MEASUREMENT AND MODALITY: THE SCALAR BASIS OF MODAL SEMANTICS

Daniel Lassiter

Ph. D. dissertation

Department of Linguistics

New York University

September 2011

(Slight revisions affecting §§3.7-3.8.1, 11/17/11)

<u>Advisors</u>:

Chris Barker (chair)

Seth Yalcin

Anna Szabolcsi

Philippe Schlenker

Chris Kennedy

ACKNOWLEDGEMENTS

I owe thanks to a number of people for help with this project, first and foremost my teachers and advisors. In seminars over the course of several years, Anna Szabolcsi and Seth Yalcin conspired unintentionally to give me the basic idea that developed into this dissertation. Anna's work on weak islands (in particular Szabolcsi & Zwarts 1993) suggested that modals did not make use of the same semantic operations as the quantifiers in terms of which they are usually analyzed; and Seth's work on epistemic adjectives indicated that at least some epistemic modals had a probabilistic semantics which could be framed using standard semantics for gradability (see the subsequently published Yalcin 2010). I started to wonder if the weak island facts could be explained if the latter were uniformly scalar, and we were off. (Weak islands did not make it into the dissertation for reasons of space, but this issue is in the background throughout: see Lassiter 2010b for a brief discussion.)

Philippe Schlenker also made important contributions, in particular by inspiring me to look at Measurement Theory and formal properties of preference; my subsequent investigations led to the core ideas of chapters 2 and 6 respectively. Chris Kennedy, whose ideas are everywhere within, also gave useful comments at various points. His help is gratefully acknowledged.

An important early influence on the ideas presented here was Chris Potts, who introduced me to decision theory and game theory. Robert van Rooij kindly sent me a draft of a paper on measurement-theoretic semantics for gradability which clarified many issues. Other people who gave useful feedback on various parts (the results of which they should not, of course, be held accountable for) include Rick Nouwen, Larry Horn, Benjamin Spector, Julien Dutant, Salvador Mascarenhas, Bert Vaux, Napoleon Katsos, Neil Myler, Chris Collins, Angelika Kratzer, Paul Portner, Pauline Jacobson, Bob Frank, Noah Goodman, Michael Franke, Tom Leu, Tim Leffel, Simon Charlow, Mike Solomon, Tricia Irwin, Peter Klecha, and Luka Crnic. Many thanks to all of them, and apologies to anyone else who I should thank and have failed to. The NYU Semantics Group also deserve a collective thanks, for discussion of various parts and for generally making NYU a great place to do semantics and pragmatics.

Special thanks to Barry Smith and everyone at the Institute of Philosophy at the School of Advanced Study, University of London for generously hosting me with a Visiting Fellowship during the final stages of writing and providing me with much intellectual stimulation and excellent wine.

Some of the material in chapters 2 and 3 were presented at *SALT 20*, the 2010 ESSLLI Student Session, a Yale Syntax Colloquium, and an Institute of Philosophy colloquium, and/or appeared as Lassiter (2010a). I talked about the material in chapter 4 at *ConSOLE XIX* in Groeningen and at King's College, Cambridge, and discussed parts of chapters 5 and 6 at *SALT 21*. Thanks to the conference organizers, editors, reviewers, and audience members who gave thoughtful comments.

Throughout my graduate career Chris Barker was an extremely helpful and generous teacher and advisor. He gave extensive comments on all parts and earlier drafts of this dissertation which saved it from numerous unclarities and errors. Chris was also an inspiring figure in his research, demonstrating that it is possible to be theoretically broad without sacrificing depth, and that disciplinary boundaries are made to be broken. Once again, thank you, Chris.

Finally, I am very grateful to my parents, to whom I owe so much, and above all to my wonderful wife, Emma.

Dissertation Overview

This dissertation argues that modal expressions and gradable expressions in English can and should be treated using the same semantic apparatus. The formal theory that I develop is closely related to standard theories of gradability, comparison, and measurement which are built on abstract degrees. However, I show that certain issues are clarified by taking an algebraic perspective which derives degrees from structures built around binary orders. This approach clarifies a wide range of issues in the semantics of gradability and modality and makes it possible to pose new questions and derive new and far-reaching conclusions regarding the structures underlying epistemic, deontic, and bouletic modals.

Chapter 1 introduces scalar semantics, gradability, and modality in general. I introduce the core hypothesis of the dissertation, that the semantic function of modals is to map propositions to points on a scale and compare them to a threshold value, just as gradable adjectives do in the theory of Kennedy (1997, 2007). I also introduce the best-known framework for modal semantics, according to which modals are quantifiers over possible worlds. After discussing some of the problems with the simplest implementation of this approach, I introduce the influential theory of Kratzer (1981, 1991), which can be seen as a hybrid of quantificational and scalar semantics for modality.

Chapter 2 introduces the Representational Theory of Measurement (RTM, Krantz, Luce, Suppes & Tversky 1971). The essential feature of RTM for our purposes is that it provides a set of mathematical tools for constructing degree-based representations from ordered qualitative structures. Since Kratzer's theory of modality is built around a structure of this kind, using RTM to undergird the degree semantics will make it possible in later chapters to construct scales from her modal semantics and evaluate their suitability as the basis of a semantics for gradable modals in later chapters.

In introducing RTM, I also argue for a new dimension of variation among adjectival scales which is crucial for the analysis of modality in later chapters. In addition to the most familiar parameters — upper- and lower-boundedness — scales can differ in how they interact with the **join** operation. I distinguish in particular between properties such as height and weight, where the degree assigned to the join $x \sqcup_i y$ of two non-overlapping objects x and y is related *additively* to the degree assigned to x and y individually, and properties such as danger and temperature, where $x \sqcup_i y$ has an *intermediate* degree of the property. This distinction is illustrated in (0.1)-(0.2), and discussed in detail in chapter 2, §2.2. (Think of $y \sqcup_i z$, the compound object consisting of y and z, as the result of pouring the contents of bowl y into bowl z, or of putting y and z on a scale together.)

(0.1) INTERMEDIATE PROPERTIES:

a. x is as hot as y.

b. x is as hot as z.

c. **Valid inference**: x is as hot as $y \sqcup_i z$.

(0.2) ADDITIVE PROPERTIES:

a. x is as heavy as y.

b. x is as heavy as z.

c. **Invalid inference**: x is as heavy as $y \sqcup_i z$.

This distinction turns out to be a crucial parameter of variation in the modal domain as well, since individual join $\sqcup_i$ corresponds to disjunction $\vee$ in the domain of propositions. As (0.3)-(0.4) indicate and I argue in detail in later chapters, deontic and bouletic modals behave like intermediate properties in this respect, while epistemic modals behave like additive properties.

(0.3) DEONTIC & BOULETIC MODALS:

a. ϕ is as good/desirable as ψ.

b. ϕ is as good/desirable as χ.

c. **Valid inference**: ϕ is as good/desirable as $(\psi \vee \chi)$.

(0.4) EPISTEMIC MODALS:

a. ϕ is as likely as ψ.

b. ϕ is as likely as χ.

c. **Invalid inference**: ϕ is as likely as $(\psi \vee \chi)$.

This distinction is important for the theory of modality in several ways: in particular, chapter 6 uses the additive/intermediate distinction in scales as a crucial component of a non-monotonic semantics for deontic and bouletic modals. As the contrast between (0.3) and (0.4) already suggests and many more puzzles discussed there demonstrate, non-monotonicity is a desirable feature of a semantic theory for these expressions, and one that is easily stated in scalar terms — but essentially ruled out *a priori* if these expressions are analyzed as quantifiers over possible worlds.

Chapter 3 examines epistemic modality from the perspective of RTM. In the first part of the chapter I focus on the prominent account of Kratzer (1991), which — as Portner (2009) points out — can be taken as the base of a degree semantics for gradable modals such as *likely* and *possible*. However, there are a number of serious logical problems inherent in the standard theory. For example, I prove that Kratzer's theory incorrectly predicts that the inference in (0.4) should be valid, and show that this validity points to several other very damaging incorrect predictions of the theory. I also use the measurement-theoretic apparatus developed in chapter 2 to demonstrate that Kratzer's theory is not able to supply consistent truth-conditions for quantitative expressions of epistemic modality, as in the following examples.

(0.5) a. It is <u>very likely</u> to rain.

b. It is <u>twice as likely</u> to rain as it is to snow.

c. It is <u>half/95% certain</u> to snow.

These issues cannot be patched up lightly, but demonstrate a deep theoretical problem: the "better possibility" relation which Kratzer's theory gives us simply does not contain enough quantitative information to serve as a model for expressions of modality in English, a fact brought out into clear relief by the measurement-theoretic analysis. These and several other problems discussed in chapter 3 raise serious doubts about whether the standard theory can be maintained in its existing form.

The second part of chapter 3 focuses on the epistemic adjectives *possible, probable, likely*, and *certain*, treating them in light of the theory of scale types derived from Rotstein & Winter (2004); Kennedy & McNally (2005); Kennedy (2007) and the discussion of measurement theory in chapter 2. I argue that degree modification data support an account in which the adjectival epistemic modals under discussion are associated with a fully closed, additive scale, a scale type which was discussed with reference to the non-modal adjectives *full* and *closed* in chapter 2. Measurement-theoretic tools make it possible to prove that this scale is equivalent to a *finitely additive probability space*, a slight variant of standard numerical probability. In effect, the degree modification data in combination with evidence for additivity (upward monotonicity) such as (0.4) show that the scale associated with adjectival epistemic modals is ordinary probability. I also show that this approach does not encounter the logical and empirical problems associated with Kratzer's theory: among other benefits, it does not validate (0.4), and it accounts for the distribution of degree modifiers with epistemic adjectives, as in (0.5). I also consider a modification of Comparative Possibility proposed by Kratzer (2012) and show that it does not resolve the problems noted here.

The third section of chapter 3 turns to the epistemic auxiliaries. A possible response to the conclusions of earlier sections might be to restrict the scalar apparatus to epistemic adjectives, but continue to treat the epistemic uses of the modal auxiliaries *must, might, should*, and *ought* using Kratzer's semantics. If so, some bridging rule is needed to capture the logical relations between the epistemic adjectives and auxiliaries. I consider a number of rules along these lines and show that they are empirically flawed in various ways. A simpler and better-motivated way to capture these logical relations is to treat the epistemic auxiliaries as having a probabilistic scalar semantics as well. In the fourth part of the chapter I discuss the problem of continuous sample spaces with respect to the treatment of possibility as non-zero probability, propose an information-theoretic semantics for question-embedding *certain*, and give a treatment of epistemic conditionals which follows closely Kratzer (1986).

Chapter 4 addresses work from experimental psychology on reasoning with *likely*, *probable*, and other expressions of epistemic modality. In the experiments under consideration (Teigen 1988; Windschitl & Wells 1998), subjects' judgments about whether an event is *likely* or *probable* are sensitive to the distribution of alternatives in context: they more likely to rate an uncertain event *A* as "likely" or "probable" if it is presented in contrast to a number of lower-ranked alternatives than if it is presented in contrast to a single event with the same total probability. In addition to their interest as an insight into the way that the vague threshold is set for *likely* and *probable*, these results are important because they have been interpreted as evidence that subjects do not reason probabilistically in making inferences using *likely* and *probable*, and as such are in conflict with the conclusions of chapter 3. Following a lead in Yalcin (2010), I show that this phenomenon is actually expected according to standard semantics for relative adjectives such as *tall*, which are sensitive to alternatives via comparison classes. This interpretation receives support from the fact that these items are sensitive to focus, which introduces a set of propositional alternative which plays the role of comparison classes for *likely* and *probable*.

