be such that there is, some Q_3. In some sense, the alien properties and kinds must actually exist if anything is to have a potentiality for them to be instantiated.

This line of reasoning can then combine with a Principle of Instantiation, as proposed by David Armstrong, to argue that both of the strategies that I have just recapitulated will fail. The Principle of Instantiation 'demand[s] that every universal be instantiated ... for each property universal [it must] be the case that it is a property of some particular[, and f]or each relation universal [it must] be the case that there are particulars between which the relation holds' (Armstrong 1989, 75). According to the Principle of Instantiation, there are no uninstantiated properties. If there are no uninstantiated properties, then of course no property could have a never-manifested potentiality to be instantiated; and it becomes at least difficult to see how objects should have potentialities for some alien property to be instantiated by themselves or anything else.

All that this shows is that the principle of instantiation is not a particularly attractive one for the proponent of the potentiality view. That, however, should not be too surprising. When it comes to properties, the potentiality view has much more to think about than actual instantiation; there is more than one intimate relation in which an object may stand to a property. One is instantiation; another is potential instantiation, and potential for another thing's potential instantiation, etc. If a property's existence is a matter of its standing in the right kind of relations to particulars (if, that is, the particular objects determine which properties there are), then we should at least allow for a principle of *potential instantiation*: there is a property of being F just in case something, at some time, instantiates, or has a potentiality to instantiate, or has an iterated potentiality for something else to instantiate that property. This principle sits very comfortably with the second of my suggested ways to accommodate alien objects: I suggested that there are possibly objects with a given uninstantiated property, Q_3, just in case something has an iterated potentiality for there to be objects with that property; now I have suggested in addition that there is an uninstantiated property, say Q_3, just in case something has an iterated potentiality for there to be objects with that property. It follows that a property exists just in case it is possibly instantiated.

Instead of adopting any such principle, the potentiality theorist might adopt a primitivist version of Platonism and say that properties simply exist, and it is an open question which of these properties anything has a potentiality to have. This view sits nicely with the first of my suggested solutions in section 5.5.1, the idea that the possibility of there being things instantiating properties that are actually uninstantiated is grounded in the potentialities of certain abstract objects, the unmanifested properties. Again, it would follow that a property exists just in case it is possibly instantiated, but its possible instantiation would not be what makes it the case that the property exists.

Again, I do not want to endorse any particular view on alien properties. I have aimed to show merely that there is a variety of attractive options for the potentiality theorist to choose from.

5.6 Physical or Metaphysical Possibility?

The arguments of the preceding section will help in addressing a further worry: does the potentiality-based account collapse metaphysical possibility into physical (or nomological, or causal) possibility?

The worry goes as follows. Surely nothing has a potentiality to do what is contrary to the laws of nature. I, for one, have no potentiality to travel faster than light; the laws of nature *constrain* my potentialities. But at the same time we want to say that the laws of nature are contingent: there might have been different laws from the ones that actually hold. So there are possibilities, namely the possibility of the laws being different, for which no object has a corresponding potentiality.

As before, the potentiality theorist may bite the bullet and say: yes, it may have seemed that the laws of nature were contingent; well, we were wrong and it turns out they are necessary. Here she would not have to stand alone. Bird (2007), Ellis (2001), Shoemaker (1980) and Swoyer (1982) have argued that the laws of nature are necessary in the strictest, metaphysical sense. Witness Shoemaker (1998, 74):

> As is suggested by the way "necessity" is used in ordinary speech, and is defined in dictionaries, causal [i.e., nomological] necessity is, pre-theoretically, the very paradigm of necessity. The main theoretical obstacle to according it that status has been the deeply rooted conviction that necessary truths should be knowable a priori. Once that obstacle has been removed, anyone who holds that causal necessity is not really necessity, or has only a second-class kind of necessity, owes us a reason for thinking this.

And again, the potentiality theorist might also go for the catch-all solution and say: nothing *in* the world may have a potentiality for the laws to be different, but the world itself does. Note that this will not be compatible with Bigelow et al. (1992)'s view of the world as 'one of a kind', one of the options I outlined above in section 5.4.

However, as with the worries that I have discussed so far, I would like to explore the piecemeal strategy and see what the potentiality theorist might say to accommodate the apparently conflicting intuition.

What is said depends crucially on the view of laws that is adopted. Accordingly, an important point to note is that the potentiality view does not come with any particular view of lawhood. It shares, to some degree, the anti-Humean outlook of the scientific essentialist views endorsed, in various versions, by Bird, Ellis, Swoyer and Shoemaker, but it is by no means committed to that view. We can distinguish four aspects in dispositional or scientific essentialism. One is essentialism about things or substances: the idea, for instance, that an electron essentially has charge *e*. Another is essentialism about properties: the idea that some properties, e.g. charge, are not mere quiddities but have an essence. (Not all scientific essentialists endorse both forms of essentialism: Bird 2007, for instance, commits himself only to the second.) And a third is the idea that those essences are their causal or dispositional profiles. Fourth, dispositions (and hence properties that are essentially dispositions) are conceived by most of these authors as akin to counterfactuals, dispositions to ... if ..., e.g., the disposition to repel another object if it is in the proximity. It is this counterfactual-like nature of dispositions that makes for 'necessary connections' in nature, on this view. The view of potentialities I am proposing has no such necessary connections built into it: potentialities are akin, I have argued, to possibility, not to the counterfactual conditional. Nor is the view I am proposing committed to both of the two essentialisms. First, do objects or substances have essences? Perhaps. I can understand an object's essence as the limitations of its potentialities. Are an object's potentialities limited by the laws of nature? That is exactly what is at issue. Nothing I have said so far should prejudge the issue. Second, do properties have essences, and in particular, do they have essentially dispositional essences? Yes: a potentiality is the property it is by being the potentiality for a particular manifestation. But a dispositional essence, on my view, does not automatically give us any necessary connections, and need not (though it might) give rise to a law of nature.

So much for the not-so-necessary connection between my proposed view of modality and dispositional essentialism. I believe that my proposed view could be combined even with a metaphysically very 'thin' understanding of the laws, such as the Lewisian best-system account. My view does not, of course, share Lewis's Humean motivation for that account, and the combination of a potentiality view of possibility with the best system account of laws may not be a very natural one. But it may hold some attraction for those who favour a picture of the world as 'open' and full of inherent possibilities (unlike the 'dead world of mechanism', to cite Ellis 2002, 60) while rejecting the idea of a 'governing' instance imposing necessities on the world, even necessities inherent, say, in the nature of properties. The anti-realist attitude towards laws that comes with the best system account may suit such a picture as much as it suits the Humean supervenience theorist. (I am not endorsing this picture, and I am not going to spell it out in any more detail. My concern is merely to point out the great variety of options available to the potentiality theory.)

On the Lewisian analysis, the laws are those regularities that are theorems of the 'best system'; the best system is that one (if there is one) among the deductive systems that truly describe our world which strikes the best balance between simplicity and strength (Lewis 1973, Lewis 1994, 478). Lewis (1983) requires in addition that all the predicates of that system refer to perfectly natural properties only. Probabilistic laws are included, without the (un-Humean) assumption of real objective chances in the world; rather, chances are assigned so as to yield, again, the best combination of strength, fit and simplicity – fit being maximized by chances that best conform to the actual frequencies, while strength and simplicity are a matter of covering as many as possible of those frequencies with as uniform an assignment of chances as possible. (See Lewis 1994.)

The best system according to these standards systematizes the way things actually behave; it might be entirely silent on their potentialities. Alternatively, it might include potentialities in much the same way in which, according to Lewis, it includes objective chance: potentialities may be assigned to any object in such a way and with such degrees as to best capture the actual behaviour of that object in terms of fit (with the object's actual behaviour), strength and simplicity (e.g., stipulating no ad hoc variation in potentialities or their degrees, and ascribing similar potentialities to things that are otherwise similar). Both of these options leave room for things to have potentialities that are not only not captured by the laws but that would, if manifested, go against them. Of course, both of these options are motivated by the Humean prejudice that potentialities aren't among the perfectly natural properties. If you are at all sympathetic to what has been put forward in this thesis so far, you will see little reason for retaining that prejudice.

Suppose, then, that some of the potentialities that things have are among the perfectly natural properties. (Perhaps all the perfectly natural properties are potentialities, a suggestion considered in chapter 3.3.5.) The best deductive system will ascribe potentialities to objects in such a way as to achieve the best balance of fit, strength, and simplicity. Could that balance tolerate the following situation: it is a law (a theorem of the best system) that all Fs are G; yet some thing, *a*, has a potentiality to be an F without being a G? I say it can.

Consider first that some potentialities that are perfectly natural properties may not be instantiated. They may, of course, be potentially instantiated – something may have a potentiality to instantiate them, or an iterated potentiality for something else to instantiate them. But those potentialities need not themselves be perfectly natural properties: if the potentiality to be F is a perfectly natural property, it does not follow that the potentiality to have the potentiality to be F, or the iterated potentiality for something to have the potentiality to be F, is a natural property too. (Recall my hypothetical fundamental properties Q_1 and Q_3 in chapter 5.5.2: even if Q_3 is a perfectly natural property, *a*'s iterated potentiality for there to be something which has Q_3 need not be perfectly natural.) But if a property is never instantiated, then the laws, on the best-system account, may be silent about it. Hence perhaps there are potentialities which are perfectly natural properties yet do not occur in the laws at all. Nonetheless, it is possible that those potentialities should be instantiated by something; and if they were instantiated, they would very likely have to occur in the then-best system, which would therefore be different from our actual best system.

Further, the best system will contain *only* predicates expressing perfectly natural properties, but there is no requirement that the best system say *all there is to say* about the perfectly natural properties. Suppose the best system says that all Fs are Gs (being F and being G may themselves be potentialities, but that does not matter for the present purpose). The best system does not thereby deny that anything has a potentiality to be F without being G; it might even contain a clause saying that some things have the potentiality to be F without being G. As long as that potentiality is never manifested by anything, it remains a law, that is, a theorem of the best system, that all Fs are G. Now, why should such a potentiality never be manifested? Consider my earlier examples for joint potentialities (4.3): a fragile glass may possess the potentiality to break to a very high degree yet be packed in styrofoam with potentialities that 'cut across' the glass's; their joint potentiality for the glass to break is of a rather low degree. If all glasses in the world were always packed in styrofoam, they would rarely, and perhaps never, get to exercise their potentiality to break. The best system, which describes the distribution of properties actually possessed by objects, might then still contain an axiom to the effect that all glasses are unbroken. Of course, glasses and their potentialities to break are unlikely to figure in any best system describing our world. But I see no reason why the same should not happen at the level of perfectly natural properties too: some things may have a (perfectly natural) potentiality to be F without being G but those potentialities may be constantly masked by some equivalent of the glass's styrofoam so as to never manifest. What actually happens is a function not of any one thing's potentialities alone, but of the joint potentialities of all things together. Many potentialities of particular things may well be 'lost' in that their contribution to the overall joint potentialities, and hence to what actually happens, is negligible. The best system, however, is a systematization of what actually happens (including, perhaps, the possession of perfectly natural potentialities); it may ascribe unmanifested potentialities but still not make room for their manifestation.

So much for potentialities and the best system account. A more natural connection, to be sure, would be with with one of my view's anti-Humean allies, an essentialist account of the laws as proposed, for instance, by Ellis (2001) and Bird (2007). I agree with the essentialists that some properties are essentially 'dispositional', or essentially potentialities. According to the scientific essentialists, it is these properties that provide the metaphysical basis for the laws of nature. Exactly how this is spelled out on my view of potentialities depends on how we capture the complex functions that would have to form those essentially dispositional properties' manifestation, a problem noted but not solved in ch. 2.5.

On that view, it will indeed be necessary that those properties which figure in the laws figure in them in exactly the way they do. But it will *not* be necessary that the properties which are actually instantiated are the *only* properties instantiated. I have earlier discussed the possibility of alien properties. Alien properties, if they have dispositional essences, would give rise to alien laws. Or more precisely, properties that are not actually ever instantiated, if they have dispositional essences, give rise to laws that do not actually govern anything that actually happens. Whether those 'alien' laws hold in some vacuous fashion or not at all is not the material point. The point is that, even if the laws are necessary in that nothing could ever go *against* them, what actually happens in the world might have been governed by a partly or entirely different set of laws involving some or only different properties. Bird (2007, 48f.) discusses this view under the label 'weak necessitarianism'. It is a view that nominally counts as necessitarianism: the laws of nature are necessary in that nothing could go against them. They do not necessarily hold non-vacuously: they might have had nothing to govern.