Since these data have previously been used to argue that humans are incapable of reasoning coherently about uncertainty, this is a matter of some psychological interest. On the linguistic side, the analysis in this chapter highlights the importance of alternatives and focus in the semantics of mid-scalar modals such as *likely* and *probable*. This feature also plays an important role in the analysis of deontic and bouletic modals in chapter 6, where I show that the "weak necessity" modals *ought* and *should*, along with *good* and *want*, are focus-sensitive and are semantically related to the relative adjectives.

Chapter 5 presents five sets of empirical problems for standard quantificational semantics for deontic and bouletic modals, all of which affect Kratzer's theory as well. The first set call into question the upward monotonicity of these expressions, which is built into quantificational semantics; I argue that they are in fact non-monotonic. The second set of puzzles involve the fact that information is often relevant to intuitive judgments of the truth of *ought*- and *should*-sentences, as shown for instance by Kolodny & MacFarlane's (2010) Miners' Paradox. I describe this puzzle and show that it is closely related to other problems involving information-sensitivity of deontic modals and desire verbs already known in the literature (e.g., Goble 1996; Levinson 2003). Kolodny & MacFarlane's proposed solution to the puzzle retains a quantificational semantics by allowing that information gain can manipulate the deontic ordering over worlds non-montonically; this appears to be the only way to resolve the puzzle while retaining a quantificational semantics for modals, and I argue that this tactic is philosophically and methodologically problematic.

The third set of puzzles involve two related facts: first, many deontic and bouletic modals are gradable and form comparatives and equatives; and second, there are more grades of deontic and bouletic modality than quantificational semantics can comfortably capture. I show that the typology is in fact uncannily similar to the typology of gradable adjectives familiar from earlier chapters, and that von Fintel & Iatridou's (2008) enrichment of Kratzer's theory to allow for a three-way split among modals fails to explain empirical differences between intermediate-strength items and universal quantifiers with respect to neg-raising (noted by Horn 1989). The fourth set of problems is specific to Kratzer's theory, which predicts widespread deontic and bouletic incomparability, ruling out even clearly reasonable comparisons such as *It is better to tresspass than it is to murder*. Kratzer's theory also fails to give meaningful truth-conditions to even weak quantitative expressions of obligation and desire such as *It is much better to give your money to charity than to gamble it on sports*.

The final puzzle in chapter 5 is that quantificational theories generally rule out the possibility of conflicts of obligation and desire, even though it is clear — and generally agreed — that such conflicts exist. Kratzer's theory, while maintaining consistency in the face of conflicting requirements, still gets the facts wrong: what we want is a semantics that can render two conflicting obligation or desire sentences *true*, but, as I show, Kratzer's theory makes them both *false*.

Chapter 6 proposes a resolution of the five sets of puzzles and a semantics for the three grades of deontic and bouletic modals exemplified by *may*, *ought/should*, and *must* respectively. In quantificational theories, including Kratzer's, the degree of obligation/desirability of a proposition ϕ is identified with the *maximum* position of any world $w \in \phi$ in the preference order. Following Goble (1996); van Rooij (1999); Levinson (2003) among others, I argue instead that the degree of obligation or desire associated with a proposition is a *weighted average* of the degrees of obligation or desire attached to the individual worlds $w \in \phi$. (This is equivalent to a construct which is well-known in the behavioral and computational sciences under the name "expected utility".) In this semantics, deontic and bouletic modals are associated with an intermediate interval scale, one of the scale types developed for gradable adjectives in chapter 2 to account for intermediate (non-additive) properties such as danger and temperature.

The proposal is shown to resolve the puzzles in chapter 5 by virtue of four differences from standard theories: it is non-monotonic; it interacts in a fine-grained and intuitively correct way with probabilistic information; it allows for deontic and bouletic gradability and comparison without predicting excessive incomparability; and it makes possible a robust notion of conflict of obligation. Quantificational theories are constitutionally unable to capture these features in a theoretically well-motivated way, but they fall out immediately from the scalar treatment of modality proposed here. Along the way I show that this theory resolves of a number of important problems in the logic of obligation and desire, including Ross' Paradox, Jackson & Pargetter's (1986) Professor Procrastinate puzzle, Kolodny & MacFarlane's (2010) Miner's Paradox, and von Fintel & Iatridou's (2008) puzzle about the relationship between "weak" and "strong necessity modals". I also explain the focus-sensitivity of the mid-scalar modals *should, ought, want*, and *good* in the same terms that the alternative-sensitivity of *likely* and *probable* was accounted for in chapter 4.

The dissertation contains a number of points of semantic interest, including new empirical and logical results which cast serious doubt on the dominant theory of modality due to Kratzer, or indeed any attempt to treat modals as quantifiers over possible worlds. Instead, I argue that modals are closely related to gradable adjectives both in their empirical characteristics and in their underlying semantics. This is a substantially new proposal, and in particular the proposal to extend this approach to all epistemic, deontic, and bouletic modals is novel. The conclusion indicates directions for an extension of the account to teleological and circumstantial modals as well as counterfactuals, and briefly considers new connections made available by the scalar approach to modality between the semantics of gradability and modality and empirical and theoretical research in psychology, economics, computer science, and beyond.

How to read this dissertation. The dissertation is written to be a single extended argument, and the best way to read it is to follow the King of Hearts' advice (Carroll 1866: ch.XII).

> "Begin at the beginning," the King said gravely, "and go on till you come to the end: then stop."

I realize, however, that some readers will wish to move as quickly as possible to one or a few topics of particular interest. Here I describe the dependencies between chapters, and then suggest selected chapters for likely groups of readers. The thickness of lines in the dependency chart indicates the degree of the dependency.

- Chapters:

    1: Scales, Gradability, and Modality.

    2: Measurement Theory, Gradability, and The Typology of Scales.

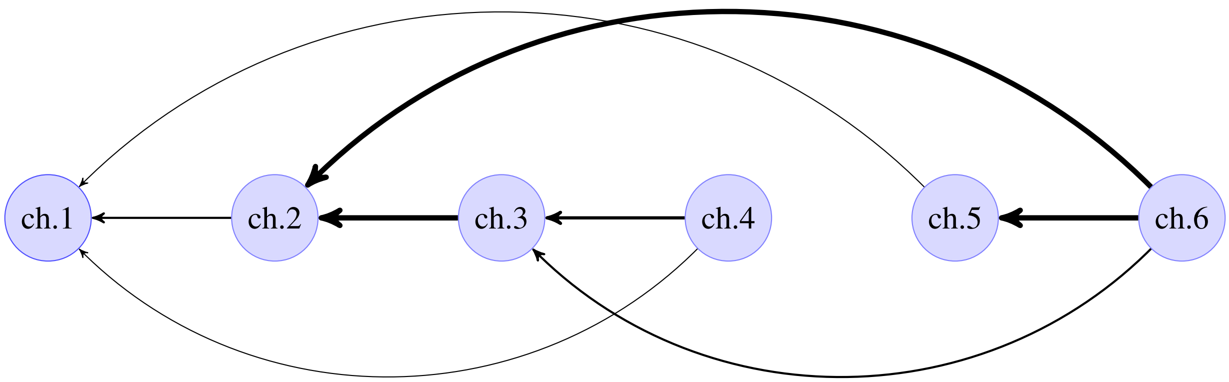
    3: The Structure of Epistemic Modality.

    4: Setting the Standard: *Probable*, Alternatives, and Rationality.

    5: Five Problems for Quantificational Semantics for Deontic and Bouletic Modality.

    6: Scalar Semantics for Deontic Modals and Desire Verbs.

- Dependencies:



- For readers primarily interested in ...
  - *Modality in general*: ch.1 (advanced readers will be able to skim); ch.2 (carefully); ch.3; ch. 4 (optional); chs. 5-6. (That's really everything, sorry; leave out 1, §§3.7-3.11, and 4 if you're really pressed for time.)
  - *Gradability and comparison*: Skim ch. 1, read ch. 2 carefully, ch.3-4, optionally 5-6.
  - *Epistemic modals*: Skim ch. 1, read ch. 2 carefully, read ch.3-4.
  - *Deontic modals and/or desire verbs*: Skim ch. 1, read ch. 2 carefully, read ch.5-6. (Skimming ch.3-4 would be useful but isn't strictly necessary.)

Readers who want to know what is wrong with dominant trends in the formal semantics of modality, without wanting all the details of my proposed solutions, may simply read ch.2-3 and 5. (2 is necessary because ch.3 will not be intelligible without it.) Chapter 5 in particular is written so as to be minimally dependent on the rest of the dissertation; it can be read on its own as a laundry list of interrelated empirical failings of the standard treatment of deontic and bouletic modals as quantifiers over possible worlds.

TABLE OF CONTENTS

Chapter 1

Scales, Gradability, and Modality

1.1 Introduction and Motivation

Recent work on modality in formal semantics (Yalcin 2007, 2010; Portner 2009; Lassiter 2010a) has highlighted the fact that many modal expressions are GRADABLE; for example, they accept at least some degree modifiers and can take part in comparatives and equatives. Although these papers have discussed epistemic modal adjectives for the most part, gradability in the modal domain goes beyond adjectives and beyond epistemic modals; some examples are given in (1.1).

(1.1) Gradability among Modals

a. How necessary is it to marinade meat before making jerkies?[1] (Degree questions)
b. Bill wants to leave as much as Sue wants to stay. (Equatives)
c. I need to go on vacation more than I need to finish this work. (Comparatives)
d. There are situations in which concerns of autonomy ought very much to matter.[2] (Degree modification)
e. It is 95% certain that our team will win. (Measure Phrases)

The modal expressions in these examples cut across syntactic categories, including adjectives *necessary, certain*, main verbs *want, need*, and the (quasi-)auxiliary *ought*.

The examples in (1.1) closely resemble examples of gradability among the better-studied classes of non-modal gradable adjectives and verbs, for example:

(1.2) Gradability among Adjectives

a. How angry can I make my teacher and still get an A? (Degree questions)
b. The Gettysburg Memorial is as old as the Eiffel Tower. (Equatives)
c. My child is cleverer than yours. (Comparatives)
d. This carton of milk is almost empty. (Degree modification)
e. Mary is six feet tall. (Measure Phrases)

(1.3) Gradability among Verb Phrases

a. How much do you like chocolate? (Degree questions)
b. John likes chocolate as much as Mary does. (Equatives)
c. I loathe *Battlefield Earth* more than any other movie. (Comparatives)
d. Harriet has almost finished her art project. (Degree modification)
e. We walked six miles before finding a gas station. (Measure Phrases)

1 http://cooking.stackexchange.com/questions/11299/how-necessary-is-it-to-marinade-meat-before-making-jerkies
2 http://prawfsblawg.blogs.com/prawfsblawg/2010/10/the-progressive-commitment-to-pornography.html

The standard approach in current formal semantics is to tie facts about gradability and comparison in the adjectival and verbal domain to SCALES — that is, to abstract representations of measurement to which gradable expressions relate their arguments. Scales are assumed to be composed of DEGREES which are partially or totally ordered. Roughly, then, *clever* is an expression which relates people to their degrees of cleverness; *loathe* is an expression which relates pairs of individuals (x,y) to the degree to which x loathes y; and so on.

This dissertation considers gradability, comparison, and other evidence for scalar semantics in the modal domain. I develop a new semantics for epistemic, deontic, and bouletic modals which is very closely related to standard scale-based theories of the semantics of non-modal gradable expressions. In particular, modals are analyzed as expressions which relate their propositional arguments to points on a scale, just like gradable adjectives.

I also argue that some modal expressions which do not show evidence of gradability — notably the modal auxiliaries *may, might*, and *must* — can be shown to have a semantics built around scales nonetheless. That is, their logical relations to expressions which are gradable, and the entailments that they license, are mysterious if these items have a quantificational semantics while other modals have a scalar semantics. Implementing this idea requires making a careful distinction between semantic **scalarity** and grammatical **gradability**, which is discussed in the next section of this introductory chapter.