Weak necessitarianism is a view that should be palatable to many critics of necessitarianism: it allows for the possibility of the world being governed by very different sets of laws. Those different laws could not, of course, have contradicted the ones that actually govern the way things happen; they could not require different patterns of interaction between the properties that are actually instantiated (though they might require different patterns of interaction between those properties and other properties that are not actually instantiated). But they could have involved properties that are just like ours, except for being involved in laws of nature that contradict those in which our (actually instantiated) properties figure. Weak necessitarianism thus accounts for the anti-necessitarian intuitions: the world could have been as the anti-necessitarian claims it could in all the phenomenal detail that you may imagine; the only caution we need to take is not to identify the properties that would then have been instantiated with those that actually are.

There are, of course, other accounts of the laws besides those that I have described. Lange (2009) has developed an ingenious variety of the anti-Humean view according to which laws are grounded in primitive 'subjunctive facts' (facts of the form: if p had been the case, then q would have been the case). I cannot address his view here, as doing so would require an account of how potentiality relates to counterfactual or subjunctive facts, an account which I do not yet have.

Another account of the laws, one that is usually thought to be intermediate between Lewis's Humean account and the necessitarian ones along the lines of the scientific essentialists, is David Armstrong's (and similar accounts proposed by Dretske 1977 and Tooley 1977). According to Armstrong (1983), its being a law that all Fs are Gs is a matter of a certain relation, nomic necessitation, holding (contingently) between the universals F-ness and G-ness. Considering that Armstrong's theory is already committed to the existence of universals and their having non-trivial higher-order properties and standing in non-trivial higher-order relations, it is but a small step for the potentiality theorist to ascribe to those universals potentialities to have those properties and stand in those relations. By assumption, the nomic necessitation relation is contingent; for the potentiality theorist, that is to say that the universals that stand in it to each other have the potentiality not to stand in it, and various potentialities to stand in it to other universals instead, hence to be such that the laws are different.

Without endorsing any particular view about the laws of nature, we can see that the potentiality theorist has a variety of options, if she does *not* want to identify metaphysical possibility with physical possibility.

5.7 Necessary Truths and Abstract Objects

It is necessary that 2+2=4; therefore it is possible that 2+2=4. This inference should be validated by the potentiality-based account of possibility (and necessity). It certainly does not follow as easily as it does on a possible-worlds based account: if nothing has an iterated potentiality for not-p, it does not obviously follow that something has an iterated potentiality for p. But that is only the first part of the problem. The second part is this. The potentiality account is committed to finding, for every proposition that is possible true, some object or objects with an iterated potentiality for that proposition to be true. But which object or objects possess a potentiality for it to be the case that 2+2=4? The obvious candidates are the numbers 2 and 4; but are there any such objects, and even if there are, can we really plausibly think of them as possessing potentialities?

It may be thought that it is my restriction to potentiality in the strong sense that makes these questions difficult to answer. I have argued that for any property F (including all 'such-that' properties), if a given object *a* is F, then *a* has a potentiality in the *weak* sense to be F. Now, of course every object is such that 2+2=4, so every object should have a potentiality to be such that 2+2=4. However, I have also argued that weak potentialities with manifestations that do not 'concern' the potentiality's possessor are grounded in joint *strong* potentialities, possessed together with those objects that the manifestation does concern. Thus I possess a weak potentiality to be such that you are in Oxford, but I possess that weak potentiality only because you and I jointly possess a potentiality (in the strong sense) to be such that you are in Oxford. Applying this to my potentiality to be such that 2+2=4, the question is: which are the objects with which I jointly possess this potentiality? In accordance with my schema, it should be the numbers 2 and 4. I have earlier appealed, albeit tentatively, to potentialities of abstract objects, but I have always offered an alternative. When it comes to necessary truths such as the truth that 2+2=4, it is difficult to see what the alternative should be. This brings us back to the very same two problems: first, if there are no abstract objects, then obviously there are no such objects as the numbers 2 and 4 to bear the relevant potentialities here. But then what does? And second, if there are abstract objects, can we really think of them as possessing potentialities?

Let me begin with an easier task: showing that necessity implies possibility when concrete objects are concerned. Suppose it is necessarily the case that *p*; in terms of potentialities, this means that nothing has an iterated potentiality to be such that not-*p*. By contraposition on the entailment of potentiality by actuality (if *x* is F, then *x* has a potentiality to be F), it follows that nothing is such that not-*p*; which entails that *p* is true. If *p* is true, then everything is such that *p*. On the weak conception of potentiality, it would follow immediately that everything has a potentiality to be such that *p*. On the stronger conception, this does not follow. But as long as *p* is 'about' some object(s) that is (or are) in general suited to have potentialities (such as a concrete object), it follows that that object has (or those objects have) a potentiality to be such that *p*. Likewise where *p* is a quantified statement, say 'something is F' or 'everything is F', any concrete object which is F will have a potentiality (on the strong conception) to be such that *p*. (If *p* is of the form 'nothing is F', since that is equivalent to 'everything is not-F', any concrete object that is not-F will have a potentiality to be such that *p*.) Hence in those cases where *p* is of such a form that it predicates something of, or quantifies over, concrete objects, there will be something or some things that has or have a potentiality to be such that *p*; hence it is possible that *p*.

The entailment of possibility by necessity, then, goes through on the potentiality-based account, so long as the proposition that is necessarily true concerns objects which are suited to have potentialities in the first place, such as concrete objects. But what of those necessary truths that concern abstract objects? What about possibiliites concerning abstract objects in general?

As in the previous section, the answer to this question depends on which other metaphysical views we take. The crucial question, of course, is: are there abstract objects?

Suppose we said yes: there are abstract objects. (I have taken this stance so far: I have assumed that properties are universals, that universals exist. But I hope that there is a nominalistic translation of what I said. I do not ultimately want the account of possibility to depend on my assumption.) Do those objects then have potentialities? I do not have a knock-down argument for this, but I can offer the following considerations.

We have already seen the notion of potentiality stretched far beyond the initial examples, dispositions and abilities. We have seen that for an object to be some way or other is sufficient for that object to possess (simultaneously) the potentiality to be that way. When, outside the philosophy room, we ascribe potentialities to objects, we usually ascribe those whose manifestation would constitute a change in the object. There is a good reason for that: if exercising the potentiality would not constitute a change in the object, we can say something stronger and more relevant by ascribing to it, not the potentiality, but the manifestation property itself. That is no reason to deny to objects the potentialities that are being manifested, even if they are being manifested throughout an object's existence. We have also seen that, with concrete objects at least, the properties that an object has necessarily, it also has potentially: if I am necessarily human, then I am human and so I have a potentiality to be human.

In denying potentialities to abstract objects, our motivation would most likely be either that they do not change, or that they do not have any (intrinsic) contingent features. Given what I just said, these two observations point to a rather different conclusion: not that we should not ascribe potentialities to abstract objects, but rather that we have an explanation for why we *do not* usually ascribe potentialities to them: we have stronger and more interesting things to say about them.

I conclude that if there are abstract objects, the potentiality-based account of possibility will fare best by ascribing to them potentialities, and that our initial resistance to such ascriptions should not mislead us into thinking that they express falsehoods. There is a good explanation for that resistance: propositions ascribing potentialities to abstract objects are not false, but whenever we are in a position to ascribe them, we are also in a position to make a more informative statement, and we have a well-known resistance to making the less informative statement.

Suppose, on the other hand, that we said no: there are no abstract objects. What appears to be talk about abstract objects is really a shorthand way of talking about perfectly unsuspicious concrete objects; what appear to be facts about an abstract object such as the number 2 are really facts about concrete, perfectly unobjectionable, objects. They might be particular facts about particular concrete objects that play the role which we thought to be played by abstracta (thus concrete inscriptions or classes of perfectly resembling concrete objects may play the role of universals; and according to Field 1980, regions of space can occupy the number role at least for the limited purpose of Newtonian physics). Or else they might be general facts, perhaps involving universal quantification (as some structuralists think), but quantifying only over the domain of unobjectionably concrete objects. Whichever option is taken, it makes my task even easier: if what appeared to be facts about abstract objects are really facts about concrete objects, then possibly true propositions that are apparently about abstract objects should be replaced with possibly true propositions that really concern concrete objects. The potentialities that account for the possibility of those propositions will then be located in some or all of those concrete objects; there is no special problem of the possibility of propositions about abstracta, since there are (on that view) no abstracta.

5.8 Conclusion

Developing a theory of possibility that is based on potentiality is, I have argued, a fruitful research programme. I have not shown more than this: that it is a promising programme. But I hope to have provided grounds for optimism, and the first steps towards executing the programme. I have shown that it it has the potential to deal with the obvious problem cases – alien objects and properties, different laws of nature, necessary truths – and indeed it has various options available in dealing with them and proving its claim to extensional adequacy.

Of course, I have not nearly spelled out the theory in as much detail as it eventually should have, but I hope to have made a plausible claim that it is at least a serious competitor. There are various gaps in my treatment which I can here only note. Perhaps the most blatant one is that I have said nothing about counterfactuals of either the 'would' or the 'might' variety. Whether the potentiality account can deal with these, and in what way, I leave as a question for another day. A second conspicuous gap is that I have not provided explicit semantics for complex modal expressions.[4] I have been concerned primarily with straightforward metaphysics. Eventually, of course, such a semantic account should be forthcoming.

In the next chapter, I will add a more formal treatment of what is a precondition for extensional adequacy and a necessary condition for the potentiality-based account to hold any promise at all: its structural or formal adequacy. I will give a formalization of potentiality based on what I have said about it in chapter 4, and derive possibility from it based on the informal definition in section 5.1. I will then show that possibility, so defined, has the right kind of logical structure: the structure, that is, of the possibility operator in normal modal logic with axiom T.

I conclude, therefore, that the metaphysics of potentiality that has been motivated and formulated in chapters 2 to 4 provides us with a promising and intuitively compelling account of possibility as being located firmly in the actual worlds and the things that populate it.

[4]In particular, I have not said anything about the 'actually' operator which is sometimes thought to spell trouble for modalists (see Melia 1992). However, Fara and Williamson (2005) argue that in fact it is not modalism but counterpart theory which cannot deal with the 'actually' operator.

Chapter 6

Logic

In this final chapter, I am going to develop a logic of potentiality and show how to derive from it the logic of possibility. The potentiality operator is construed following the model of Wiggins (1976)'s *de re* necessity operator 'Nec'. The derived logic of possibility will be standard: that is, it will be a normal modal logic, and it will validate the inference from actuality to possibility.

A modal logic's normality can be guaranteed by various minimal combinations of axioms and rules of inference (Chellas 1980, p.114-118, provides a list of such combinations). One, though not the standard, option is this: The possibility operator $\Diamond$ must

1. be closed under logical equivalence: if $\vDash \phi \equiv \psi$, then $\Diamond\phi \equiv \Diamond\psi$;
2. be closed under, and distribute over, disjunction: $\vDash (\Diamond(\phi \vee \psi) \equiv (\Diamond\phi \vee \Diamond\psi)$;
3. and in addition validate the theorem $\vDash \neg\Diamond\bot$.

(Cf. Chellas 1980, 118, Theorem 4.5(5).) In addition to yielding a *normal* modal logic, any logic of $\Diamond$ that purports to formalize *metaphysical* possibility must validate the T theorem:

$$\vDash \phi \rightarrow \Diamond\phi$$

I will prove that this is the case, and that therefore my potentiality-based logic of possibility is at least the system **T** as known from standard modal logic. The potentiality-based account thereby shows that it can meet the minimal requirement of formal adequacy.

I am not yet sure whether, or which of, the stronger systems of modal logic will be validated by possibility defined in terms of potentiality. I will provide a conditional proof of the axiom typical of S4, but say nothing about S5. For the time being, I am content to point out that the minimal requirement is met.

6.1 Syntax

6.1.1 Vocabulary

The following are primitive descriptive expressions of the language L:

- a denumerable set of sentence letters $p_0, p_1, p_2, ...$
- a denumerable set of singular n-place predicates for all $n \geq 1$, $F_0, F_1, F_2, ...$
- a denumerable set of plural n-place predicates for all $n \geq 1$, $G_0, G_1, G_2, ...$
- a denumerable set of mixed (singular and plural) n-place predicates for all $n \geq 1$, $H_0, H_1, H_2, ...$
- a denumerable set of individual constants $a_0, a_1, a_2, ...$

- a denumerable set of plural constants $b_0, b_1, b_2, ...$

In addition, there is

- a denumerable set of individual variables $y_0, y_1, y_2...$
- a denumerable set of plural variables $z_0, z_1, z_2, ...$

and the following syncategorematic expressions:

- $\exists, \forall, \neg, \vee, \wedge, \rightarrow, \equiv, \lambda$, POT

as well as the parentheses '(' and ')'.

A singular term is either an individual constant or an individual variable; a plural term is either a plural constant or a plural variable; a term is either a singular term or a plural term.

6.1.2 Formation Rules

There are two kinds of complex expressions of L, sentences and one-place predicates, defined by the following rules:

- Every sentence letter is a sentence.
- If Φ is an n-place predicate with m places for singular terms and $(n - m)$ places for plural terms (where $n \geq m$), $t_1...t_m$ are singular terms and $t_{m+1}...t_n$ are plural terms, $\Phi t_1...t_n$ is a sentence.

  (The clause applies in an obvious way to mixed predicates; 'pure' singular and plural predicates are included as special cases: where Φ is a singular predicate, we have $n = m$; where Φ is a plural predicate, we have $m = 0$. I adopt the convention that the singular and plural terms are ordered so that the predicate is followed first by all the singular terms (if any), which are then followed by all the plural terms (if any).)

- If ϕ is a sentence, then so is $\neg\phi$.
- If ϕ and ψ are sentences, then so are $(\phi \vee \psi), (\phi \wedge \psi), (\phi \rightarrow \psi)$ and $(\phi \equiv \psi)$.
- If ϕ is a sentence and x a variable, then $\exists x\phi$ and $\forall v\phi$ is a sentence.
- If ϕ is a sentence and x a variable, then $\lambda x.\phi$ is a one-place predicate.

  (If x is an individual variable, $\lambda x.\phi$ is a singular predicate; if x is a plural variable, $\lambda x.\phi$ is a plural predicate.)
- If Φ is a one-place predicate, then so is POT[Φ].

  (If Φ is a singular predicate, then so is POT[Φ]; if Φ is a plural predicate, then so is POT[Φ].)

Outer brackets may be dropped. Free and bound variables, open and closed sentences are defined in the usual way.

6.2 Semantics

6.2.1 Models and Assignments

A model is a pair $< D, v >$, where

- D is a (non-empty) set of individuals,
- v is a valuation that assigns values to primitive descriptive expressions as follows:
  - to each individual constant an element of D

- to each plural constant a subset of D[1]
- to each predicate a property or relation of the appropriate type,
- to each sentence letter a proposition.

All well-formed expressions receive values in a model relative to an assignment of values to the variables. An assignment s is a function that assigns to each individual variable an element of D, and to each plural variable a subset of D. Closed sentences and predicates are assigned the same values relative to all assignment functions.

If s is an assignment and x a variable, $s[d/x]$ is the assignment that differs from s at most in assigning d to x.

6.2.2 Semantic Rules

The semantic rules extend the valuation function to give values, relative to an assignment, to complex expressions.

I use t, t_1, t_2 as a metavariables for terms (singular and plural, names and variables), x as a metavariable for variables (individual or plural), Φ, Ψ as metavariables for predicates (singular, plural or mixed), and ϕ, ψ etc. as metavariables for sentences (open or closed).

Let v^s be the function that assigns values to expressions in accordance with v, relative to the assignment function s. Then if ϕ is a primitive descriptive expression, $v^s(\phi) = v(\phi)$; if x is a variable, $v^s(x) = s(x)$.

Let $\Pi(\phi, x, v, s)$ be the propositional function that takes every object d to the proposition $v^{s[d/x]}(\phi)$, where ϕ is a formula, x a variable, v a valuation function and s an assignment. (If x is a singular variable, the function will take elements of D to propositions; if x is a plural variable, the function will take subsets of D to propositions.)

[1] I assign sets to plural constants (and variables), purely for reasons of convenience. Different treatments of the logic of plurals should be compatible with everything I say in this chapter, but I have not checked this.

Here are the semantic rules:

Sentences

- If ϕ is of the form $\Phi t_1 ... t_n$, Φ is a mixed predicate, $t_1 ... t_m$ ($0 \leq m \leq n$) are the singular terms of ϕ and $t_{m+1} ... t_n$ are its plural terms, $P = v^s(\Phi)$ and $d_1 = v^s(t_1), ... d_n = v^s(t_n)$, then $v^s(\phi)$ = the proposition that $d_1, ... d_m$ and the members of d_{m+1}, ..., the members of d_n instantiate P.
  (Again, the clause applies in an obvious way to mixed predicates, but singular and plural predicates are included as special cases.)
- If ϕ is of the form $\neg\psi$, then $v^s(\phi)$ = the negation of $v^s(\psi)$.
- If ϕ is of the form $(\psi_1 \vee \psi_2)$, $(\psi_1 \wedge \psi_2)$, $(\psi_1 \rightarrow \psi_2)$, or $(\psi_1 \equiv \psi_2)$, $v^s(\phi)$ = the disjunction of $v^s(\psi_1)$ and $v^s(\psi_2)$; the conjunction of $v^s(\psi_1)$ and $v^s(\psi_2)$; the material implication from $v^s(\psi_1)$ to $v^s(\psi_2)$; or the biconditional of $v^s(\psi_1)$ and $v^s(\psi_2)$, respectively.
- If ϕ is of the form $\exists x\psi$ and $f = \Pi(\psi, x, v, s)$, then (i) if x is an individual variable, $v^s(\phi)$ = the (infinite) disjunction of $f(d)$ over all $d \in D$; or (ii) if x is a plural variable, $v^s(\phi)$ = the (infinite) disjunction of $f(d)$ over all $d \subseteq D$.
- If ϕ is of the form $\forall x\psi$ and $f = \Pi(\psi, x, v, s)$, then (i) if x is an individual variable, $v^s(\phi)$ = the (infinite) conjunction of $f(d)$ over all $d \in D$; or (ii) if x is a plural variable, $v^s(\phi)$ = the (infinite) conjunction of $f(d)$ over all $d \subseteq D$.

Predicates

- If Φ is of the form $\lambda x.\phi$ and $f = \Pi(\phi, x, v, s)$, then (i) if x is an individual variable, $v^s(\Phi)$ = the property of being a d such that $f(d)$ is a true proposition; (ii) if x is a plural variable, $v^s(\Phi)$ = the property of being objects xx, where each of the xx is a member of D, such that $f(\{d : d$ is one of the $xx\})$ is a true proposition.

- If Φ is of the form POT[Ψ] and $P = v^s(\Psi)$, then $v^s(\Phi)$ = the property of having a potentiality to instantiate P.

I do not assume any particular theory of properties or propositions, but I do assume the following:

A1 An n-place property and an ordered n-tuple of objects (or ordered n-tuple of sets of objects) together uniquely determine a proposition (that is, propositions are Russellian, not Fregean).

A2 All propositions are either true or not true. The negation of a proposition p is true iff p is not true, the disjunction of propositions p and q is true iff at least one of p and q is true, etc.

A3 For any objects xx and proposition p, xx instantiate the property of being such that p is true just in case p is true.

Lemma 1 For every model $M = < D, v >$: where $\phi = \exists x\psi$ and (i) x is an individual variable, $v^s(\phi)$ is true just in case for some $d \in D$, $v^{s[d/x]}(\phi)$ is true; or, if (ii) x is a plural variable, $v^s(\psi)$ is true just in case for some $d \subseteq D$, $v^{s[d/x]}(\psi)$ is true.

Proof. (i) Let x be an individual variable. If $\phi = \exists x\psi$ and $f = \Pi(\psi, x, v, s)$, then $v^s(\phi)$ = the infinite disjunction of $f(d)$ over all $d \in D$. A disjunction, finite or infinite, is true just in case at least one of its disjuncts is. So for at least one $d \in D$, $f(d)$ must be a true proposition. $f(d) = v^{s[d/x]}(\psi)$; so, for at least one $d \in D$, $v^{s[d/x]}(\psi)$ is a true proposition. (ii) Let x be a plural variable. If $\phi = \exists x\psi$ and $f = \Pi(\psi, x, v, s)$, then $v^s(\phi)$ = the infinite disjunction of $f(d)$ over all $d \subseteq D$. A disjunction, finite or infinite, is true just in case at least one of its disjuncts is. So for at least one $d \subseteq D$, $f(d)$ must be a true proposition. $f(d) = v^{s[d/x]}(\psi)$; so, for at least one $d \subseteq D$, $v^{s[d/x]}(\psi)$ is a true proposition.

Note: As illustrated by the proof of Lemma 1, the proofs for singular and plural terms run in exact parallelism. In the following, I will ignore plural terms in my proofs. All of them are easily adapted to cover plural terms and predicates as well, by simply replacing the relevant expressions in accordance with the following schema: for $d \in D$, write $d \subseteq D$; for '*d* instantiates property *P*', write 'the members of *d* instantiate property *P*', etc.

6.3 The System P

6.3.1 Rules and Theorems

The system **P** is the set of its theorems. A sentence ϕ of the language L is a theorem of **P** just in case for all models $M = < D, v >$ and assignments s, $v^s(\phi)$ is a true proposition. I write $\vDash_P \phi$, abbreviated to $\vDash \phi$, for 'ϕ is a theorem of **P**'.

I accept classical logic:

Theorems If ϕ is a truth-functional tautology, then $\vDash \phi$.

(This is guaranteed by the assumptions I made about propositions.)

Modus Ponens If $\vDash \phi$ and $\vDash \phi \to \psi$, then $\vDash \psi$.

Proof. Suppose $\vDash \phi$ and $\vDash \phi \to \psi$. This means that for all models $M = < D, v >$ and assignments s, $v^s(\phi)$ is a true proposition and $v^s(\phi \to \psi)$ is a true proposition, i.e. the material implication from $v^s(\phi)$ to $v^s(\psi)$ is a true proposition. By A2, that material implication is a true proposition just in case, if $v^s(\phi)$ is a true proposition, then so is $v^s(\psi)$. Since $v^s(\phi)$ is, by assumption, a true proposition, $v^s(\psi)$ must be true too. Since M and s were arbitrary, the same goes for every model and assignment: ψ must be assigned a true proposition in every model, relative to every assignment. Hence $\vDash \psi$.

This proof of Modus Ponens should give a feel for how my non-standard semantics validates classical truth-functional logic.

Given any theorem of the form $\vDash \phi \rightarrow \psi$, Modus Ponens may be used to formulate a derived rule of the form: If $\vDash \phi$, then $\vDash \psi$. For convenience, here is a list of such rules, as well as one theorem, of classical logic that will be appealed to in proofs below:

(I) If $\vDash \phi \equiv \psi$, $\vDash \phi \equiv \phi_1$ and $\vDash \psi \equiv \psi_1$, then $\vDash \phi_1 \equiv \psi_1$.

(Closure of $\equiv$ under logical equivalence)

(II) If $\vDash \phi \equiv \psi$, then $\vDash \exists x\phi \equiv \exists x\psi$.

(Closure of $\exists$ under logical equivalence)

(III) $\vDash \exists x(\phi \vee \psi) \equiv (\exists x\phi \vee \exists x\psi)$

(Closure under, and distribution over, disjunction of $\exists$)

(IV) If $\vDash \phi \equiv \psi$, then $\vDash \phi \rightarrow \psi$

(V) If $\vDash \phi \rightarrow \psi$ and $\vDash \psi \rightarrow \chi$, then $\vDash \phi \rightarrow \chi$.

(Transitivity of $\rightarrow$)

(VI) If $\vDash \neg\phi$ and $\vDash \neg\psi$, then $\vDash \phi \equiv \psi$.

(Logical equivalence of all contradictions)

In addition, there are theorems and rules validated by the semantics of POT and λ. The following are schemas for theorems of **P**. First, a rule and a theorem for λ:

Closureλ If $\vDash \phi \equiv \psi$, then $\vDash (\lambda x.\phi(t) \equiv \lambda x.\psi(t))$

(Tλ**)** $\vDash \phi \equiv \lambda x.\phi(t)$, where x is not free in ϕ.

And secondly, a rule and three theorems containing POT:

Closure$_{POT}$ If $\vDash \Phi t \equiv \Psi t$, then $\vDash$POT$[\Phi](t) \equiv$ POT$[\Psi](t)$

(POT$\vee$) $\vDash$ POT$[\lambda x.(\phi \vee \psi)](t) \equiv$ (POT$[\lambda x.\phi](t) \vee$ POT$[\lambda x.\psi](t)$)

(T$_{POT}$) $\vDash \Phi t \rightarrow$ POT$[\Phi](t)$

(POT*C*) $\vDash \neg$ POT$[\lambda x.\bot](t)$

($\bot$ is an abbreviation for an arbitrary contradiction.)

6.3.2 Proofs

I will begin with the proofs for (Closureλ) and (Tλ). They will be very explicit in their appeal to the semantics set out above. After that, my proofs will take a slightly more condensed form. In particular, I will omit reference to the semantic clauses and to assumptions A1 to A3.

Axioms containing POT will be proved by relating the semantics for expressions of L to the metaphysics of potentiality developed in chapter 4.