As a general theory of modality, the approach developed here is novel and quite different from standard approaches to modal semantics which treat most or all modals as expressing quantification over possible worlds. In addition for providing a natural account of the gradability of many modals as illustrated in (1.1) — which quantificational theories do not — the logical behavior of modal expressions on the scalar alternative is quite different from the behavior of quantifiers. Most obviously, quantificational theories are predict that all modals are **upward monotonic**. I give a variety of arguments for the conclusion that, while epistemic modals are indeed upward monotonic, deontic and bouletic modals are **non-monotonic**. Unlike quantificational semantics, the scalar approach makes it possible to model monotonic and non-monotonic modalities alike, and to explain the difference using a parameter of variation which is also reflected in the semantics of non-modal gradable adjectives.

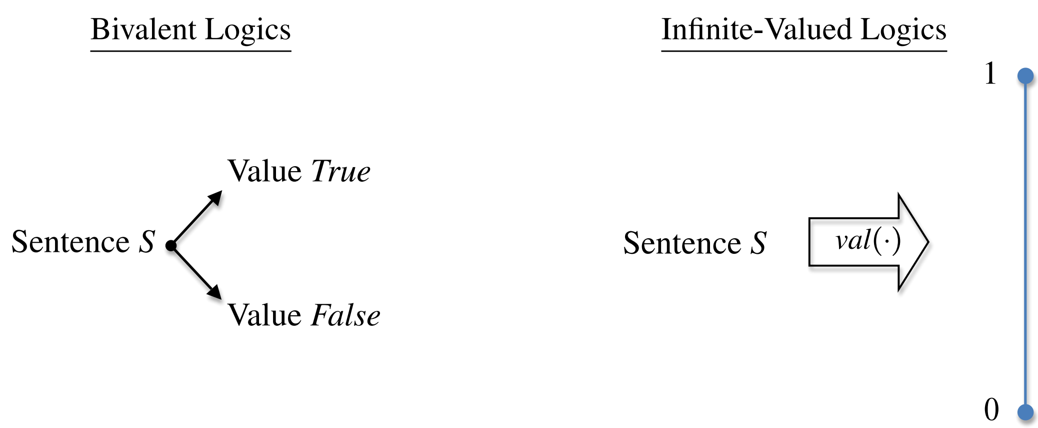
Chapters 1 and 2 give the necessary theoretical and technical background on scalar semantics, quantificational and hybrid semantics for modality, and Measurement Theory. In chapters 3-6 I analyze the behavior of modals as scalar expressions, using entailment data, corpus data, and experimental results to assign these expressions to scale types which are independently attested in the semantics of gradable adjectives. I also show that, in a variety of domains, quantificational theories make incorrect predictions about the logical behavior of epistemic, deontic, and bouletic modals, while the predictions of the scalar alternative that I develop are borne out.

1.2 Scalarity and Gradability

1.2.1 What is Scalar Semantics?

Scalar approaches to the semantics of various domains have gained increasing popularity in formal semantics in recent years, and have been employed in the analysis of gradable adjectives, telicity in verbs, prepositional phrases, common nouns, and elsewhere. Traditionally it has been assumed that the semantic effect of expressions is to partition their domain into two values, an EXTENSION and an ANTI-EXTENSION. In the simplest case, scalar representations can be seen as an enrichment of the classical assumption to allow for many values, often an infinite number; the scale is just the collection of all possible values of the representation, along with an ordering on the values. Two-valued representations can be re-captured from scalar representations, when appropriate, by the use of THRESHOLD VALUES.[3]

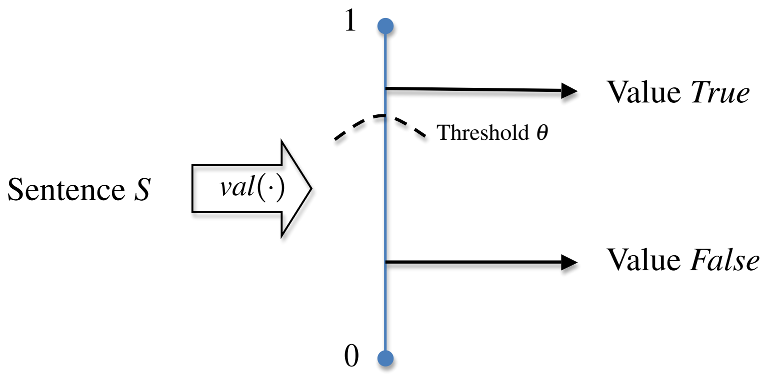
For instance, in bivalent semantics the valuation function *val* maps well-formed and meaningful sentences to one of two values, True and False. A well-known (though controversial) proposal for a scalar enrichment of truth-values is fuzzy logic, where to the range of *val* is not $\{True, False\}$ but the infinite subset $[0,1]$ of the real numbers (Zadeh 1965, 1978). Fuzzy logic is just one of a variety of infinite-valued logics with this basic form.



Infinite-valued logics do not necessarily reject bivalence completely, though; it is always possible to recover bivalent truth-values from a linearly ordered infinite-valued representation by utilizing thresholds. A threshold is a distinguished value on the scale which is used to partition the domain of *val* into two subsets, which (in the simplest case) can be identified with the bivalent truth-values *True* and *False*. I will generally use θ as a variable over possible threshold values. So, for example, if θ is (arbitrarily) set at .7 then we have the following three-step mapping from sentences to bivalent truth-values:

3 Three- or more-valued representations can also be modeled in this way: for instance, for a three-valued scalar semantics for gradable adjectives with an intermediate borderline area we need two thresholds.

Three-Step Bivalent Logic



Here bivalent truth-values are determined by first mapping a sentence to a value in $[0,1]$, then establishing a threshold value, and finally by comparing the value of the sentence to the threshold in order to determine whether it should count as True or False relative to that threshold.

The best-known variety of scalar semantics among linguists applies this approach to gradable adjectives, which pick out properties of individuals, events, or states. In the simplest case on which we will focus, the scalar treatment of properties mirrors the fuzzy-logic approach to truth-values very closely. That is, classically properties have been though of as sets of objects: "x has property P" is true if and only if $x \in A$, where A is a set containing all and only the individuals or objects which have property P. Just as classical bivalent logic presupposes that all sentences can be assigned one of two values (True or False), the classical theory of properties presupposes that individuals are related to properties in one of two ways — they are either in the set or not:

(1.4) $[\![tall]\!]^{\mathcal{M},w,g} = \lambda x_e[x \in \mathbf{tall}]$, where $\mathbf{tall} \subseteq \mathbf{D}_e$.

Scalar semantics for properties generalizes this approach by allowing that adjectives and other property-denoting expressions may map objects not only to the values True (in the set) or False (not in the set), but to a possibly infinite range of values called **degrees**. Unlike the case of truth-values, perhaps, this approach is widely thought to yield reasonable results in the case of gradable adjectives: we have a clear intuition that many properties are graded in this way.

One well-analyzed example is the property of height, as instantiated e.g. by the adjectives *tall* and *short*. Individuals can be *tall* to various degrees; on the variety of scalar semantics that we will focus on throughout this dissertation, this intuition is explained by treating *tall*, not as a function from individuals to truth-values as in (1.4), but as a function from individuals to degrees of height as in (1.5).

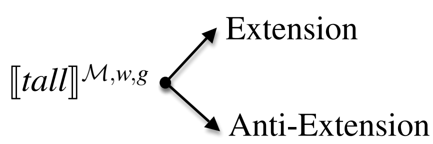
(1.5) $[\![tall]\!]^{\mathcal{M},w,g} = \lambda x_e[\mathbf{height}(x)]$,
where $\mathbf{height}(x)$ is a function from $\mathbf{D}_{tall} \subseteq \mathbf{D}_e \rightarrow [0,\infty)$.

This account — due in particular to Bartsch & Vennemann (1973); Kennedy (1997, 2007) — captures the basic features of gradability in a straightforward way.[4] The essential idea is that gradable

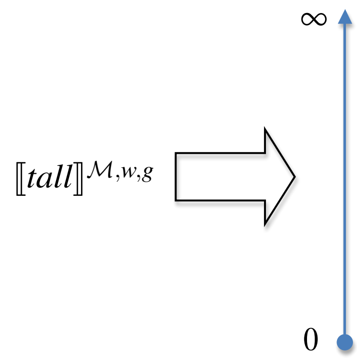
4 It is not universally accepted, though; for enrichments, alternatives, and discussion, see among others von Stechow

adjectives are **measure functions**, i.e. functions from objects (etc.) to their positions on a scale. The difference between the set-theoretic and scalar conceptions of properties is essentially the same as the contrast between bivalent and infinite-valued approaches to truth-values:

Set-theoretic Approach

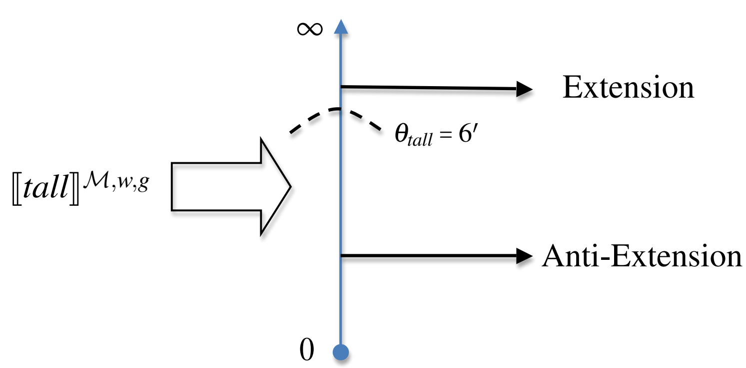


Scalar Approach



Despite the fact that height is a property that comes in degrees, *tall* and *short* do sometimes function to partition their domains into (at least) two sets, as in the classical approach. In addition to *Mary is tall to degree d*, we can just say *Mary is tall* or *Mary is not tall*. This unmodified form of these adjectives is called the POSITIVE FORM. In the measure function analysis, the denotation of the positive form is calculated by a three-step procedure. First, we find the height of the individual argument of the adjective; second, we establish a threshold value; and finally we compare the individual's height to the threshold value. Supposing (arbitrarily) that the threshold for counting as *tall*, θ_{tall}, is 6 feet:

Three-step Set-theoretic Representation



How the threshold value is determined for the positive form of gradable adjectives is a much-discussed and complex issue; at a minimum, discourse context, world knowledge, and lexical

1984; Kennedy 1997, 2001; Heim 2001, 2006; Schwarzschild & Wilkinson 2002; Schwarzschild 2004. The choice among these accounts is important, but does not matter for our purposes here; I adopt this approach because of its simplicity and ease of integration with the measurement-theoretic perspective to be introduced shortly.

semantics play a role. There is also considerable debate about whether there is a unique threshold value operative in a given context; although I will assume for simplicity's sake that there is, everything I say here could easily be generalized to allow for indeterminate, fuzzy, or probabilistic threshold values.

A point that bears highlighting here is that there is nothing special about gradable adjectives with respect to this construction. Many natural language expressions are standardly analyzed as partitioning their domains into two sets: their extension and its complement, True and False, etc. In principle, the extension of any expression with these characteristics could turn out to be determined by means of a scalar representation along the lines just sketched. As long as there is some way of determining a scale, a measure function, and a threshold value, it is always possible to use scalar representations to determine classical set-theoretic/bivalent extensions. Of course, we have to ask on a case-by-case basis whether there is evidence for a scalar representation.