Proof of Closureλ

Suppose $\vDash \phi \equiv \psi$, i.e. for all models $M = < D, v >$ and all assignments s, $v^s(\phi \equiv \psi)$, the biconditional of $v^s(\phi)$ and $v^s(\psi)$, is a true proposition. By A2, then, for all models $M = < S, v >$ and assignments s, $v^s(\phi)$ is true iff $v^s(\psi)$ is true, i.e. $v^s(\phi)$ and $v^s(\psi)$ are either both true or both false. In particular, for all $d \in D$, $v^{s[d/x]}(\phi)$ is true iff $v^{s[d/x]}(\psi)$ is true, i.e. $v^{s[d/]}(\phi)$ and $v^{s[d/x]}(\psi)$ are either both true or both false.

Take an arbitrary model $M = < D, v >$ and an arbitrary assignment s. Let $f = \Pi(\phi, x, v, s)$, the function that takes every $d \in D$ to the proposition $v^{s[d/x]}(\phi)$, and let $f' = \Pi(\psi, x, v, s)$, the function that takes every $d \in D$ to the proposition $v^{s[d/x]}(\psi)$. Since for every $d \in D$, $v^{s[d/x]}(\phi)$ is true iff $v^{s[d/x]}(\psi)$ is true, it follows that for all d, $f(d)$ is a true proposition just in case $f'(d)$ is a true proposition.

Now, $v^s(\lambda x.\phi)$ = the property of being a $d \in D$ such that $f(d)$ is true. By A3, every $d \in D$ instantiates that property just in case $f(d)$ is true. By exactly parallel reasoning, every $d \in D$ instantiates the property $v^s(\lambda x.\psi)$ just in case $f'(d)$ is true. So, for every $d \in D$: d has $v^s(\lambda x.\phi)$ if and only if $f(d)$ is true; $f(d)$ is true if and only if $f'(d)$ is true; $f'(d)$ is true if and only if d has $v^s(\lambda x.\psi)$; hence d has $v^s(\lambda x.\phi)$ if and only if d has $v^s(\lambda x.\psi)$.

Furthermore, for every term t, $v^s(\lambda x.\phi(t))$ is true just in case $v^s(t)$ instantiates $v^s(\lambda x.\phi)$, and $v^s(\lambda x.\psi(t))$ is true just in case $v^s(t)$ instantiates $v^s(\lambda x.\psi)$. Take a term t and object $d \in D$ such that $d = v^s(t)$. Then $v^s(\lambda x.\phi(t))$ is true just in case d instantiates $v^s(\lambda x.\phi)$; d instantiates $\lambda x.\phi$ if and only if d instantiates $v^s(\lambda x.\psi)$; and d instantiates $v^s(\lambda x.\psi)$ just in case $v^s(\lambda x.\psi(t))$ is true; hence $v^s(\lambda x.\phi(t))$ is true if and only if $v^s(\lambda x.\psi(t))$ is true. Hence the biconditional of $v^s(\lambda x.\phi(t))$ and $v^s(\lambda x.\psi(t))$ is true; hence $v^s(\lambda x.\phi(t) \equiv \lambda x.\psi(t))$ is true.

Since M and s were arbitrary, the same holds for all models and assignments:

$\vDash \lambda x.\phi(t) \equiv \lambda x.\psi(t)$.

Proof of (Tλ)

Take a model $M = < D, v >$, an assignment s, and a variable x such that x is not free in ϕ, and let $f = \Pi(\phi, x, v, s)$. That is, for any $d \in D$, $f(d) = v^{s[d/x]}(\phi)$. Since x is not free in ϕ, $v^{s[d/x]}(\phi) = v^s(\phi)$, and hence $f(d) = v^s(\phi)$, for every $d \in D$.

Now, $v^s(\lambda x.\phi(t))$ is true just in case the object $v^s(t)$ instantiates the property $v^s(\lambda x.\phi)$. $v^s(\lambda x.\phi)$ = the property of being a $d \in D$ such that $f(d)$ is true. By A3, any $d \in D$ instantiates that property if and only if $f(d)$ is true; since $f(d) = v^s(\phi)$, it follows that every $d \in D$ instantiates $v^s(\lambda x.\phi)$ if and only if $v^s(\phi)$ is true; in particular, $v^s(t)$ instantiates $v^s(\lambda x.\phi)$ if and only if $v^s(\phi)$ is true. Hence $v^s(\lambda x.\phi(t))$ is true just in case

$v^s(\phi)$ is true. So the biconditional of $v^s(\phi)$ and $v^s(\lambda x.\phi(t))$ is true; hence $v^s(\phi \equiv \lambda x.\phi(t))$ is true. Since M, s were arbitrary, the same holds for all models and assignments: $\vDash \phi \equiv \lambda x.\phi(t)$.

Proof of Closure$_{POT}$

Suppose $\vDash \Phi t \equiv \Psi t$. Take an arbitrary model $M = < D, v >$ and assignment s and let $P_1 = v^s(\Phi)$, $P_2 = v^s(\Psi)$ and $d = v^s(t)$. Then as a matter of logic (the logic of system **P**), d instantiates P_1 if and only if d instantiates P_2; d's instantiating P_1 is logically equivalent to d's instantiating P_2.

Now suppose that $v^s(\text{POT}[\Phi](t))$ is true. Then d has the potentiality to instantiate P_1. Since d's instantiating P_1 is logically equivalent to d's instantiating P_2, it follows by principle (C2) in chapter 4.7.1 that d has the potentiality to instantiate P_2, and hence that $v^s(\text{POT}[\Psi](t))$ is true. So if $v^s(\text{POT}[\Phi](t))$ is true, then so is $v^s(\text{POT}[\Psi](t))$. By exactly parallel reasoning, if $v^s(\text{POT}[\Psi](t))$ is true, then so is $v^s(\text{POT}[\Phi](t))$. Hence $v^s(\text{POT}[\Phi](t))$ is true if and only if $v^s(\text{POT}[\Psi](t))$ is true; so $v^s(\text{POT}[\Phi](t) \equiv \text{POT}[\Psi](t))$ is true. Since M, s were arbitrary, the same holds for all models and assignments: $\vDash \text{POT}[\Phi](t) \equiv \text{POT}[\Psi](t)$.

Proof of (POT$\vee$)

Take a model $M = < D, v >$ and an assignment s. Let $f(d) = \Pi(\phi, x, v, s)$, $f'(d) = \Pi(\psi, x, v, s)$, and $f''(d) = \Pi((\phi \vee \psi), x, v, s)$. For any $d \in D$, $f''(d)$ is a disjunction with exactly two disjuncts: $f(d)$ and $f'(d)$.

Given principle (D) in chapter 4.7.2, an object $d \in D$ has a potentiality to be such that $f''(d)$, the disjunction of $f(d)$ and $f'(d)$, is true just in case d has a potentiality to be such that $f(d)$ is true or d has a potentiality to be such that $f'(d)$ is true. Take $d \in D$

such that $d = v^s(t)$. Then we can again reason by a chain of equivalences:

v^s(POT[$\lambda x.(\phi \vee \psi)$]$(t)$ is true

iff

d has a potentiality to be such that $f''(d)$ is true

iff

d has a potentiality to be such that $f(d)$ is true or d has a potentiality to be such that $f'(d)$ is true

iff

v^s(POT[$\lambda x.\phi$](t)) is true or v^s(POT[$\lambda x.\psi$](t)) is true

iff

v^s(POT[$\lambda x.\phi$](t))$\vee$ POT[$\lambda x.\psi$](t)) is true.

So: v^s(POT[$\lambda x.(\phi \vee \psi)$]$(t)$) is true iff v^s(POT[$\lambda x.\phi$](t))$\vee$ POT[$\lambda x.\psi$](t)) is true; hence v^s(POT[$\lambda x.(\phi \vee \psi)$]$(t) \equiv$ (POT[$\lambda x.\phi$]$(t) \vee$ POT[$\lambda x.\psi$](t)))) is true. Since M, s were arbitrary, (POT$\vee$) is a theorem.

Proof of (T$_{POT}$)

Take a model $M = < D, v >$ and an assignment s. Let $d = v^s(t)$ and $P = v^s(\Phi)$, and suppose that $v^s(\Phi t)$ is true. Then d instantiates P. By (T) in chapter 4.9, if d instantiates P, it follows that d has a potentiality to instantiate P, and hence that v^s(POT[Φ](t)) is true. So if $v^s(\Phi t)$ is true, then so is v^s(POT[Φ](t)); $v^s(\Phi t \rightarrow$POT[Φ](t)) is true. Since M, s were arbitrary, (T$_{POT}$) is a theorem.

Proof of (POT*C*)

For (POT*C*), I cannot appeal directly to chapter 4, but I can appeal to the connection between potentiality and the modal verb 'can' and the idea that nothing can be such that a contradiction is true.

Take a model $M = < D, v >$ and an assignment s. Let $f = \Pi(\bot, x, v, s)$. Since a sentence that is a contradiction remains a contradiction under any assignment, $f(d)$ will be a contradiction for all $d \in D$. Let $d = v^s(t)$. d cannot be such that $f(d)$ is true, for $f(d)$ is a contradiction. Hence it is not true that d has a potentiality to be such that $f(d)$ is true. So $v^s(\neg\text{POT}[\lambda x.\bot](t)$ is true. Since M and s were arbitrary, the same holds for all models and assignments; (POT*C*) is a theorem.

6.4 The Operator ◇

I now introduce an operator ◇, which can serve as a first approximation to possibility. It is defined as an existential generalization on POT:

(Def◇) $\Diamond\phi =_{df} \exists x\, \text{POT}[\lambda x.\phi](x)$, x the first variable not free in ϕ.

Intuitively, $\Diamond\phi$ says that something has a potentiality to be such that ϕ.

6.4.1 Theorems and a Rule

Closure◇ If $\vDash \phi \equiv \psi$, then $\vDash \Diamond\phi \equiv \Diamond\psi$.

(◇∨) $\vDash \Diamond(\phi \vee \psi) \equiv (\Diamond\phi \vee \Diamond\psi)$

(T◇) $\vDash \phi \rightarrow \Diamond\phi$

(◇*C*) $\vDash \neg\Diamond\bot$

The following corollary captures an intuitive connection between potentiality and possibility: if an object has (or some objects have) a potentiality to Φ, then it is possible that that object Φ's (or that those objects Φ):

Corollary 1 $\vDash \mathrm{POT}[\Phi](t) \rightarrow \Diamond(\Phi t)$

6.4.2 Proofs

Proof of Closure$\Diamond$

Suppose $\vDash \phi \equiv \psi$, and let x be the first variable not free in ϕ or ψ. By (Tλ), we have in addition

$$\vDash \phi \equiv \lambda x.\phi(x)$$

and

$$\vDash \psi \equiv \lambda x.\psi(x)$$

By classical logic (I),

$$\vDash \lambda x.\phi(x) \equiv \lambda x.\psi(x).$$

And then by Closure$_{POT}$, it follows that

$$\vDash \mathrm{POT}[\lambda x.\phi](x) \equiv \mathrm{POT}[\lambda x.\psi](x)$$

By classical logic (II), it follows that

$$\vDash \exists x\, \mathrm{POT}[\lambda x.\phi](x) \equiv \exists x\, \mathrm{POT}[\lambda x.\psi](x),$$

so by (Def$\Diamond$)

$$\vDash \Diamond\phi \equiv \Diamond\psi$$

Proof of (◇∨)

Given (POT∨), we already have

$$\models \text{POT}[\lambda x.(\phi \vee \psi)](x) \equiv (\text{POT}[\lambda x.\phi](x) \vee \text{POT}[\lambda x.\psi](x))$$

for any x, including x the first variable not free in ϕ or ψ. By classical logic (II), it follows that

$$\models \exists x\, \text{POT}[\lambda x.(\phi \vee \psi)](x) \equiv \exists x\, (\text{POT}[\lambda x.\phi](x) \vee \text{POT}[\lambda x.\psi](x))$$

The right-hand side of this equivalence being logically equivalent, by classical logic (III), to $\exists x\, \text{POT}[\lambda x.\phi](x) \vee \exists x\, \text{POT}[\lambda x.\psi](x)$, it follows again by classical logic (I), that

$$\models \exists x\, \text{POT}[\lambda x.(\phi \vee \psi)](x) \equiv (\exists x\, \text{POT}[\lambda x.\phi](x) \vee \exists x\, \text{POT}[\lambda x.\psi](x))$$

Then by (Def◇),

$$\models \Diamond(\phi \vee \psi) \equiv (\Diamond\phi \vee \Diamond\psi)$$

Proof of (T◇)

Take a model $M = < D, v >$ and an assignment s, and suppose that $v^s(\phi)$ is a true proposition. Let x be the first variable not free in ϕ. Then by (Tλ), we know that $v^s(\phi \rightarrow \lambda x.\phi(x))$ is also true; that is, if $v^s(\phi)$ is true, then so is $v^s(\lambda x.\phi(x))$. Given the initial assumption, then, $v^s(\lambda x.\phi(x))$ must be true. By (T$_{POT}$), we know further that $v^s(\lambda x.\phi(x) \rightarrow \text{POT}[\lambda x.\phi](x)$ is true; that is, if $v^s(\lambda x.\phi(x))$ is true, then so is $v^s(\text{POT}[\lambda x.\phi](x))$. Hence from the initial assumption that $v^s(\phi)$ is true, it follows that so is $v^s(\text{POT}[\lambda x.\phi](x))$.