This dissertation is essentially an extended argument for the usefulness of this approach in a domain to which it has rarely been applied, the analysis of modality. I will argue that epistemic, deontic, and bouletic modals have a semantics built on scales — not just gradable modal adjectives and verbs such as *likely*, *obligatory*, *need*, and *want*, but also non-gradable modal auxiliaries such as *may* and *must*, and part-time gradable modals such as *ought* and *should*. The claims is that modals denote (or have denotations which make crucial use of) measure functions on propositions, i.e. functions that map propositions to points on a scale. As in the case of scalar semantics for other expression-types, I will argue that modal sentences get their truth-values by comparing the position of their propositional argument on the relevant scale to a threshold value determined by a combination of lexical semantics and discourse context. A rough idea of the kind of truth-conditions of that modal sentences receive on this proposal can be seen in (1.6):

(1.6) a. $[\![\phi \textit{ is likely}]\!]^{\mathcal{M},w,g} = 1$ if and only if $\textbf{likely}(\phi) \geq \theta_{likely}$
b. $[\![\phi \textit{ must be the case}]\!]^{\mathcal{M},w,g} = 1$ if and only if $\textbf{must}(\phi) \geq \theta_{must}$

So, for example, *It is likely to rain* will be true just in case the measure function denoted by *likely* maps the proposition *it rains* to a point on the appropriate scale which meets or exceeds the relevant threshold value.

This is a quite different approach from the standard analysis of modals as quantifiers; I will argue that the scalar theory makes better predictions in a variety of domains. A major source of inspiration for my theory will be the extensive literature on the semantics of gradable adjectives. In numerous ways, English modals resemble gradable adjectives in the details of the scale structure and their interaction with operators of various kinds, as I will show.

1.2.2 What is Gradability?

Chapters 3-6 will argue that some modal expressions that are not gradable nevertheless have their semantic effect via a three-step process involving scales, measure functions, and threshold values. Obviously this invites the questions: what is gradability, what distinguishes the gradable expressions from the non-gradable ones, and how is it possible to be scalar without being gradable?

As I will use the term, an expression is GRADABLE if it interacts with other linguistic expressions whose grammatical function is to manipulate the threshold value. Some examples are in (1.7):

(1.7) a. Joan is <u>5 feet</u> tall.

b. Harry is tall<u>er than Larry</u>.

c. Sam is <u>very</u> tall.

Although the method of deriving a classical extension for *tall* sketched in the last section looks rather roundabout, the intermediate steps come in handy when we are called upon to deal with threshold-manipulating operators like *very* and *5 feet*. The set of people who are *very tall* is a subset of the ones who are *tall*, in any context; in many contexts, the set of people who are *5 feet tall* will be a superset of those who are *tall*.

By using a three-step process involving scales and threshold values to determine a classical extension for these expressions, we can account for the differences between *tall*, *very tall*, and *5 feet tall* in a straightforward way: *very* and *5 feet* temporarily change the threshold value which is used to determine a classical extension. So, for example, even if the global value of θ_{tall} is 6 feet, (1.7a) will come out true if Joan's height is at least 5 feet, because θ_{tall} has been reset to 5 feet for the purpose of evaluating this expression. A compositional implementation of this approach will be given in the next section.

As this discussion implies, a precondition for an expression's being gradable is that it must be associated with a scale and a threshold value: otherwise the threshold-manipulating operators would have nothing to operate upon. However, the opposite implication does not necessarily hold — logically, there could be scalar expressions which are not gradable because they have a threshold which cannot be manipulated grammatically. If such expressions exist, they determine their extensions using scales as an intermediary, but have threshold values which are either fixed once and for all, or sensitive to contextual factors but not to grammatical manipulation.

This distinction between scalarity and gradability, though subtle, is important for the theory of modality, I will argue. Although evidence that an expression *U* is gradable is *ipso facto* evidence that *U* detemines its extension using scales as an intermediary, a lack of evidence for gradability does not necessarily imply that scales are not implicated in *U*'s semantics. There is indirect but compelling evidence that scales and threshold values are implicated in the semantics of epistemic, deontic, and bouletic modals, even those which do not combine with operators which manipulate the threshold value.

1.3 Semantics of Gradable Adjectives and the Typology of Scales

This section gives a quick sketch of how the scalar analysis of gradable adjectives is implemented compositionally, more details of which can be found in Kennedy (1997, 2007). I also briefly discuss the semantics of the positive form, vagueness, the role of comparison classes, and the several types of scales which have been shown to be relevant in the semantics of gradable adjectives. Since adjectives provide the best-studied class of gradable expressions, this treatment is the gold standard for scalar semantics; many details, including the discussion of adjective type, scale type, and comparison classes, will play an important role in the scalar semantics of modality developed in later chapters.

1.3.1 Compositional Implementation of the Scalar Analysis

The measure function analysis of gradable adjectives assumes that, in addition to the standard e for individuals, t for truth-values, and s for worlds, there is a fourth basic type d (for "degree"). *Degrees*, in approaches of this type, are usually thought of as abstract representations of measurement organized into linearly ordered scales. Formally, scales are structures at least as rich as $\langle D, \leq \rangle$, where $\leq$ is a reflexive, transitive, and antisymmetric binary order.[5] It is usually assumed that the ordering is connected, and that it is dense for at least some expressions, and possibly all (Fox & Hackl 2006; Nouwen 2008). When they are connected and dense, scales with this abstract structure can also be thought of as intervals on the real numbers $\mathbb{R}$ (e.g., Kennedy & McNally 2005, though I will argue below that this identification is somewhat misleading).

While non-gradable adjectives like *British* and *geological* continue to be of type $\langle e,t \rangle$ in degree semantics, gradable adjectives like *tall* and *happy* which take individual arguments are treated as functions of type $\langle e,d \rangle$, i.e. functions from individuals to degrees. The general semantic form of a gradable adjective A is

(1.8) $[\![A]\!]^{\mathcal{M},w,g} = \lambda k_{\alpha}[\mathbf{A}(k)]$

where $\mathbf{A}$ is the measure function appropriate to the adjective A and k is a variable of type α, as appropriate for the adjective in question. *Tall*, for example, expresses a function which takes an argument of type e and returns that individual's degree of height – effectively, to a real number in the range $[0, \infty)$. (Here and throughout, italicized words and phrases like *tall* and *five feet* represent English expressions, while boldfaced expressions like **tall** and **5 feet** represent their model-theoretic translations.)

(1.9) $[\![tall]\!]^{\mathcal{M},w,g} = \lambda x_e[\mathbf{height}(x)]$

Kennedy's (1997) measure function analysis treats the comparative morpheme as a three-place relation between measure functions, degrees, and individuals:

(1.10) $[\![more/-er]\!]^{\mathcal{M},w,g} = \lambda A_{\langle e,d \rangle} \lambda d_d \lambda x_e[\mathbf{A}(x) \geq d]$

(1.10) requires a syntactic structure where *-er* combines first with the main adjective and then with the comparative clause. The comparative clause itself denotes a definite description of a degree, and is derived via ellipsis within the comparative clause and movement of a silent operator/*wh*-word (following Bresnan 1973; Chomsky 1977 a.o.). This movement triggers a further degree abstraction and a maximization operation (cf. von Stechow 1984). (For simplicity I ignore the possibility of phrasal comparatives.) For example:

(1.11) a. $[\![Op_i \text{ Harry is tall } t_i]\!]^{\mathcal{M},w,g} = \mathbf{max}(\lambda d[[\![tall]\!]^{\mathcal{M},w,g}([\![Harry]\!]^{\mathcal{M},w,g}) \geq d])$

b. $= \mathbf{max}(\lambda d[\mathbf{height}(\mathbf{Harry}) \geq d])$

5 Note however that the foundational work of Cresswell (1976) treated degrees as equivalence classes of individuals under a weak or quasi-order, rather than abstract points; cf. also Rullmann (1995). As the discussion in chapter 2 will make clear, these perspectives can be construed in such a way that they are equivalent. However, taking equivalence classes to be fundamental makes it possible to state theories which cannot be replicated in degree semantics: see Szabolcsi & Zwarts (1993) for one such example.

(1.12) Mary is taller than Harry is.

a. LF: *Mary is [[more tall] than [Op_i Harry (is tall t_i)]*

b. =1 iff: $[\lambda A_{\langle e,d\rangle}\lambda d_d \lambda x_e[A(x) > d]]([\![tall]\!]^{\mathcal{M},w,g})$ $([\![Op_i \textit{ Harry is tall } t_i]\!]^{\mathcal{M},w,g})$

c. =1 iff: $\textbf{height}(\textbf{Mary}) > \textbf{max}(\lambda d[\textbf{height}(\textbf{Harry}) \geq d])$

Although the literature has generally concentrated on gradable adjectives of type $\langle e,d\rangle$, it is not difficult to adapt the measure function analysis to gradability for arbitrary types; in general, for expressions of Boolean type $\langle \alpha,t\rangle$, the corresponding gradable type is $\langle \alpha,d\rangle$. The denotations for the comparative and other degree expressions can likewise be given type-polymorphic denotations which allow them to be applied to arbitrary gradable types as needed. So, for example, instead of treating *more/-er* as being of type $\langle\langle e,d\rangle,\langle d,\langle e,t\rangle\rangle\rangle$, we could write a type-polymorphic denotation which could equally well be applied to verbal comparatives:

(1.13) $[\![more/-er]\!]^{\mathcal{M},w,g} = \lambda K_{\langle \alpha,d\rangle}\lambda d_d \lambda k_\alpha[K(k) > d]$

Similar type-polymorphic denotations could be constructed for *as*, *almost*, the positive morpheme **pos** to be introduced shortly, and other degree operators which can modify expressions other than individual-modifying adjectives, e.g. the proposition-embedding adjectives with which we will frequently be concerned in the coming chapters.

1.3.2 The Positive Form, Vagueness, and Comparison Classes

Kennedy (2007) argues that the positive form (with no overt degree modification) is derived via a silent morpheme (or type-shifting operation) **pos**:

(1.14) $[\![\textbf{pos}]\!]^{\mathcal{M},w,g} = \lambda A_{\langle e,d\rangle}\lambda x_e[A(x) > \theta_A]$

θ_A is just a free variable here; how its value is determined in context is a complex and controversial question which touches on issues relating to vagueness, the semantics of comparison classes, and adjective type discussed in this and the next subsection.

The literature on vagueness is vast, and I will not have a great deal to say on the topic in this dissertation (though the interested reader may consult Lassiter 2011b for my favored account, which utilizes probabilistic threshold values; cf. also Schmidt, Goodman, Barner & Tenenbaum 2009; Frazee & Beaver 2010). For the purposes of this work, we can think of vagueness as a kind of pervasive context-sensitivity in determining the threshold value, as argued for example by e.g. Fara (2000); Barker (2002); Kennedy (2007). Roughly, certain adjectives in the positive form and with some degree modifiers determine their threshold value by reference to a “norm”, “expected value”, or “standard value” whose value is constrained (but perhaps not fully determined) by features of the semantics of the expression, its grammatical environment, and the discourse context.

For example, the sentences in (1.15) can both be true, even if the elephant is much bigger than the flea:

(1.15) a. This flea is big.

b. This elephant is not big.

This difference is plausibly explained by assuming that *big* is interpreted with respect to a standard value which is sensitive to features of the discourse, and perhaps some class of objects with which the objects in question are implicitly compared. On this account, then, (1.15a) means something like "This flea is big relative to the relevant norm *N*", where *N* is given by context.

According to many authors, the threshold value is constrained in part by reference to an implicit or explicit COMPARISON CLASS. On this account, the context- and norm-sensitivity of (1.15a) comes down to roughly "This flea is big relative to the expected value for comparison class *C*", where *C* is some set of objects of which the flea in question is a member. On the plausible assumption that the implicit comparison class relevant to evaluating (1.15a) is the set of fleas, while the relevant comparison class for (1.15b) is the set of elephants, we have the beginnings of an explanation of how these two sentences can be simultaneously true.

APs with explicit comparison classes are typically of the form *A for a NP*, as in (1.16).

(1.16) a. Harry is heavy for a jockey.
b. Harry is heavy for a sumo wrestler.

The use of an explicit comparison class brings with it the requirement that the individual to which the adjective is applied is a member of the comparison class: thus (1.16a) is infelicitous unless Harry is a jockey, and (1.16b) is infelicitous unless he is a sumo wrestler (Kennedy 2007). In Kennedy's analysis, this is evidence for that comparison classes exert their semantic effect by restricting the domain of the measure function. For concreteness' sake I will make this assumption when we discuss comparison classes in chapters 4 and 6, although the details of this analysis are not vital here.

1.3.3 Adjective Type

Recent work has emphasized the importance of ADJECTIVE TYPE and SCALE TYPE in the semantics of gradable adjectives (Rotstein & Winter 2004; Kennedy & McNally 2005; Kennedy 2007). The fact that adjectives and scales come in a variety of forms, and that these distinctions have grammatical repercussions, will be reflected in our treatment of modality as well.

Unger (1971) seems to have been the first to notice that gradable adjectives come in two types, which he called ABSOLUTE and RELATIVE. Relative adjectives like *tall* and *expensive* are the ones which have generally occupied philosophers concerned with vagueness; it is generally unclear just how tall someone has to be in order to count as tall, or what dollar amount makes an item expensive, even if we have a fully specified comparison class. Closely related to relative adjectives are what I will call HIGH DEGREE ADJECTIVES, exemplified by *huge, tiny*, and *ecstatic*. These adjectives also combine with comparison classes, and in the positive form mean roughly that their argument has a <u>much</u> greater degree of size/happiness than the norm or expected value (subject to the same caveats as relative adjectives about persistent vagueness, and questions about how exactly the comparison class achieves its semantic effect).

However, Unger points out that not all adjectives behave this way: for example, whether or not an object is *flat* is plausibly an all-or-nothing affair. If an object (say, a road) has any bumps in it, it is not "flat" but at best "almost flat" or "approximately flat". This differs considerably from heights and costs: a road must be maximally flat if it is to be flat at all, while someone can be tall without being maximally tall (whatever this would even mean). Adjectives like "flat" are those which Unger calls **absolute**.

Kennedy & McNally (2005); Kennedy (2007) point out that the absolute adjectives cleave further into two groups. MAXIMUM (or "maximum-standard") adjectives like *flat*, *full*, *straight*, and *safe* require that an object have a maximal degree of the property in question in order to count as instances of the concept. So, for example, if I tell you that my beer glass is full, it is strange to continue by asserting that it could be fuller (Kennedy 2007). In contrast, MINIMUM (or "minimum-standard") adjectives like *bent* and *dangerous* require only that an object have a non-minimal degree of the property in question; for example, an antenna is bent if it has any amount of bend in it. Unlike maximum adjectives, there is generally no oddity in saying that something is bent, but also that it could be more bent.

As Kennedy discusses in detail, absolute (minimum and maximum) adjectives share a number of properties. For instance, the positive form of both types of adjective typically has sharp boundaries and thus a lack (or near-lack) of vagueness. Absolute adjectives are also much less sensitive to comparison classes: *This is bent for an antenna* seems strange, at best a funny way to say *This antenna is bent*. Kennedy argues that we can account for all of these properties if absolute adjectives require that their threshold be an extreme scalar value. On this analysis, if A is a minimum-standard adjective, then θ_A must be the minimum point on $\mathcal{S}_A$, the scale associated with A — an antenna is bent just in case it has a non-zero degree of bend. Likewise, if A is a maximum-standard adjective, then θ_A must be the maximum point on $\mathcal{S}_A$ — a glass is full just in case it has a maximal degree of fullness. In other words, if A is maximum-standard, then *x is* **pos** *A* is true if and only if $A(x) = \mathbf{max}(\mathcal{S}_A)$. If A is minimum-standard, then *x is* **pos** *A* is true if and only if $A(x) > \mathbf{min}(\mathcal{S}_A)$.

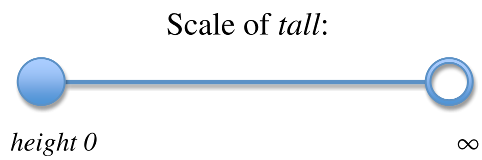
1.3.4 Scale Structure

Because absolute adjectives in the positive form constrain θ to be at the minimum or maximum point of the scale, it follows that an adjective can only be absolute if it is associated with a scale which has a minimum or maximum as appropriate. For this reason scale structure places constraints on adjective type: we cannot speak meaningfully of maximum and minimum values with all types of adjectives. Recently a theory of scale types focusing on BOUNDEDNESS has been developed by Rotstein & Winter (2004); Kennedy & McNally (2005); Kennedy (2007). The crucial observation is that scales may vary in the presence or absence of a lower bound, and independently in the presence or absence of an upper bound. Furthermore, this variation can be related to a number of linguistically interesting properties in addition to the relative/absolute distinction.

For instance, *tall* is presumably not associated with any maximum value: at least, as an ontological fact, there is no upper limit on possible heights. A number of theorists have suggested that this fact about *tall* and similar adjectives has linguistic consequences. For instance, von Stechow (1984); Rullmann (1995) argue that (1.17) is semantically ill-formed because there is no unique or maximal height h such that Sam is not h-tall, and so the comparative clause makes reference to an undefined maximum *Sam is not d-tall* (cf. (1.11-12)).

(1.17) # Mary is taller than Sam isn't.

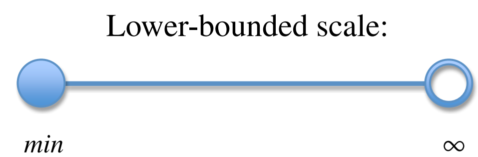
This account of (1.17) relies on the assumption that *tall* can be associated with a scale formed of all of the possible heights between 0 and infinity:



The ill-formedness of (1.17) is not a special fact about heights, though: this sentence would be infelicitous with various other adjectives replacing *tall*, such as *rich*.

(1.18) # Mary is richer than Sam isn't.

Rich, too, is intuitively associated with a scale with a lower bound ($0) but no upper bound — you can keep getting richer forever if you have enough time, energy, and luck. We can represent what *tall* and *rich* share by generalizing $\mathcal{S}_{tall}$ to the notion of a LOWER-BOUNDED scale:



Note that we use *Min* in the general case, since we cannot assume that every scale will be readily related to numerical values as *rich* and *tall* are, or that 0 is the minimum for all scales that are.

Supposing for present purposes that scales associated with natural language expressions are always connected, if we allow for all logically possible variations with respect to boundedness properties, the typology of possible scales with respect to boundedness is as the figure below. Kennedy & McNally (2005) discuss this typology and show that all four of these possibilities are instantiated among gradable adjectives in English.

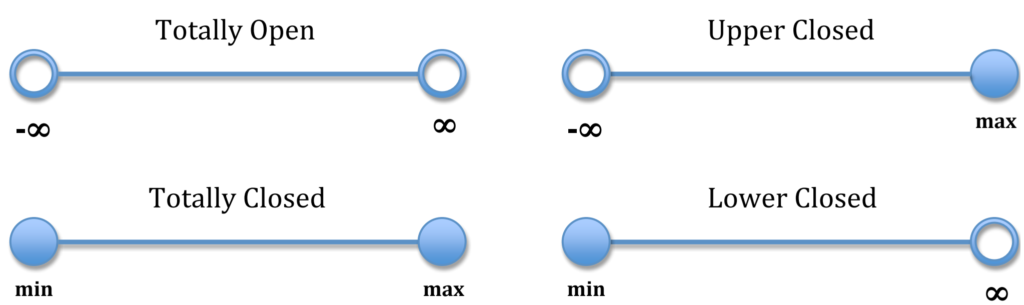


Fig. 1.1. Possible scale types with respect to boundedness properties.

There is a clear connection between the notion of adjective type discussed above and the scale types just given. In order to get the interpretation that we ascribed to them, maximum adjectives like *full* and *flat* must be associated with a scale which has a maximum element. This limits them to upper-closed and fully closed scales. Likewise, minimum adjectives like *bent* and *dangerous* cannot be associated with a scale which lacks a minimum element, which limits them to either lower-closed or fully closed scales.

Rotstein & Winter (2004); Kennedy & McNally (2005); Kennedy (2007) give a number of empirical tests for adjective type and boundedness properties. As an example, consider the degree modifier *slightly*. *x is slightly A* is true, roughly, just in case *x* has the property picked out by *A* to a small but non-zero degree. If we want to cash out this intuition more precisely, note that we have to assume that it makes sense to talk about a zero degree of the property denoted by *A* — that is, that *A* can sensibly be associated with a scale with a minimum element. If *A*'s scale does not have a minimum, we expect semantic anomaly. The presence or absence of a minimum element on the relevant scales, then, can be invoked to explain why the sentences in (1.19) are acceptable while those in (1.20) are not:

(1.19) a. This neighborhood is slightly dangerous.
b. This antenna is slightly bent.

(1.20) a. # This neighborhood is slightly safe.
b. # This antenna is slightly straight.

If the sentences in (1.20) can be interpreted at all, they must be taken to describe e.g. *how much* of the neighborhood is safe, rather than *the degree to which* the neighborhood is safe.

The explanation of (1.19)-(1.20) given by the authors cited is that *dangerous* is associated with a lower-closed scale. Intuitively this corresponds to the observation that a neighborhood can get more and more dangerous ad infinitum; however, there is a minimum amount of danger that it can have, namely complete safety. As a result modification by *slightly*, which is restricted to adjectives whose scale has a minimum element, is acceptable. However, *safe* has a scale which is the inverse of the scale of *dangerous*, and so is upper-closed. As a result, modification by *slightly* is not permitted because the scale has no minimum. Kennedy & McNally (2005); Kennedy (2007) give a number of other arguments which converge on these conclusions, some of which will be reviewed in chapter 3.

Further examples of tests for adjective type and scale type are *almost, completely*, and proportional modifiers. Rotstein & Winter (2004) show that, if an adjective can be modified by *almost*, its scale has a maximum:

(1.21) a. This neighborhood is almost safe/#dangerous.
b. This antenna is almost straight/#bent.

Likewise, Kennedy & McNally show that *completely*-modification is acceptable with a degree-modifying meaning only when an adjective has a scale with a maximum element:

(1.22) a. This neighborhood is completely safe/#dangerous.
b. This antenna is completely straight/#bent.

Note that *completely* is sometimes possible with scales with no maximum, but in these cases it indicates emphasis, correction, or high speaker confidence rather than maximization.

(1.23) **Mary**: The president is not tall.
Sue: Uh-uh! He is completely tall.

On the other hand, *proportional* modifiers like *half*, *90%*, and *mostly* measure the relative distance of an object from both the maximum and the minimum of a scale, and thus can only modify adjectives which are associated with a fully closed scale. As we have seen, neither *safe* nor *dangerous* fulfill this requirement, but e.g. *full/empty* and *open/closed* do; this explains the data in (1.24).

(1.24) a. # This neighborhood is half/90%/mostly dangerous/safe.
b. This glass is half/90%/mostly full/empty.
c. This window is half/90%/mostly open/closed.

Again, the examples marked as infelicitous can be given an interpretation on which the adverb quantifies the proportion of the spatial area of the neighborhood which is safe/dangerous, but this is not the degree-modifying reading that we are interested in.

Boundedness and adjective type will crop up repeatedly in the discussion of modality in later chapters, particularly when we discuss the epistemic adjectives *possible, probable, likely*, and *certain* in chapter 3.

1.4 Two Perspectives on Scales

As I have presented it, following in particular Kennedy (1997, 2007), scales are ordered sets of degrees. What are degrees? According to Kennedy and others (e.g., von Stechow 1984; Bierwisch 1989; Heim 2001), degrees are abstract representations of measurement. On this account, degrees of height or happiness exist, and these scales have the structure that they do, independent of whether any objects in the world actually possess those degrees of height or happiness.