Now let $d = v^s(x)$. Then $v^{s[d/x]}(\text{POT}[\lambda x.\phi](x)) = v^s(\text{POT}[\lambda x.\phi](x))$, hence $v^{s[d/x]}$

(POT[$\lambda x.\phi$](x)) is true. By lemma 1, then, v^s($\exists x$POT[$\lambda x.\phi$](x)) is true; hence (by definition) $v^s(\Diamond\phi)$ is true.

So: if $v^s(\phi)$ is true, then so is $v^s(\Diamond\phi)$; hence $v^s(\phi \to \Diamond\phi)$ is true.

Since M and s were arbitrary, the same holds for all models relative to all assignments: $\vDash \phi \to \Diamond\phi$.

Proof of ($\Diamond C$)

Suppose (for reductio) that for some model $M = < D, v >$ and some assignment s, $v^s(\Diamond\bot)$ was true. By definition, that is to say that v^s($\exists x$ POT[$\lambda x.\bot$](x)) is true, for x the first variable not free in $\bot$. By lemma 1, there must be a $d \in D$ such that $v^{s[d/x]}$(POT[$\lambda x.\bot$](x)) is a true proposition. But by (POTC), we also know that $v^{s[d/x]}$($\neg$POT[$\lambda x.\bot$](x) is true, and consequently that $v^{s[d/x]}$(POT[$\lambda x.\bot$](x) is not true. Hence the initial assumption must be rejected: $v^s(\Diamond\bot)$ cannot be true, and its negation, $v^s(\neg\Diamond\bot)$, must be true.

Since M and s were arbitrary, the same holds for all models and assignments: $\vDash \neg\Diamond\bot$.

Proof of Corollary 1

By (Tλ), we have

$$\vDash \Phi t \equiv \lambda x.(\Phi t)(t),$$

And from that by Closure$_{POT}$ it follows that

$$\vDash \text{POT}[\Phi](t) \equiv \text{POT}[\lambda x.\Phi t](t)$$

By classical logic (IV), we also have the weaker theorem

$$\vDash \text{POT}[\Phi](t) \to \text{POT}[\lambda x.\Phi t](t).$$

Now, take a model $M = < D, v >$ and an assignment s, and suppose that v^s(POT[Φ](t)) is a true proposition. Since POT[Φ](t) $\rightarrow$ POT[$\lambda x.\Phi t$](t) is a theorem, v^s(POT[Φ](t) $\rightarrow$ POT[$\lambda x.\Phi t$](t)) must be true. So if v^s(POT[Φ](t)) is true, then so is v^s(POT[$\lambda x.\Phi t$](t)). We have assumed that the antecedent holds; so it follows that v^s(POT[$\lambda x.\Phi t$](t)) is true. Now take $d \in D$ such that $v^s(t) = d$. Then v^s(POT[$\lambda x.\Phi t$](t)) = $v^{s[d/x]}$(POT[$\lambda x.\Phi t$](x)). Since, given our initial assumption, v^s(POT[$\lambda x.\Phi t$](t)) is true, it follows that $v^{s[d/x]}$ (POT[$\lambda x.\Phi t$](x)) is true. And then by lemma 1, v^s($\exists x$ POT[$\lambda x.\Phi t$](x)) is true; hence, by definition, $v^s(\Diamond\Phi t)$ is true.

So: if v^s(POT[Φ](t)) is true, then so is $v^s(\Diamond\Phi t)$; v^s(POT[Φ](t) $\rightarrow \Diamond\Phi t$) is true. And since M and s were arbitrary, the same holds for all models and assignments: $\vDash$ POT[Φ](t) $\rightarrow \Diamond\Phi t$.

6.5 Possibility: the Operator $\Diamond$*

6.5.1 Introduction: Possibility and Iterated Potentiality

The operator $\Diamond$ is not an adequate expression of metaphysical possibility: it is an existential generalization on non-iterated potentiality, while metaphysical possibility is to be understood as an existential generalization on *iterated* potentiality (or so I have argued). One way of formally rendering that understanding would be by defining an operator for iterated potentiality and then defining the new possibility operator in terms of it. While this will not ultimately be my strategy, it will be instructive to see how it would be done.

An iterated potentiality, I have said, is a potentiality whose manifestation consists in something (perhaps the iterated potentiality's possessor, but perhaps not) having a potentiality (whose manifestation may in turn consist in something having a potentiality, etc.) to be such that p (chapter 4.5). We can count the number of iterations on a potentiality: a potentiality to be F is once iterated, a potentiality for something to have a potentiality to be such that p is twice iterated, etc. Here is a suggestion for formally capturing the notion of an n-times iterated potentiality:

Syntax of POT^n: if ϕ is a sentence and x a variable not free in ϕ, then for any natural number $n \geq 1$, POT$^n[\lambda x.\phi]$ is a predicate.

Definition of POT^n: 1. POT$^1[\lambda x.\phi](t) =_{df}$ POT$[\lambda x.\phi](t)$

2. POT$^{n+1}[\lambda x.\phi](t) =_{df}$ POT$[\lambda x.\exists x($POT$^n[\lambda x.\phi](x))](t))$

We can then define a generalized operator for potentiality that is iterated any number of times, say POT*:

Syntax of POT*: if ϕ is a closed sentence and x a variable not free in ϕ, then POT*$[\lambda x.\phi]$ is a predicate.

Semantics of POT*: If v is a valuation, s an assignment, ϕ a sentence and x a variable not free in ϕ: Let f be the function that takes every natural number $n \geq 1$ to the property $v^s($POT$^n[\lambda x.\phi])$. Then $v^s($POT*$[\lambda x.\phi])$ = the property of possessing at least one of the $f(n)$, for all natural number $n \geq 1$.

The idea that possibility is an existential generalization on iterated potentiality would then be captured by a definition such as the following:

(*) $\Diamond\phi =_{df} \exists x$ POT*$[\lambda x.\phi](x)$, for x the first variable not free in ϕ.

Why, then, am I not going to go for this strategy? Because we can make things easier than this. Consider that, if $\Diamond\phi$ is to be true according to (*), then at least one of the following will be true (by the semantic clause for POT* and the definition of POT^n):

- $\exists x$ POT$[\lambda x.\phi](x)$
- $\exists x$ POT$[\lambda x.\exists x$ POT$[\lambda x.\phi](x)](x)$
- $\exists x$ POT$[\lambda x.\exists x$ POT$[\lambda x.\exists x$ POT$[\lambda x.\phi](x)](x)](x)$

....

In representing an iterated potentiality, we apply an existential generalization to a POT sentence in very much the same way as we do in defining the possibility operator; only in the former case we do so only inside the scope of a λ operator which is itself embedded in the scope of a POT operator. Formally, that difference in treatment seems spurious.

Let us try an alternative treatment then. We can retain the original definition of the possibility operator, $\diamond$, in terms of non-iterated potentiality:

(Def$\diamond$) $\diamond\phi \equiv \exists x$ POT$[\lambda x.\phi](x)$, for x the first variable not free in ϕ.

Substituting $\diamond$ in accordance with this definition, we could rewrite the above sequence of sentences as follows:

- $\exists x$ POT$[\lambda x.\phi](x)$
- $\exists x$ POT$[\lambda x.\diamond\phi](x)$
- $\exists x$ POT$[\lambda x.\diamond\diamond\phi](x)$

....

That is to say, we can understand an iterated potentiality as a potentiality whose manifestation is a possibility. Or indeed, we could leave talk of iterated potentialities completely out of the picture and apply (Def$\diamond$) more thoroughly, thus transforming our sequence of sentences into

- $\Diamond\phi$
- $\Diamond\Diamond\phi$
- $\Diamond\Diamond\Diamond\phi$

....

It seems, then, that we need not define an operator for iterated potentiality to capture that sequence; all we need to iterate is the possibility operator that has already been defined.

In omitting the formalization of iterated potentialities, we are not omitting the iterated potentialities themselves: given the metaphysics developed in chapter 4, iterated potentialities are what makes a sentence of the form $\Diamond\phi$, $\Diamond\Diamond\phi$, etc., true: intuitively, these sentences say that something has a once, twice, or however many times iterated potentiality for it to be the case that ϕ. In giving up the formalization of iterated potentiality, we merely relinquish the opportunity to ascribe an iterated potentiality to a particular object (or to particular objects): we are left only with a device for expressing that *something* has an iterated potentiality. In the long run, we may wish to ascribe iterated potentialities to particular objects, and I have indicated how we should go about doing so. My present concern, however, is the definition of a possibility operator based on the logic of potentiality outlined in the previous sections of this chapter. For that concern it will be perfectly sufficient to iterate $\Diamond$ as defined by (Def$\Diamond$). I will now go on to spell out this idea in detail.

6.5.2 Syntax and Semantics

I begin by introducing an operator $\Diamond^n$, which can be defined inductively as follows:

(Def$\Diamond^n$) 1. $\Diamond^0\phi = \phi$

2. $\Diamond^{n+1}\phi = \Diamond\Diamond^{n}\phi$

Intuitively, $\Diamond^n$ is merely an abbreviation for a sequence of exactly n occurrences of $\Diamond$.

$\Diamond^*$ is a generalization over $\Diamond^n$. Its syntax is obvious:

- If ϕ is a sentence, then so is $\Diamond^*\phi$.

Let $\Xi(\phi, v, s)$ be the function that takes every natural natural number n to the proposition $v^s(\Diamond^n\phi)$. Then we can add the following clause to the semantic rules given in section 6.2.2:

- If ϕ is of the form $\Diamond^*\psi$ and $f = \Xi(\psi, v, s)$, then $v^s(\phi)$ = the disjunction of $f(n)$ over all $n \in \mathbb{N}$.

6.5.3 Theorems and a Rule for $\Diamond^*$

$\Diamond^*$, like $\Diamond$, is closed under logical equivalence and generates exactly similar theorems. Here they are:

Closure$\Diamond^*$ If $\vDash \phi \equiv \psi$, then $\vDash \Diamond^*\phi \equiv \Diamond^*\psi$

($\Diamond^*\vee$) $\vDash \Diamond^*(\phi \vee \psi) \equiv (\Diamond^*\phi \vee \Diamond^*\psi)$

(T$\Diamond^*$) $\phi \rightarrow \Diamond^*\phi$

($\Diamond^*C$) $\neg\Diamond^*\bot$

A corollary parallel to Corrollary 1 can also be proved:

Corollary 2 $\vDash \mathrm{POT}[\Phi](t) \rightarrow \Diamond^*(\Phi t)$

In addition, there is reason to believe that the semantics of $\Diamond^*$ validates the theorem typical of S4:

(4◇*) $\Diamond^*\Diamond^*\phi \rightarrow \Diamond^*\phi$

(4◇*) can be proved conditional on one assumption the truth of which I have not yet been able to establish: the assumption that ◇ distributes over infinite disjunction.

6.5.4 Proofs

My general strategy in proving those theorems containing ◇* that correspond to theorems containing ◇ will be the following.

In a first part, I show that the theorem in question holds when ◇* is replaced throughout by $\Diamond^n$, for every natural number n. This is done by showing (A) that the theorem containing $\Diamond^n$ is valid for $n = 0$; and (B) that for every natural number n, if the theorem is valid for n, it is also valid for $n + 1$. By mathematical induction, (A) and (B) together deliver the desired result, that the theorem in question holds for every value of n. In a second part, then, I use that result to prove the theorem containing ◇*.

As in section 6.3.2, I will begin with a relatively explicit proof, which should demonstrate the kind of reasoning involved in proofs of this kind. The rest of the proofs will be more abbreviated.

Proof of Closure◇*

Part 1: (Closure$\Diamond^n$) If $\vDash \phi \equiv \psi$, then $\vDash \Diamond^n\phi \equiv \Diamond^n\psi$.

(A) $\phi \equiv \psi$ is definitionally identical to $\Diamond^0\phi \equiv \Diamond^0\psi$. Since every formula follows from itself:

If $\vDash \phi \equiv \psi$, then $\vDash \Diamond^0\phi \equiv \Diamond^0\psi$.