This is probably the mainstream perspective in formal semantics, but it is not universal. Cresswell (1976); Klein (1991); Sassoon (2010); van Rooij (2010); Bale (2011) and others have argued that degrees should be thought of as **equivalence classes**: sets of objects all of which bear the "exactly as *P* as" relation to each other, for the relevant property *P* (cf. also Rullmann 1995). These authors take their inspiration in this regard from the Representational Theory of Measurement, an algebraic approach to measurement which has been highly influential in psychology, philosophy, and economics. For measurement theorists, degrees do exist, but they exist as an abstraction from the real-world objects which instantiate them and the qualitative relations that these objects bear to each other. Fundamental to this approach are BINARY ORDERS with varying amounts of structure, and CONCATENATION OPERATIONS which relate simple and compound objects.

Measure functions mapping objects to real numbers are employed in measurement theory as well, but care is taken to ensure that the numerical representations do not carry any information that is not already inherent in the qualitative structures underlying them. So, we can freely talk about objects such as "the degree to which Sam is tall", but the existence of this degree is dependent on the prior existence of a qualitative structure representing heights, containing among other things a set of individuals who bear the "exactly as tall as" relation to Sam.

The degree-based and measurement-theoretic perspectives are sometimes thought to be in competition, and they may well carry different philosophical commitments; for example, the measurement-theoretic perspective might be more attractive to someone who wishes to avoid ontological commitment to abstracta. As far as formal semantics is concerned, though, there is nothing to choose here: any degree semantics can be translated into an equivalent measurement-theoretic implementation, as we will discuss in some detail in the next chapter. (The reverse does not hold, though: measurement theory is more expressive than degree semantics.) As a result, the choice of whether to include degrees in our ontology is essentially a matter of convenience or philosophical proclivity; measurement theory provides a rigorous way to construct qualitative representations without degrees or numbers that are equivalent to quantitative representations with degrees or numbers.

Although the degree-based and measurement-theoretic analyses considered in this dissertation are logically equivalent, the algebraic perspective of measurement theory is very useful to adopt here, and will be the subject of some detailed formal discussion in chapter 2, for several reasons. First, the process of constructing scales using measurement theory — rather than simply treating them as unanalyzed primitives — will suggest new possible parameters of variation in scale type, several of which, I will argue, are in fact instantiated in natural language scales, and vital for the understanding of modality in natural language. Needless to say, since the two approaches are equivalent, nothing that I will propose makes the use of measurement theory obligatory. The situation is comparable to the relationship between formal logic and its algebraic treatment: even though an algebraic re-formulation of (say) propositional logic is provably equivalent to the more familiar style of presentation, certain aspects of the theory become clearer from an algebraic perspective, and certain methods of proof become available which were previously obscured. Similarly, measurement theory as I use it here does not add anything vital to standard degree semantics, but it allows us as theorists to adopt a different perspective on our familiar degree semantics which suggests new ways of viewing problems and new connections.

Second, measurement theory provides a well-understood method for constructing degree-based representations from qualitative orderings such as those underpinning Kratzer's theory of modality. Since Kratzer's theory is the standard one among linguists, and it relies heavily on a binary relation of comparative possibility, it provides a natural starting point for the project of devising a scalar semantics for gradable modals. When we begin to undertake this project in chapter 3, the tools of measurement theory will give us exactly what we need to be explicit about the predictions of this theory and to consider its strengths and weaknesses.

For these and other reasons discussed there, chapter 2 is dedicated to a formal presentation of the aspects of the Representational Theory of Measurement that are most relevant for this dissertation, in particular the method of constructing measure functions from qualitative orderings. I will also use chapter 2 to argue for an expanded range of scale types which can be given a natural formulation in measurement-theoretic terms, which will play an important role in the scalar semantics for modals developed in chapters 3-6.

Now I turn to an overview of modal semantics, focusing on the formal structure of the standard theory in linguistics (Kratzer 1981, 1991). This theory will provide our main starting-point and the benchmark with which to compare the alternative modal semantics proposed in this dissertation.

1.5 Modality and Modal Semantics

1.5.1 Overview

The term "modal" is used in at least two different ways. Sometimes it is used to pick out a syntactic category, the MODAL AUXILIARIES *may, might, can, could, should, would, must*, and perhaps *ought*.

I will use "modal" in a more expansive way to refer to expressions which have a particular semantic flavor. As Portner (2009: 1) puts it:

> [M]odality is the linguistic phenomenon whereby grammar allows one to say things about, or on the basis of, situations which need not be real.

This is more of a pointer than a definition – Portner precedes it with the proviso "I am not too comfortable trying to define modality" – but it provides a reasonable characterization of modality as a semantic phenomenon.

Construed this way, a wide variety of natural language expressions have (or have been claimed to have) modal semantics, going well beyond the small set of modal auxiliaries: conditionals, *because*-clauses, imperfective verbs, the future tense, expressions of mood, evidentials, many attitude verbs, and probably much more. I will not use the term "modal" this broadly here, though. I am primarily interested in modal expressions that take propositions as arguments (perhaps in addition to other arguments) and fall into one of four syntactic categories: auxiliaries, verbs, adjectives, and sentential adverbs.

(1.25) AUXILIARIES
- a. Harry <u>should</u> be in Sacramento by now.
- b. My brother <u>can</u> bench press 250 pounds.
- c. All cameras <u>must</u> be checked at the door.

(1.26) VERBS
- a. I <u>need</u> to go to Sacramento.
- b. My mother <u>wants</u> to be on television.
- c. You are <u>required</u> to wait behind the line.

(1.27) ADJECTIVES
- a. We are <u>unable</u> to fulfill your request.
- b. It is <u>likely</u> that we have missed our train.
- c. It is <u>impermissible</u> to fake illness to get out of work.

(1.28) ADVERBS
- a. <u>Evidently</u>, we have missed our train.
- b. We will <u>possibly</u> be in Houston next week.
- c. <u>Obligatorily</u>, children are picked up by 3PM.

Since these expressions come in several syntactic categories, we might expect that their semantics will vary somewhat as well. Nevertheless, the modals in (1.25)-(1.28) are usually analyzed as having a common semantics built on what I will call standard modal logic, the semantic framework associated with e.g. Hintikka (1962); Kripke (1963) and much following work in logic, philosophy, and recently computer science (cf. e.g. Goldblatt 1987, 2003; Fagin, Halpern, Moses & Vardi 2003; Halpern 2003; Shoham & Leyton-Brown 2009).

Modals are traditionally thought to come in several semantic types, and certain of the auxiliary modals are ambiguous between two or more of these types. For example, *must* can be interpreted **epistemically** ("It *must* be, given what is known", **deontically** ("You *must* do this, according to the laws"), **teleologically** ("You *must* do this in order to accomplish your goals"), and perhaps **bouletically** ("I *must* have this"). Another important modal type is **dynamic** or **circumstantial** modality, which refers to abilities and potentials, and is exemplified by *can* in (1.25b) and *unable* in (1.27a).

Modal adjectives, adverbs, and verbs are generally pickier about their modal flavor: for instance, *want* is restricted to bouletic modality, *permissible* to deontic modality, and *likely* to epistemic modality. Note in connection with the latter that, although epistemic modality is traditionally contrasted with **doxastic** modality ("given what is believed"), the term "epistemic modal" is widely used even when talking about beliefs which are not necessarily true. The term "doxastic modal" would probably be better for *likely* and related expressions, but, in keeping with common practice, I will not distinguish the two.

In this dissertation I will mostly be interested in epistemic, deontic, and bouletic modality. In the next few sections I will present standard modal logic and the influential modification of this approach due to Kratzer (1981, 1991). Kratzer's theory is built on a comparative relation and is, in certain ways, tantalizingly similar to a degree-based semantics.

A note on terminology: throughout the dissertation the term "quantificational semantics" is reserved for theories that make use of quantification over worlds, as in standard modal logic and Kratzer's theory. Actually, the scalar alternative that I will propose can also be implemented using quantification, this time over degrees (or equivalence classes of propositions, see ch.2). However, I will be careless about this subtlety in the interest of not having to repeat the phrase "theories where modals are quantifiers over possible worlds" ad nauseam.

1.5.2 Standard Modal Logic

The classic analysis of modality treats modal expressions essentially as restricted quantifiers over possible worlds. The restriction is provided by an accessibility relation R, which comes in various types associated with the modal flavors (epistemic, doxastic, deontic, bouletic, dynamic, etc.) just discussed. Certain expressions, e.g. *want*, are lexically associated with one or several accessibility relations – in the case of *want*, R must be bouletic – while others are freer, e.g. *can*, which can be associated with (at least) epistemic, doxastic, deontic, or dynamic R.

Many modal expressions receive a plausible interpretation in standard modal logic as (implicit) existential or universal quantifiers over accessible worlds. Fixing an accessibility relation R:

(1.29) a. $[\![\textit{necessarily } \phi]\!]^{\mathcal{M},w,g} = 1$ iff for all w' such that wRw': $[\![\phi]\!]^{\mathcal{M},w',g} = 1$.

b. $[\![\textit{must } \phi]\!]^{\mathcal{M},w,g} = 1$ iff for all w' such that wRw': $[\![\phi]\!]^{\mathcal{M},w',g} = 1$.

c. $[\![\textit{possibly } \phi]\!]^{\mathcal{M},w,g} = 1$ iff for some w' such that wRw': $[\![\phi]\!]^{\mathcal{M},w',g} = 1$.

d. $[\![\textit{might } \phi]\!]^{\mathcal{M},w,g} = 1$ iff for some w' such that wRw': $[\![\phi]\!]^{\mathcal{M},w',g} = 1$.

e. $[\![\phi \textit{ is impossible}]\!]^{\mathcal{M},w,g} = 1$ iff for no w' such that wRw': $[\![\phi]\!]^{\mathcal{M},w',g} = 1$.

Although this approach works well for expressions with modal force at the "extremes", such as *must, might, possible*, and *impossible*, it is more difficult to apply to intermediate grades of modality or to comparative modalities (Kratzer 1991). For example, consider the intermediate modality *probable*, and comparative modalities such as *It is (morally) better that ϕ than it is that ψ*. Since the set of accessible worlds is unordered, the best we seem to be able to do in standard modal logic is:

(1.30) $[\![\phi \textit{ is probable}]\!]^{\mathcal{M},w,g} = 1$ iff, for the relevant epistemic accessibility relation R, there are more worlds w' such that wRw' and $[\![\phi]\!]^{\mathcal{M},w',g} = 1$ than there are worlds w'' such that wRw'' and $[\![\neg\phi]\!]^{\mathcal{M},w'',g} = 1$.

(1.31) $[\![\phi \textit{ is better than } \psi]\!]^{\mathcal{M},w,g} = 1$ iff, for the relevant deontic accessibility relation R, there are more worlds w' such that wRw' and $[\![\phi]\!]^{\mathcal{M},w',g} = 1$ than there are worlds w'' such that wRw'' and $[\![\psi]\!]^{\mathcal{M},w'',g} = 1$.

Assuming we want to allow for the possibility that the set W of possible worlds is infinite, these truth-conditions are problematic. If there happen to be an infinite number of ϕ- worlds, both of these sentence-types will come out as trivially false no matter what. In addition, it seems unlikely that counting worlds would give us the right truth-conditions: quite clearly, whether ϕ is morally better than ψ has no direct connection to the number of possible worlds that instantiate these two propositions.