(B) Suppose that Closure$\Diamond^n$ already held for a given fixed value of $n \in \mathbb{N}$. Suppose further that $\vDash \phi \equiv \psi$; it follows that $\vDash \Diamond^n\phi \equiv \Diamond^n\psi$. By Closure◇, it follows that $\vDash \Diamond\Diamond^n\phi \equiv \Diamond\Diamond^n\psi$. Then by definition, $\vDash \Diamond^{n+1}\phi \equiv \Diamond^{n+1}\psi$. So:

If $\vDash \phi \equiv \psi$, then $\vDash \Diamond^{n+1}\phi \equiv \Diamond^{n+1}\psi$.

By mathematical induction, (A) and (B) together entail that Closure$\Diamond^n$ holds for every value of $n \in \mathbb{N}$.

Part 2: Closure◇*.

Suppose $\vDash \phi \equiv \psi$. By Closure$\Diamond^n$, it follows that $\vDash \Diamond^n\phi \equiv \Diamond^n\psi$, for every $n \in \mathbb{N}$.

Now take a model $M = < D, v >$ and an assignment s. Let $f = \Xi(\phi, v, s)$ and $f' = \Xi(\psi, v, s)$.

Suppose $v^s(\Diamond^*\phi)$ is true. Since $v^s(\Diamond^*\phi)$ = the disjunction of $f(n)$ over all $n \in \mathbb{N}$, at least one disjunct $f(k) = v^s(\Diamond^k\phi)$ of that disjunction must be true. By Closure$\Diamond^n$, $v^s(\Diamond^k\phi \equiv \Diamond^k\psi)$ is true. It follows that, since $v^s(\Diamond^k\phi) = f(k)$ is true, $v^s(\Diamond^k\psi) = f'(k)$ must be true too. Hence $v^s(\Diamond^*\psi)$, the disjunction of $f'(n)$ for all $n \in \mathbb{N}$ has at least one true disjunct, $f'(k)$, and must therefore itself be true. So: if $v^s(\Diamond^*\phi)$ is true, then so is $v^s(\Diamond^*\psi)$.

By exactly parallel reasoning, if $v^s(\Diamond^*\psi)$ is true, then so is $v^s(\Diamond^*\phi)$.

Taking both directions together, $v^s(\Diamond^*\psi)$ is true iff $v^s(\Diamond^*\phi)$ is true. Hence $v^s(\Diamond^*\phi \equiv \Diamond^*\psi)$ is true. Since M and s were arbitrary, the same holds for all models relative to all assignments: $\vDash \Diamond^*\phi \equiv \Diamond^*\psi$.

Proof of ($\Diamond^*\vee$)

Part 1: ($\Diamond^n\vee$) $\Diamond^n(\phi \vee \psi) \equiv (\Diamond^n\phi \vee \Diamond^n\psi)$

(A) $\Diamond^0(\phi \vee \psi) =_{df} (\phi \vee \psi) =_{df} (\Diamond^0\phi \vee \Diamond^0\psi)$. Hence

$\vDash \Diamond^0(\phi \vee \psi) \equiv (\Diamond^0\phi \vee \Diamond^0\psi)$

(B) Suppose we have ($\Diamond^n\vee$), for a given fixed value of $n \in \mathbb{N}$. By Closure$\Diamond$,

$\vDash \Diamond\Diamond^n(\phi \lor \psi) \equiv \Diamond(\Diamond^n\phi \lor \Diamond^n\psi)$

Further, by ($\Diamond\lor$):

$\vDash \Diamond(\Diamond^n\phi \lor \Diamond^n\psi) \equiv (\Diamond\Diamond^n\phi \lor \Diamond\Diamond^n\psi)$

Applying classical logic (I) to these two formulas, we get

$\vDash \Diamond\Diamond^n(\phi \lor \psi) \equiv (\Diamond\Diamond^n\phi \lor \Diamond\Diamond^n\psi)$

So by (Def$\Diamond^n$),

$\vDash \Diamond^{n+1}(\phi \lor \psi) \equiv (\Diamond^{n+1}\phi \lor \Diamond^{n+1}\psi)$

By mathematical induction, ($\Diamond^n\lor$) holds for all $n \in \mathbb{N}$.

Part 2: Proof of ($\Diamond^*\lor$)

Take a model $M = < D, v >$ and an assignment s, and let $f = \Xi(\phi, v, s), f' = \Xi(\psi, v, s), f'' = \Xi((\phi \lor \psi), v, s)$.

Given ($\Diamond^n\lor$), for all $n \in \mathbb{N}$, $f''(n)$ is true iff $f(n)$ is true or $f'(n)$ is true.

The truth of $v^s(\Diamond^*(\phi \lor \psi) \equiv (\Diamond^*\phi \lor \Diamond^*\psi)$ is proved by proving both directions of the biconditional.

(i) Suppose $v^s(\Diamond^*(\phi \lor \psi))$ is true. Then for at least one $n \in \mathbb{N}$, $f''(n)$ must be true. If $f''(n)$ is true, then $f(n)$ or $f'(n)$ must be true. If $f(n)$ is true, it follows that $v^s(\Diamond^*\phi)$ must be true; if $f'(n)$ is true, it follows that $v^s(\Diamond^*\psi)$ must be true. Since at least one of $f(n)$ and $f'(n)$ must be true, it follows that $v^s(\Diamond^*\phi \lor \Diamond^*\psi)$ is true.

(ii) Suppose $v^s(\Diamond^*(\phi\lor\psi))$ is not true. Then for no $n \in \mathbb{N}$, $f''(n)$ is true; accordingly, for no $n \in \mathbb{N}$ either $f(n)$ or $f'(n)$ is true. So neither $v^s(\Diamond^*\phi)$ nor $v^s(\Diamond^*\psi)$ can be true. Therefore their disjunction, $v^s(\Diamond^*\phi \lor \Diamond^*\psi)$ cannot be true either.

Taking (i) and (ii) together: $v^s(\Diamond^*(\phi \lor \psi))$ is true iff $v^s(\Diamond^*\psi \lor \Diamond^*\psi)$ is true, and so $v^s(\Diamond^*(\phi \lor \psi) \equiv (\Diamond^*\psi \lor \Diamond^*\psi))$ is true. Since M, s were arbitrary, ($\Diamond^*\lor$) is a theorem.

Proof of (T$\Diamond$*)

The proof of (T$\Diamond$*) is trivial: take a model $M = < D, v >$ and an assignment s, and let $f = \Xi(\phi, v, s)$. If $v^s(\phi)$ is true, then by definition $v^s(\Diamond^0\phi)$ is true, hence $f(0)$ is true, hence the disjunction of $f(n)$ over all $n \in \mathbb{N}$ is true; hence $v^s(\Diamond^*\phi)$ is true. So $v^s(\phi \to \Diamond^*\phi)$ is true. M and s were arbitrary, so $\vDash \phi \to \Diamond^*\phi$.

This triviality should not cause any concern, however; given (T$\Diamond$), (T$\Diamond$*) could be proved in a non-trivial way. Proving it in this way simply speeds things up.

Proof of ($\Diamond$*C)

Part 1: ($\Diamond^n$C) $\neg\Diamond^n\bot$

(A) $\Diamond^0\bot$ is identical, by definition, to $\bot$, and $\neg\bot$ is a theorem of classical logic. Hence

$$\vDash \neg\Diamond^0\bot$$

(B) Suppose that we have, for a given fixed value of $n \in \mathbb{N}$, $\vDash \neg\Diamond^n\bot$. Given $\vDash \neg\bot$, it follows by classical logic (VI) that $\vDash \bot \equiv \Diamond^n\bot$. Hence by Closure$\Diamond$, $\vDash \Diamond\bot \equiv \Diamond\Diamond^n\bot$, which is by definition identical to $\vDash \Diamond\bot \equiv \Diamond^{n+1}\bot$.

So for all models $M = < D, v >$ and assignments s, $v^s(\Diamond\bot)$ is true iff $v^s(\Diamond^{n+1}\bot)$ is true. Given ($\Diamond C$), $v^s(\Diamond\bot)$ cannot be true; so $v^s(\Diamond^{n+1}\bot)$ is not true and $v^s(\neg\Diamond^{n+1}\bot)$ is true. Since M and s were arbitrary, it follows that $\vDash \neg\Diamond^{n+1}\bot$.

By mathematical induction, ($\Diamond^n\bot$) holds for every value of $n \in \mathbb{N}$.

Part 2: Proof of ($\Diamond^*\bot$)

Take a model $M = < D, v >$ and an assignment s, and let $f = \Xi(\bot, v, s)$. Then $v^s(\Diamond^*\bot)$ is the disjunction of $f(n)$ over all $n \in \mathbb{N}$. By part 1, for all $n \in \mathbb{N}$, $v^s(\neg\Diamond^n\bot)$

is true, and so $f(n) = v^s(\Diamond^n \bot)$ is not true. Accordingly, $v^s(\Diamond^* \bot)$, the disjunction of $f(n)$ over all $n \in \mathbb{N}$, is not true; hence, $v^s(\neg\Diamond^*\bot)$ is true. Since M, s were arbitrary, $\vDash \neg\Diamond^*\bot$.

Proof of Corollary 2

Take an arbitrary model $M = < D, v >$ and an assignment s, and let $f = \Xi(\Phi t, v, s)$. Then $v^s(\Diamond^*\Phi t)$ is the disjunction of $f(n)$ over all $n \in \mathbb{N}$.

Suppose that $v^s(\text{POT}[\Phi](t))$ is true. By Corollary 1, it follows that $v^s(\Diamond\Phi t)$ is true. Since $v^s(\Diamond\Phi t) = v^s(\Diamond^1\Phi t) = f(1)$, it follows that the disjunction of $f(n)$ over all $n \in \mathbb{N}$ has one true disjunct, hence is itself true. So if $v^s(\text{POT}[\Phi](t))$ is true, then $v^s(\Diamond^*\Phi t)$ is true; $v^s(\text{POT}[\Phi](t) \rightarrow \Diamond^*\Phi t)$ is true. Since M, s were arbitrary, $\vDash \text{POT}[\Phi](t) \rightarrow \Diamond^*\Phi t$.

Note that the converse of Corollary 2 does not hold. A counterexample is provided by replacing Φ with the predicate $\lambda x.\neg\exists y(y = x)$.

Conditional Proof of (4◇*)

To prove (4◇*), we first need to prove two lemmas.

Lemma 2. $\Diamond^{n+1}\phi = \Diamond^n\Diamond\phi$

> *Proof:* By definition, $\Diamond^{n+1}\phi = \Diamond\Diamond^n\phi = \Diamond\Diamond\Diamond^{n-1}\phi = ...$ etc., until a formula is reached where the final superscript is 1. That formula will be a sequence of exactly $n + 1$ occurrences of ◇, followed by ϕ.
>
> Also by definition, $\Diamond^n\Diamond\phi = \Diamond\Diamond^{n-1}\Diamond\phi = \Diamond\Diamond\Diamond^{n-1}\Diamond\phi = ...$, etc., until the superscript 1 is reached. Again, that formula will be a sequence of exactly $n + 1$ occurrences of ◇, followed by ϕ. Hence the two formulas are the same.

Lemma 3. $\Diamond^n\Diamond^m\phi = \Diamond^{m+n}\phi$

> *Proof.* By lemma 2, $\Diamond^n\Diamond^m\phi = \Diamond^{n-1}\Diamond\Diamond^m\phi$, which by (Def◇*) is identical to

$\Diamond^{n-1}\Diamond^{m+1}\phi$. By alternating applications of lemma 2 and (Def$\Diamond$*), the formula can be repeatedly transformed in this way until $\Diamond^{n-n}\Diamond^{m+n}\phi = \Diamond^{0}\Diamond^{m+n}\phi = \Diamond^{m+n}\phi$ is reached.

Second, I need to make the following assumption.

Assumption: For any number n, $\Diamond^n$ distributes over infinite disjunctions, i.e. $\vDash \Diamond^n(\phi_1 \vee \phi_2 \vee ...) \rightarrow (\Diamond^n\phi_1 \vee \Diamond^n\phi_2 \vee ...)$ holds even when there are infinitely many disjuncts ϕ_i.

Why should we believe this assumption to be true? It is true if $\Diamond$ distributes over infinite disjunctions, which in turn will be true if POT does. So the truth of the assumption depends on the behaviour of potentialities with infinitely disjunctive manifestations. Intuitively, the assumption that potentiality distributes over infinite disjunction seems correct: given that we can think of determinables as (entailing, if not identical to) disjunctions of their determinates, and many determinables have infinitely many determinates, my reasoning in chapter 4.7.2 would suggest that potentiality distributes over infinite disjunctions. However, I do not wish to fully endorse this principle until the relation between degrees of potentialities and probability has been more thoroughly investigated. For we can think of having a potentiality as having it to some degree greater than zero. If degrees of potentialities behave like probabilities, then they may be faced with familiar problems: having a non-zero probability does not, or at least not without problems, distribute over infinite disjunctions. I have suggested that we should take that observation to tell against a probabilistic model for degrees of potentialities, rather than against the distribution of potentiality over infinite disjunction (see chapter 4.10). But for the time being, my argument for (4$\Diamond$*) will be hypothetical: if the stated assumption is true, then (4$\Diamond$*) is a theorem. Here is the proof.

Take a model $M = < D, v >$ and an assignment s. For all $n \in \mathbb{N}$, let $f(n) = \Xi(\phi, v, s)$ and $f'(n) = \Xi(\Diamond^*\phi, v, s)$. Suppose $v^s(\Diamond^*\Diamond^*\phi)$ is true. $v^s(\Diamond^*\Diamond^*\phi)$ is the infinite disjunction of $f'(n)$ over all $n \in \mathbb{N}$. Each disjunct of that disjunction, however, is itself an infinite disjunction: the disjunction of $f(n)$ over all $n \in \mathbb{N}$. Schematically, then, $\Diamond^*\Diamond^*\phi$ is a formula of the form

$$\Diamond^0(\Diamond^0\phi \vee \Diamond^1\phi \vee ...) \vee \Diamond^1(\Diamond^0\phi \vee \Diamond^1\phi \vee ...) \vee \Diamond^2(\Diamond^0\phi \vee \Diamond^1\phi \vee ...) \vee ...$$

where '...' indicates continuation *ad infinitum*. If my hypothetical assumption is true, then the n-step possibility operator at the beginning of each disjunct distributes over the infinite disjunction in its scope. Schematically, that is to say that our infinite disjunction is equivalent to this infinite disjunction:

$$(\Diamond^0\Diamond^0\phi \vee \Diamond^0\Diamond^1\phi \vee ...) \vee (\Diamond^1\Diamond^0\phi \vee \Diamond^1\Diamond^1\phi \vee ...) \vee (\Diamond^2\Diamond^0\phi \vee \Diamond^2\Diamond^1\phi \vee ...) \vee ...$$

And using lemma 3, we can then do away with the duplication of diamonds, and contract each pair of consecutive diamonds into one with a new superscript:

$$(\Diamond^0\phi \vee \Diamond^1\phi \vee ...) \vee (\Diamond^1\phi \vee \Diamond^2\phi \vee ...) \vee (\Diamond^2\phi \vee \Diamond^3\phi \vee ...) \vee ...$$

This infinite disjunction, though infinitely repetitive, is logically equivalent to the disjunction of $f(n)$ over all $n \in \mathbb{N}$, and hence to $v^s(\Diamond^*\phi)$. So if the stated assumption holds, then $v^s(\Diamond^*\Diamond^*\phi \rightarrow \Diamond^*\phi)$ is true. Since M, s were arbitrary, $(4\Diamond^*)$ is a theorem, conditional on the assumption of distribution over infinite disjunctions.

6.5.5 Possibility and Necessity

I have introduced two distinct possibility operators, $\Diamond$ and $\Diamond^*$, and shown that both are rightly associated with possibility, since both satisfy the formal requirements for an operator to express possibility: closure under logical equivalence, closure under and distribution over disjunction, implication by actuality (T), and non-applicability to contradictions ($\Diamond C$); $\Diamond^*$ in addition probably yields the characteristic S4 theorem. I have suggested that $\Diamond^*$ expresses metaphysical possibility. My task, it would seem, is finished. But there are two further interesting lines that are worth pursuing for a moment. One, of course, is necessity. Not that I needed to prove anything about it: the normality of my modal logic has been established through the theorems and rules for $\Diamond$ and $\Diamond^*$, so we can be assured that introducing their corresponding necessity operators in the usual way will preserve that normality. Nonetheless, it may be interesting to have a closer look. The second line is the relation between the starred and the non-starred operators. I will discuss that relation in the next section and also prove an interesting theorem, putting to use the observations about necessity.

Let me, then, begin with necessity. We can define two necessity operators in the standard way:

$$\Box\phi =_{df} \neg\Diamond\neg\phi$$

$$\Box^*\phi =_{df} \neg\Diamond^*\neg\phi$$

Since we have seen that both $\Diamond$ and $\Diamond^*$ yield a normal modal logic, we can be assured that the two necessity operators defined in this standard way will behave in an equally normal way. In particular, we know that the following will be theorems and rules of **P**:

Nec If $\vDash \phi$, then $\vDash \Box\phi$

(K) $\vDash \Box(\phi \to \psi) \to (\Box\phi \to \Box\psi)$

(T□) $\vDash \Box\phi \to \phi$

Nec* If $\vDash \phi$, then $\vDash \Box^*\phi$

(K*) $\models \Box^*(\phi \to \psi) \to (\Box^*\phi \to \Box^*\psi)$

(T□) $\models \Box^*\phi \to \phi$

(4□*) $\models \Box^*\phi \to \Box^*\Box^*\phi$

All of these can be derived from the theorems given for ◇ and ◇*, respectively, in the usual way. It is the operator □* that expresses metaphysical necessity.

Note that, once we have defined □, we could from there go on to introduce □* without appeal to ◇*, but in a a way that is entirely parallel to the way in which ◇* has been introduced. Here it is.

First, I define $\Box^n$:

1. $\Box^0\phi = \phi$
2. $\Box^{n+1}\phi = \Box\Box^n\phi$

Intuitively, $\Box^n$ is merely an abbreviation for a sequence of exactly n occurrences of □.

□* is a generalization over $\Box^n$. Its syntax is obvious:

If ϕ is a sentence, then so is $\Box^*\phi$.

Let $\rho(\phi, v, s)$ be the function that takes every natural natural number n to the proposition $v^s(\Box^n\phi)$. Then we could add the following clause to the semantic rules:

(*) If ϕ is of the form $\Box^*\psi$ and $f = \rho(\psi, v, s)$, then $v^s(\phi)$ = the conjunction of $f(n)$ over all $n \in \mathbb{N}$.

I will, however, *not* add that clause to the semantic rules; rather, I will understand □* as defined in terms of ◇*. But importantly, both strategies are equivalent when it comes to the truth conditions for $\Box^*\phi$, though they may (depending on one's view of propositions) be seen as assigning different semantic values to the formula. The following lemma states their truth-conditional equivalence, which I will appeal to in the next section.

Lemma 4 For all models $M = < D, v >$ and assignments s, if $f = \rho(\phi, v, s)$, $v^s(\Box^*\phi)$ is true just in case the conjunction of $f(n)$ over all $n \in \mathbb{N}$ is true.

Proof: Take an arbitrary model $M = < D, v >$ and assignment s, and let $f = \rho(\phi, v, s)$ and $f' = \Xi(\neg\phi, v, s)$; i.e., for all $n \in \mathbb{N}$, $f(n) = v^s(\Box^n\phi)$ and $f'(n) = v^s(\Diamond^n\neg\phi)$.

Then by the definition of $\Box^*$, $v^s(\Box^*\phi) = v^s(\neg\Diamond^*\neg\phi)$ = the negation of $v^s(\Diamond^*\neg\phi)$ = the negation of the disjunction of $f'(n)$ over all $n \in \mathbb{N}$.

The negation of a disjunction is true iff the conjunction of the negations of all its disjuncts is true; hence $v^s(\Box^*\phi)$ is true iff the conjunction of the negation of $f'(n)$, over all $n \in \mathbb{N}$, is true.

Since for every $n \in \mathbb{N}$, $f'(n) = v^s(\Diamond^n\neg\phi)$, the negation of $f'(n)$, for any $n \in \mathbb{N}$, is $v^s(\neg\Diamond^n\neg\phi) = v^s(\Box^n\phi)$. Hence $v^s(\Box^*\phi)$ is true iff the conjunction of $v^s(\Box^n\phi) = f(n)$ over all $n \in \mathbb{N}$ is true.

6.5.6 $\Diamond$ and $\Diamond^*$, $\Box$ and $\Box^*$

Now to the next line: how does $\Diamond^*$ relate to $\Diamond$, and $\Box^*$ to $\Box$?

A model for their relationship can be found in dynamic logic, which has the modal operators $< a >$ and $[a]$ with roughly the following semantics: let a be a programme, that is, an operation which takes states of the world to (potentially different) states of the world. Let s, t stand for such states. Then $< a > \phi$ is true at a state s iff there is some state t such that a can take s to t and at t, ϕ is true; $[a]\phi$ is true at s iff at all states t such that a can take s to t, ϕ is true. The interesting parallel, however, is not with this semantics (which is a species of possible-worlds semantics), but with the further operators $< a^* >$ and $[a^*]$. Where $< a > \phi$ means, informally, that running the programme a could take the world to a state where ϕ (or its semantic value) is true, $< a^* >$ means, informally, that running the programme *a any (nonnegative) number of times* could take the world to a state where ϕ (or its semantic value) is true; equally, mutandis mutatis, for $[a]$ and $[a^*]$.

As with my $\Diamond^*$ and $\Box^*$ operators, the asterisk is used in dynamic logic to indicate iteration. Rather than pursuing this intuitive parellel, however, I would like to point out that my semantics validates the same theorems containing the operators $\Diamond$, $\Diamond^*$ and $\Box$, $\Box^*$ as does dynamic logic, mutatis mutandis. In particular, the system and semantics as given so far validate the 'induction' axiom of dynamic logic, which provides an interesting connection between the starred and non-starred necessity operators. Here is the original (from Harel 1983, 512):

(Ind) $[a^*](\phi \rightarrow [a]\phi) \rightarrow (\phi \rightarrow [a^*]\phi)$

(Ind) can be reformulated, via the interdefinability of $[a]$ and $< a >$ and classical logic, as

(Ind') $< a^* > \phi \rightarrow (\phi \vee < a^* > (\phi \wedge < a > \phi)$

I will now prove the theorem corresponding to (Ind) for the necessity operators of my system **P**. However, to do so, I first need to derive some preliminary results.

Preliminary Results

This proof requires two preliminary results. The first is simple:

Lemma 5 $\Box^{n+1}\phi = \Box^n\Box\phi$

Proof: The proof is exactly parallel to that of Lemma 2 in section 6.5.4.

The second result that I will need is the K axiom for $\Box^n$, for every value of $n \in \mathbb{N}$. The proof will proceed in the familiar way, showing the theorem first to hold for $n = 0$, and then to hold for $n + 1$ whenever it holds for a number n.

(K^n) $\models \Box^n(\phi \rightarrow \psi) \rightarrow (\Box^n\phi \rightarrow \Box^n\psi)$

Proof: (A) For $n = 0$, (K^n) amounts to $\models (\phi \rightarrow \psi) \rightarrow (\phi \rightarrow \psi)$, which is a truth-functional tautology and therefore a theorem of **P**.

(B) Suppose we had, for a given value of n,

$$\models \Box^n(\phi \rightarrow \psi) \rightarrow (\Box^n\phi \rightarrow \Box^n\psi)$$

By **Nec**,

$$\models \Box(\Box^n(\phi \rightarrow \psi) \rightarrow (\Box^n\phi \rightarrow \Box^n\psi)).$$

Furthermore, as an instance of theorem (K),

$$\models \Box(\Box^n(\phi \rightarrow \psi) \rightarrow (\Box^n\phi \rightarrow \Box^n\psi)) \rightarrow (\Box\Box^n(\phi \rightarrow \psi) \rightarrow \Box(\Box^n\phi \rightarrow \Box^n\psi)).$$

By modus ponens, we can detach the consequent and get

$$\models \Box\Box^n(\phi \rightarrow \psi) \rightarrow \Box(\Box^n\phi \rightarrow \Box^n\psi).$$

Another instance of theorem (K) is

$$\models \Box(\Box^n\phi \rightarrow \Box^n\psi) \rightarrow (\Box\Box^n\phi \rightarrow \Box\Box^n\phi).$$

Applying classical logic (V), we get:

$$\models \Box\Box^n(\phi \rightarrow \psi) \rightarrow (\Box\Box^n\phi \rightarrow \Box\Box^n\phi).$$

Which is, by Lemma 5, identical to

$$\models \Box^{n+1}(\phi \rightarrow \psi) \rightarrow (\Box^{n+1}\phi \rightarrow \Box^{n+1}\phi).$$

Proof of (Ind□)

I can now prove the theorem that corresponds to (Ind):

(Ind□) $\Box^*(\phi \rightarrow \Box\phi) \rightarrow (\phi \rightarrow \Box^*\phi)$

Take an arbitrary model $M = < D, v >$ and assignment s and let $f = \rho(\phi, v, s)$ and $f' = \rho((\phi \rightarrow \Box\phi), v, s)$. Then by lemma 4, $v^s(\Box^*(\phi \rightarrow \Box\phi)$ is true iff the conjunction of $f(n)$ over all $n \in \mathbb{N}$ is true; and $v^s(\Box^*\phi)$ is true iff the conjunction of $f'(n)$ over all $n \in \mathbb{N}$ is true.