A further problem with the use of standard tools of modal logic here is that the truth-conditions of complex expressions are stipulated in the meta-language rather than being derived compositionally. This is particularly damaging in the case of *ϕ is better than ψ*, a comparative sentence: presumably, the truth-conditions of this sentence ought to be derived using the same formal apparatus that we used to treat non-modal comparatives such as *John is taller than Mary* above. A related problem involves intermediate grades of modality with degree modifiers such as *It is <u>somewhat</u> probable that ϕ* and *ϕ is <u>much</u> better than ψ*. Such sentences are transparently related to complex degree expressions such as *Mary is <u>somewhat</u> happy* and *Sue is <u>much</u> funnier than Bill*, which have a compositional interpretation in a degree- or delineation-based theory of gradability and comparison. Presumably a compositional interpretation is needed for their modal counterparts as well (cf. Yalcin 2007, 2010; Portner 2009).

1.5.3 Kratzer's Semantics

Kratzer (1981, 1991) presents a revised modal logic, closely related to Lewis's (1973) semantics for counterfactuals, which deals with several of the problems that we noted for standard modal logic. Kratzer retains the assumption that most modal expressions are as restricted universal or existential quantifiers over accessible worlds, but her modals have much more complicated restrictions than in standard modal logic. This allows Kratzer to give reasonable truth-conditions for intermediate grades of modality, as well as modal comparatives.

Kratzer's semantics for modality relies on an ordering of COMPARATIVE POSSIBILITY on worlds, denoted $\succcurlyeq_{\mathbf{g}(w)}$. This order is derived from the interaction of two contextual parameters: a modal base **f** and an ordering source **g**.[6] The MODAL BASE **f** is a function which, given a world,

6 Two notes on $\succcurlyeq$ and related notation.

First, I will use $\succcurlyeq_{\mathbf{g}(w)}$ to indicate "closer to the ideal provided by the ordering source **g** at world w", while Kratzer, returns a set of propositions that are relevant to the evaluation of the modal expression. In the case of epistemic modality, for example, the modal base is the set of propositions known to/believed by the speaker (or whoever else the contextually appropriate person(s) are).

The ORDERING SOURCE **g** is a function which, applied to a world w, returns a set of propositions which induces an ordering over the modal base. In the case of deontic modals, for example, the ordering source is some contextually relevant set of laws, orders, norms, etc., and worlds are ranked by how close they come to satisfying all of the propositions in $\mathbf{g}(w)$. The ordering is determined by the rule in (1.32) (where $u \in p$ abbreviates $[\![p]\!]^{\mathcal{M},u,g} = 1$):

(1.32) For all worlds $u, v \in W$:
$u \succcurlyeq_{\mathbf{g}(w)} v$ if and only if $\{p : p \in \mathbf{g}(w) \wedge v \in p\} \subseteq \{p : p \in \mathbf{g}(w) \wedge u \in p\}$

That is, u is at least as good/normal/etc. as v iff u satisfies every law (norm, etc.) that v does. u is strictly better than v, $u \succ v$, iff u satisfies every law that v does, and v does not satisfy every law that u does.

For a concrete example, suppose that we have three norms in play, and a domain of worlds which have the following properties:

(1.33) NORMS:
N1. Children obey their parents.
N2. No tresspassing.
N3. No murder.

(1.34) w_1: N1-N3 are obeyed.
w_2: Only N1 violated.
w_3: Only N2 violated.
w_4: Only N3 violated.
w_5: Only N1 and N2 violated.
w_6: Only N1 and N3 violated.
w_7: Only N2 and N3 violated.
w_8: N1, N2, and N3 all violated.

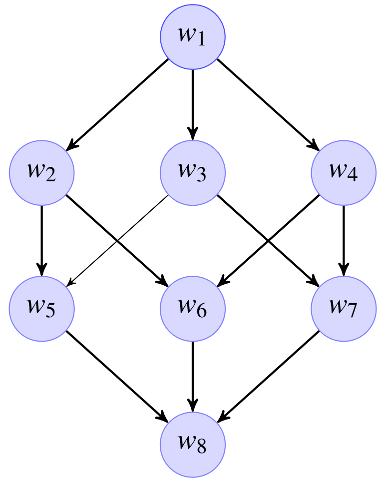
This example gives us a relation of comparative possibility $\mathbf{g}(w)$ with the following structure. (Reflexive and transitive arrows are left implicit to avoid clutter.)

---

following Lewis, uses $\preccurlyeq_{\mathbf{g}(w)}$. This choice makes sense within Lewis' theory of counterfactuals, but is confusing in the current context, since it seems to suggest 'less than or equal to', while the orderings we are interested in with respect to gradable modals correspond intuitively to an ordering in terms of 'greater than or equal to'. When we compare these notions explicitly to the orderings induced by gradable adjectives, the current notation will be preferable.

Second: here and throughout I will use $\succcurlyeq$ and the related symbols $\succ$, $\approx$, $\preccurlyeq$, $\prec$ for qualitative orderings of worlds, objects, etc., and reserve $\geq$, $>$, $=$, $\leq$, $<$ for orderings of numbers and degrees.

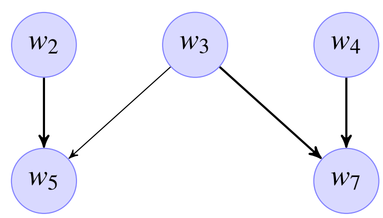
Relation $\succcurlyeq_{\mathbf{g}(w)}$ with domain $\{w_1, w_2, ..., w_8\}$:



Note that, in cases like this in which the propositions in $\mathbf{g}(w)$ are consistent and independent and there are enough worlds in the modal base to instantiate all possible combinartions, the relation $\succcurlyeq_{\mathbf{g}(w)}$ has the same structure as the subset relation $\subseteq$ by which it is defined in (1.32) (i.e., a Boolean algebra).

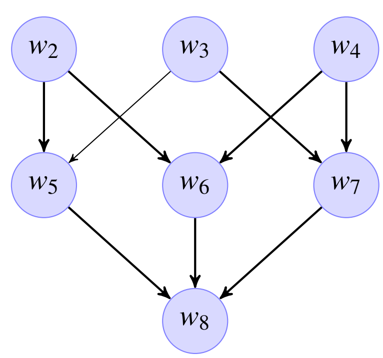
Kratzer takes care to construct the order $\succcurlyeq_{\mathbf{g}(w)}$ so that it is also well-defined in cases in which there is conflict between norms/expectations. For instance, imagine that the norms are the same as in the above example, but someone's parent has instructed them to commit murder. In this case there is no possibility of violating no norms, and so there is no top-ranked world. If there is only one relevant way to obey parents and one way to commit murder, then we get the impoverished set of possibilities:

(1.35) w_2: Only N1 violated.
w_3: Only N2 violated.
w_4: Only N3 violated.
w_5: Only N1 and N2 violated.
w_7: Only N2 and N3 violated.



If there are other ways of disobeying parents and committing murder which render these independent, however, we can treat the conflict by simply removing the top-ranked world.

(1.36) w_2: N1 violated.
w_3: N2 violated.
w_4: N3 violated.
w_5: N1 and N2 violated.
w_6: N1 and N3 violated.
w_7: N2 and N3 violated.
w_8: N1, N2, and N3 violated.



In the first example, the best world was the one where no norms were violated, w_1; as it happens, the set of ideal worlds relative to $\mathbf{g}(w)$ will always be $\bigcap \mathbf{g}(w)$ when this set is not empty. In the second and third examples, where $\bigcap \mathbf{g}(w)$ is empty, the best worlds are the worlds which violate the fewest norms: w_2, w_3, and w_4.

This way of ordering worlds in terms of their closeness to an ideal makes some intuitive sense for deontic modals; certainly, at a minimum we want a world w' which obeys all the norms or laws that w'' does and then some to come out as better than w''. However, in the epistemic domain, it is less obvious what propositions to include in the ordering source. Kratzer (1981) suggests that, for epistemic modals, the ordering source is composed of propositions representing the "normal state of affairs": "Worlds in which the normal course of events is realized are a complete bore, there are no adventures or surprises".

In order to give truth-conditions to certain types of modal sentences, including modal comparatives, we need to define a relation on propositions in terms of the relation $\succcurlyeq_{\mathbf{g}(w)}$ on worlds. Kratzer does this in the following way:

(1.37) COMPARATIVE POSSIBILITY:
ϕ is at least as good a possibility as ψ (in w, relative to $\mathbf{f}$ and $\mathbf{g}$) iff:
For all $u \in \bigcap \mathbf{f}(w)$: if $u \in \psi$, then there is a world $v \in \bigcap \mathbf{f}(w)$ such that $v \succcurlyeq_{\mathbf{g}(w)} u$ and $v \in \phi$.

In other words, ϕ is at least as good/likely/desirable/etc. as ψ if and only if every ψ-world in the modal base is weakly dominated by some ϕ-world in the modal base. An equivalent condition is the requirement that there cannot be any ψ-worlds in the modal base which either (a) outrank or (b) are not comparable with all ϕ-worlds. Following notation in Halpern (1997), throughout the dissertation I will use the abbreviation $\succcurlyeq^{s}_{\mathbf{g}(w)}$ for the comparative relation on propositions "at least as good a possibility as" which is derived from the relation $\succcurlyeq_{\mathbf{g}(w)}$ on worlds as in (1.37):

(1.38) $\succcurlyeq^{s}_{\mathbf{g}(w)} =_{df} \{(\phi, \psi) \mid \forall u \in \psi \; \exists v \in \phi : v \succcurlyeq_{\mathbf{g}(w)} u\}$, where $u, v \in \bigcap \mathbf{f}(w)$.

Kratzer (1991) retains the core idea from standard modal logic that *must*, *necessarily*, etc. are universal quantifiers over worlds, and that *might*, *possibly*, etc. are existential quantifiers. However, rather than treating them as quantifiers over a set of worlds which is pragmatically determined once and for all, Kratzer thinks of them as quantifiers whose restriction is determined in a somewhat more complicated fashion by the modal base and ordering source. *Must* (and other strong modals, presumably) are defined as in (1.39):

(1.39) $[\![must\ \phi]\!]^{\mathcal{M},w,g} = \forall u \exists v [v \succcurlyeq_{\mathbf{g}(w)} u \wedge \forall z : z \succcurlyeq_{\mathbf{g}(w)} v \rightarrow z \in \phi]$ $\quad (u, v, z \in \bigcap \mathbf{f}(w))$

In any case, the effect of (1.39), as Kratzer explains, is that "a proposition is a necessity if and only if it is true in all accessible worlds which come closest to the ideal established by the ordering source".

The definition is intended to be maximally general, but it is admittedly a bit obscure as stated. There are two motivations for giving the definition in this way. First, if the propositions in the modal base are not consistent (so that $\bigcap \mathbf{g}(w) = \varnothing$), there will not be any worlds which dominate **all** other worlds. This is a result of the fact that $\succcurlyeq_{\mathbf{g}(w)}$ is defined in terms of a subset relation, so that the ordering is connected only in special cases. This is an important feature of Kratzer's semantics. Less crucially, Kratzer does not wish to assume that any particular branch of the ordering has a set of maximal elements. As far as I can make out, the latter assumption is only relevant if there are infinitely many propositions in $\mathbf{g}(w)$, a possibility that we can safely ignore here.[7]

Kratzer adds that "[t]he definition would be less complicated if we could quite generally assume the existence of such 'closest' worlds". We can get a more intuitive sense of what (1.39) amounts to if we make this assumption and treat *must* as a universal quantifier over the set of $\succcurlyeq_{\mathbf{g}(w)}$-undominated worlds, i.e. those v for which there is no v' s.t. $v' \succ_{\mathbf{g}(w)} v$:

(1.40) $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) =_{df} \{v \mid v \in \bigcap \mathbf{f}(w) \wedge \neg\exists v' \in \bigcap \mathbf{f}(w) : v' \succ_{\mathbf{g}(w)} v\}$

If every $\mathbf{g}(w)$-branch has one or more maximal worlds, then (1.39) is equivalent to (1.41).[8]

(1.41) $[\![must\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $\forall u : u \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) \rightarrow u \in \phi$.