Now make two assumptions:

(a) $v^s(\phi) = f(0)$ is true;

(b) $v^s(\Box^*(\phi \rightarrow \Box\phi)$ is true.

The following is an instance of theorem (K^n) and holds for every natural number n:

$$\Box^n(\phi \rightarrow \Box\phi) \rightarrow (\Box^n\phi \rightarrow \Box^n\Box\phi)$$

By Lemma 5, it can be rewritten as

$$\Box^n(\phi \rightarrow \Box\phi) \rightarrow (\Box^n\phi \rightarrow \Box^{n+1}\phi)$$

From assumption (b) it follows, as we have seen, that the conjunction of $f'(n)$ over all $n \in \mathbb{N}$ is true; then $f'(n) = \Box^n(\phi \rightarrow \Box\phi)$ must be true for every natural number n. Hence we can detach the consequent of our instance of (K^n) and conclude that for every $n \in \mathbb{N}$, $v^s(\Box^n\phi \rightarrow \Box^{n+1}\phi)$ must be true. So for every $n \in \mathbb{N}$, if $f(n) = v^s(\Box^n\phi)$ is true, then so is $f(n+1) = v^s(\Box^{n+1}\phi)$. Together with assumption (a), this entails by mathematical induction that $f(n)$ is true for all $n \in \mathbb{N}$. So the conjunction of $f(n)$ over all $n \in \mathbb{N}$ is true, and then $v^s(\Box^*\phi)$ is true.

By double conditionalisation, we get: if $v^s(\Box^*(\phi \to \Box\phi)$ is true, then if $v^s(\phi)$ is true, $v^s(\Box^*\phi)$ is true; i.e., $v^s(\Box^*(\phi \to \Box\phi) \to (\phi \to \Box^*\phi))$ is true. Since M, s were arbitrary, the same holds for all models and assignments:

$$\models \Box^*(\phi \to \Box\phi) \to (\phi \to \Box^*\phi)$$

Bibliography

Adams, Robert Merihew. 1974. “Theories of Actuality.” *Noûs* 8:211–231.

Almog, Joseph. 1989. “Logic and the World.” *Journal of Philosophical Logic* 18:197–220.

Armstrong, David. 1983. *What is a Law of Nature?* Cambridge: Cambridge University Press.

Armstrong, David M. 1989. *Universals: An Opinionated Introduction*. Westview.

Ashwell, Lauren. 2010. “Superficial Dispositionalism.” *Australasian Journal of Philosophy* 88:1–19.

Bigelow, John, Ellis, Brian, and Lierse, Caroline. 1992. “The World as One of a Kind: Natural Necessity and Laws of Nature.” *British Journal for the Philosophy of Science* 43:371–388.

Bigelow, John and Pargetter, Robert. 1988. “Quantities.” *Philosophical Studies* 54:287–304.

Bird, Alexander. 1998. “Dispositions and Antidotes.” *The Philosophical Quarterly* 48:227–234.

—. 2007. *Nature’s Metaphysics*. Oxford: Oxford University Press.

—. 2009. “Monistic Dispositional Essentialism.” Manuscript.

Black, Robert. 2000. “Against Quidditism.” *Australasian Journal of Philosophy* 78:87–104.

Blackburn, Simon. 1990. “Filling in Space.” *Analysis* 50:62–65.

Borghini, Andrea and Williams, Neil E. 2008. “A Dispositional Theory of Possibility.” *Dialectica* 62:21–41.

Brown, Mark A. 1988. “On the Logic of Ability.” *Journal of Philosophical Logic* 17:1–26.

Chellas, Brian F. 1980. *Modal Logic: An Introduction*. Cambridge: Cambridge University Press.

Chisholm, Roderick. 1967. “Identity through Possible Worlds: Some Questions.” *Noûs* 1:1–8.

Clarke, Randolph. 2008. "Intrinsic Finks." *The Philosophical Quarterly* 58:512–518.

Coates, Jennifer. 1983. *The Semantics of the Modal Auxiliaries*. London: Croom Helm.

Cross, Troy. 2005. "What is a Disposition?" *Synthese* 144:321–341.

DeRose, Keith. 1991. "Epistemic Possibilities." *Philosophical Review* 100:581–605.

Divers, John and Melia, Joseph. 2002. "The Analytic Limit of Genuine Modal Realism." *Mind* 111:15–36.

Dretske, Fred. 1977. "Laws of Nature." *Philosophy of Science* 44:248–268.

Ellis, Brian. 2001. *Scientific Essentialism.* Cambridge: Cambridge University Press.

—. 2002. *The Philosophy of Nature*. Chesham: Acumen.

Everett, Anthony. 2009. "Intrinsic Finks, Masks, and Mimics." *Erkenntnis* 71:191–203.

Fara, Michael. 2005. "Dispositions and Habituals." *Noûs* 39:43–82.

Fara, Michael and Williamson, Timothy. 2005. "Counterparts and Actuality." *Mind* 114:1–30.

Field, Hartry. 1980. *Science without numbers*. Oxford: Blackwell.

Fine, Kit. 1994. "Essence and Modality." *Philosophical Perspectives* 8:1–16.

—. 1995a. "The Logic of Essence." *Journal of Philosophical Logic* 24:241–273.

—. 1995b. "Ontological Dependence." *Proceedings of the Aristotelian Society* 95:269–289.

—. 1995c. "Senses of Essence." In Walter Sinnott-Armstrong (ed.), *Modality, Morality, and Belief. Essays in Honour of Ruth Barcan Marcus*, 53–73. Cambridge: Cambridge University Press.

—. 2000. "Semantics for the Logic of Essence." *Journal of Philosophical Logic* 29:543–584.

—. 2002. "The Varieties of Necessity." In T.S. Gendler and J. Hawthorne (eds.), *Conceivability and Possibility*, 253–282. Oxford: Oxford University Press.

Frawley, William (ed.). 2006. *The Expression of Modality*. Berlin/New York: Mouton de Gruyter.

Harel, David. 1983. "Dynamic Logic." In Dov Gabbay and F.Guenthner (eds.), *Handbook of Philosophical Logic*, volume II: Extensions of Classical Logic, chapter 10, 497–604. Dordrecht: Reidel.

Holton, Richard. 1999. "Dispositions all the way round." *Analysis* 59:9–14.

Horty, John F. and Belnap, Nuel. 1995. "The Deliberative Stit: A Study of Action, Omission, Ability, and Obligation." *Journal of Philosophical Logic* 24:583–644.

Jacobs, Jonathan D. 2010. "A powers theory of modality: or, how I learned to stop worrying and reject possible worlds." *Philosophical Studies* 151:227–248.

Johnston, Mark. 1992. "How to Speak of the Colors." *Philosophical Studies* 68:221–263.

Jubien, Michael. 2007. "Analyzing Modality." *Oxford Studies in Metaphysics* 3:99–139.

Karakostas, Vassilios. 2007. "Nonseparability, Potentiality, and the Context-Dependence of Quantum Objects." *Journal for General Philosophy of Science* 38:279–297.

—. 2009. "Humean Supervenience in the Light of Contemporary Science." *Metaphysica* 10:1–26.

Kenny, Anthony. 1976. "Human Abilities and Dynamic Modalities." In Juha Manninen and Raimo Tuomela (eds.), *Essays on Explanation and Understanding*, 209–232. Dordrecht: D. Reidel.

Kratzer, Angelika. 1977. "What 'must' and 'can' must and can mean." *Linguistics and Philosophy* 1:337–355.

—. 1981. "The Notional Category of Modality." In H. J. Eikmeyer and H. Rieser (eds.), *Words, Worlds, and Contexts. New Approaches in Word Semantics*. Berlin/New York: de Gruyter.

—. 1991. "Modality." In Arnim von Stechow and Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 639–650. Berlin/New York: de Gruyter.

Kripke, Saul. 1981. *Naming and Necessity*. Oxford: Blackwell, 2nd edition.

Lange, Marc. 2009. *Laws and Lawmakers. Science, Metaphysics, and the Laws of Nature*. Oxford University Press.

Langton, Rae and Lewis, David. 1998. "Defining 'Intrinsic'." *Philosophy and Phenomenological Research* 58:333–345.

Lewis, David. 1970. "How to Define Theoretical Terms." *Journal of Philosophy* 67:427–444.

—. 1973. *Counterfactuals*. Oxford: Blackwell.

—. 1976. "The Paradoxes of Time Travel." *American Philosophical Quarterly* 13:145–152.

—. 1979. "Scorekeeping in a Language Game." *Journal of Philosophical Logic* 8:339–359.

—. 1983. "New Work for a Theory of Universals." *Australasian Journal of Philosophy* 61:343–377.

—. 1986a. *On the Plurality of Worlds*. Oxford: Blackwell.

—. 1986b. *Philosophical Papers, Volume II*. Oxford: Oxford University Press.

—. 1994. "Humean Supervenience Debugged." *Mind* 103:473–490.

—. 1997. "Finkish Dispositions." *Philosophical Quarterly* 47:143–158.

Manley, David and Wasserman, Ryan. 2008. "On Linking Dispositions and Conditionals." *Mind* 117:59–84.

Martin, C. B. 1994. "Dispositions and Conditionals." *The Philosophical Quarterly* 44:1–8.

Maudlin, Tim. 2007. "Why Be Humean?" In *The Metaphysics within Physics*, chapter 2, 50–77. Oxford University Press.

McKitrick, Jennifer. 2003. "A Case for Extrinsic Dispositions." *Australasian Journal of Philosophy* 81:155–174.

Melia, Joseph. 1992. "Against Modalism." *Philosophical Studies* 68:35–56.

Molnar, George. 1999. "Are Dispositions Reducible?" *The Philosophical Quarterly* 49:1–17.

—. 2003. *Powers. A Study in Metaphysics*. Oxford: Oxford University Press.

Mondadori, Fabrizio and Morton, Adam. 1976. "Modal Realism: The Poisoned Pawn." *Philosophical Review* 85:3–20.

Mumford, Stephen. 1998. *Dispositions*. Oxford: Oxford University Press.

Mumford, Stephen and Anjum, Rani Lill. 2010. "A powerful theory of causation." In Anna Marmodoro (ed.), *The Metaphysics of Powers: Their Grounding and Their Manifestation*, 143–159. London / New York: Routledge.

Oppy, Graham. 2000. "*Humean* Supervenience?" *Philosophical Studies* 101:77–105.

Peacocke, Christopher. 1999. *Being Known*. Oxford: Oxford University Press.

Plantinga, Alvin. 1974. *The Nature of Necessity*. Oxford University Press.

Popper, Karl R. 1959. "The Propensity Interpretation of Probability." *The British Journal for the Philosophy of Science* 10:25–42.

Prior, Elizabeth. 1985. *Dispositions*. Aberdeen: Aberdeen University Press.

Sainsbury, Mark. 1997. "Easy Possibilities." *Philosophy and Phenomenological Research* 57:907–919.

Schaffer, Jonathan. 2003. "Is There a Fundamental Level?" *Noûs* 37:498–517.

Shoemaker, Sydney. 1980. "Causality and properties." In *Identity, cause, and mind: Philosophical Essays*, 206–233. Oxford: Oxford University Press (2003).

—. 1998. "Causal and Metaphysical Necessity." *Pacific Philosophical Quarterly* 79:59–77.

Stalnaker, Robert. 1976. "Possible Worlds." *Noûs* 10:65–75.

—. 2003. *Ways a World Might Be*. Oxford: Oxford University Press.

Swoyer, Chris. 1982. "The Nature of Causal Laws." *Australasian Journal of Philosophy* 60:203–223.

Thomason, Richmond and Stalnaker, Robert. 1973. "A Semantic Theory of Adverbs." *Linguistic Inquiry* 4:195–220.

Tooley, Michael. 1977. "The Nature of Laws." *Canadian Journal of Philosophy* 7:667–698.

Wiggins, David. 1976. "The De Re 'Must': a Note on the Logical Form of Essentialist Claims." In Gareth Evans and John McDowell (eds.), *Truth and Meaning: Essays in Semantics*, 285–312. Oxford: Oxford University Press.

Williamson, Timothy. 1994. *Vagueness*. London / New York: Routledge.

—. 1998. "Bare Possibilia." *Erkenntnis* 48:257–273.

—. 1999. "Existence and Contingency." *Proceedings of the Aristotelian Society* 73:181–203.

—. 2000. *Knowledge and its Limits*. Oxford: Oxford University Press.

—. 2002. "Necessary Existents." In A. O'Hear (ed.), *Logic, Thought and Language*, 233–251. Cambridge: Cambridge University Press.

—. 2007. *The Philosophy of Philosophy*. Oxford: Blackwell.

—. forthcoming. "Necessitism, Contingentism and Plural Quantification." *Mind* .