That is, if every $\mathbf{g}(w)$-branch has maximal worlds, then *must* ϕ is true iff ϕ is true in all worlds which are maximal in their respective $\succcurlyeq_{\mathbf{g}(w)}$-branches. So, for example, in (1.35) and (1.36) $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w))$ is $\{w_2, w_3, w_4\}$. Viewed this way, Kratzer's definition of *must* and other strong modals is quite close to the definitions from standard modal logic that we saw above.

Kratzer also follows standard modal logic in treating possibility as the dual of necessity: that is, a proposition is a possibility (etc.) if and only if its negation is not a necessity.

(1.42) $[\![might\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $[\![must\ \neg\phi]\!]^{\mathcal{M},w,g} = 0$
$= 1$ iff $\exists u \forall v : [\neg(v \succcurlyeq_{\mathbf{g}(w)} u)] \vee [\exists z : z \succcurlyeq_{\mathbf{g}(w)} v \ \wedge \ z \in \phi]$ (where $u, v, z \in \bigcap \mathbf{f}(w)$)

If we assume again that every $\mathbf{g}(w)$-branch has one or more maximal worlds, then (1.42) becomes equivalent to (1.43):

(1.43) $[\![might\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $\exists u : u \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) \wedge u \in \phi$.

That is, *might* ϕ is true iff there is at least one ϕ-world among the worlds which come closest to the ideal established by the ordering source. This definition, too, is closer to standard modal logic than it might seem at first glance.

Kratzer's theory has various advantages over standard modal logic, however. The use of an ordering over accessible worlds rather than an unstructured set of worlds makes it possible to give reasonable truth-conditions for expressions that standard modal logic either gets wrong or cannot account for at all. For example, Kratzer (1991) shows that her theory accounts for the failure of certain inference patterns involving modals and conditionals which are predicted to be valid in standard modal logic (see chapters 5-6 for more discussion of these issues). Most importantly from the perspective of this dissertation, Kratzer is able to give truth-conditions to modal judgments intermediate between the extremes picked out by *impossible*, *possible*, *necessary*, etc. For example, Kratzer (1991) gives the following truth-conditions for some intermediate grades of modality:

(1.44) $[\![probably\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff ϕ is a better possibility than $\neg\phi$ (i.e. $\phi \succcurlyeq^{s}_{\mathbf{g}(w)} \neg\phi$)

---

7 The original motivation for worrying about the cardinality of the ordering source, as far as I can make out, goes back to Lewis's (1973) discussion of the limit assumption in counterfactuals. Though this question has important ramifications in the context of counterfactuals, I don't know of any corresponding reason to consider it to be a crucial issue in the semantics of modals, and I will simplify here by assuming that ordering sources are always finite.

8 Every branch has one or more maximal worlds iff $\forall z \in \bigcap \mathbf{f}(w)\ \exists z' \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) : z' \succcurlyeq_{\mathbf{g}(w)} z$.

(1.45) *There is a slight possibility that* ϕ is true in w iff

a. $\phi \cap \bigcap \mathbf{f}(w) \neq \varnothing$ (i.e., ϕ is compatible with the modal base); and

b. $\neg\phi$ is probable in w (by the definition given above).

(1.46) *There is a good possibility that* ϕ is true in w iff *It is probable that* $\neg\phi$ is false (by the definition in (1.44)).

Again, this is reasonable: a good possibility is not necessarily probable, but its negation should not be probable either.

(1.47) *It is more likely that* ϕ *than it is that* ψ is true iff ϕ is more possible than ψ by (1.37). (i.e. $\phi \succcurlyeq^{s}_{\mathbf{g}(w)} \psi$)

The justification of (1.47) is clear.

As Kratzer shows, these definitions predict the validity of a number of reasonable patterns of inference. For example:

(1.48) a. If *It is probable that* ϕ is true, then *It is probable that* $\neg\phi$ is false.

b. *There is a good possibility that* ϕ and *There is a good possibility that* $\neg\phi$ may both be true.

c. *There is a slight possibility that* ϕ is true iff *There is a good possibility that* ϕ is false, but ϕ is not impossible (i.e., if $\phi \cap \bigcap \mathbf{f}(w) \neq \varnothing$).

d. If *There is a slight possibility that* ϕ is true, then *It is more likely that* $\neg\phi$ *than it is that* ϕ is true.

Overall, this theory holds out the promise of deriving reasonable truth-conditions for gradable modal expressions, and has been shown to have numerous other virtues as well.

One problem which Kratzer's theory shares with standard modal logic, however, is the lack of a compositional treatment of complex modal expressions. For example, the definition in (1.47) treats *It is more likely that ... than it is that ...* as if it were a single discontinuous lexical item. In light of the similarity between this sentence-type and other comparatives, this is of course rather dubious: presumably this sentence should be decomposed into a statement about degrees, as other natural language comparative sentences are. However, it may well be that core features of Kratzer's approach can be retained while remedying this defect as Portner (2009) suggests; to my knowledge, however, no one has attempted to work out the details of such an analysis. In later chapters we will attempt to do just this, noting some rather severe problems which arise for Kratzer's theory along the way.

1.6 Conditionals

Although my focus in this dissertation is not on conditionals, it is impossible to give a comprehensive treatment of modality without adopting some theory of conditional semantics, and issues involving conditionals will crop up at various crucial points (especially in chapters 5-6). Throughout I will assume an analysis of conditionals as restrictors based on Kratzer (1986) — a treatment which is more or less standard in linguistics, and versions of which have seen recent popularity in philosophy as well (e.g., Yalcin 2007; Kolodny & MacFarlane 2010; Egré & Cozic 2011; Rothschild 2011). According to this analysis, the conditional is not a sentential connective but rather a device of domain restriction. That is, in a sentence *If* ϕ *then* ψ, the antecedent functions to restrict the domain of a modal expression contained in ψ to the worlds in which ϕ holds. (In order to make the theory work we must assume that when there is not an overt modal in the conditional consequent there is a covert one; the arguments for and against such an assumption will not detain us here.) Assuming Kratzer's theory of modality as described in §1.5.3, we can state this analysis as follows, where the interpretation is relativized to **f** and **g** represent the modal base and ordering source (Kratzer 1991: 648-9).

(1.49) $[\![\textit{If } \phi \textit{ then } \psi]\!]^{\mathcal{M},w,g}_{\mathbf{f},\mathbf{g}} = [\![\psi]\!]^{\mathcal{M},w,g}_{\mathbf{f}',\mathbf{g}}$ where, for all w, $\mathbf{f}'(w) =_{df} \mathbf{f}(w) \cup \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$.

Since $\mathbf{f}(w)$ represents the set of propositions known at world w, this analysis can be stated in plain English: to evaluate *If* ϕ *then* ψ, just pretend that you know ϕ and then evaluate ψ on the basis of that assumption. Since modal sentences get their truth-conditions on the basis of an ordering $\succcurlyeq_{\mathbf{g}(w)}$ on the worlds that satisfy **all** of the propositions in $\mathbf{f}(w)$, adding ϕ to $\mathbf{f}(w)$ will have the effect of eliminating from the order $\succcurlyeq_{\mathbf{g}(w)}$ all worlds in which ϕ does not hold.

There is a different way to state the analysis which has the same semantic effect, but will be more useful to us here since it does not rely on special features of Kratzer's theory of modality. Instead of interpreting a conditional by adding its antecedent temporarily to the modal base, we can treat the antecedent of the conditional as a restrictor of the binary order over worlds $\succcurlyeq_{\mathbf{g}(w)}$ to worlds which satisfy the antecedent. First define the restriction operator $\upharpoonright$ as:

(1.50) The **restriction** $\succcurlyeq\upharpoonright B$ of an order $\succcurlyeq$ to a set B is defined as $\{(x,y) \mid x \succcurlyeq y \wedge x \in B \wedge y \in B\}$.

Restricting a binary order to a set B means removing from the order any pair for which either member is not in B. We can now achieve the semantic effect of (1.49) in two steps. First, we relativize the interpretation of the sentence to a single parameter **h** which, in the ordinary case, gives us an order equivalent to Kratzer's $\succcurlyeq_{\mathbf{g}(w)}$. Second, we define the conditional so that semantic effect of the antecedent is to restrict the order so derived for the purpose of evaluating the consequent.

(1.51) $[\![\textit{If } \phi \textit{ then } \psi]\!]^{\mathcal{M},w,g}_{\mathbf{h}} = [\![\psi]\!]^{\mathcal{M},w,g}_{\mathbf{h}'}$, where

a. For all w, $\mathbf{h}(w) =_{df} \succcurlyeq_{\mathbf{g}(w)}$ as defined in (1.32) above.

b. For all w, $\mathbf{h}'(w) =_{df} \mathbf{h}(w) \upharpoonright \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$.

In (1.51) we allow the antecedent to restrict the order $\succcurlyeq_{\mathbf{g}(w)}$ directly, rather than indirectly as in (1.49). The reader may verify that, in the context of Kratzer's theory of modality, these two approaches are equivalent.[9] However, (1.51) is more general because it can be applied to any theory of modality which makes use of a binary order over worlds, and not just to one which calculates

9 Or can be made to be: in a handful of cases we have to change the definitions slightly in order to accommodate items for which Kratzer's official proposals do not make reference to $\mathbf{g}(w)$. For example, the first clause of the definition of *slight possibility* in (1.45) is $\phi \cap \bigcap \mathbf{f}(w) \neq \varnothing$, the requirement that there are some ϕ-worlds in the modal base. We can make this clause dependent on **g** by changing it to $\exists w \exists w'[w' \in \phi \wedge w \succcurlyeq_{\mathbf{g}(w)} w']$, the requirement that some ϕ-worlds appear somewhere in the ordering induced over the modal base; then restricting the ordering source as in (1.51) will have the same effect on the interpretation of *slight possibility* as restricting the modal base in (1.50).

a binary order from two contextual parameters in the way that Kratzer's does. Since the theory that I propose in later chapters has this character, I will assume throughout the dissertation this interpretation of conditionals.

1.7 Overview and Preview of Chapter 2

This chapter gave an overview of scalar semantics, quickly summarized a compositional semantics of gradability and comparison which treats scalar expressions as measure functions, described Kratzer's modal semantics and its advantages over standard modal logic as a theory of modality in natural language, and described briefly the semantics for indicative conditionals that is assumed here. Some of the high points of this discussion are that

- Scalar expressions determine a classical (two-valued) extension by a three-step process of mapping their argument to a value on a scale, establishing a threshold value, and comparing the value of their argument to the threshold value.
- In degree-based approaches, scales are partially or totally ordered sets of degrees which vary at least according to the presence or absence of minimum and maximum elements.
- Gradable adjectives come in at least four types:
  - Relative adjectives like *tall* and *rich*;
  - High degree adjectives like *huge* and *ecstatic*;
  - Minimum adjectives like *bent* and *dangerous*;
  - Maximum adjectives like *full, safe* and *straight*.
- Kratzer's (1991) theory of modality improves in a number of ways on standard modal logic, but continues to treat modal expressions as quantifiers or comparatives whose interpretation is non-compositional and relativized to two different types of relatively unconstrained contextual parameters;
- Despite the grammatical similarities between complex expressions of modality and gradable expressions more generally, only very recently has there been any attempt to give them a unified semantics.

The similarities between gradable and modal expressions suggest that we need to develop a theory which explains their commonalities. Furthermore, it is clear that this theory should be compositional and, where appropriate, should use the resources of existing theories of gradability. It remains to be seen just how closely related gradable expressions and modals are, however.

In the remainder of this dissertation I will argue that we should give up the venerable assumption that modals are a complicated sort of quantifiers. Instead, I argue, the commonalities between gradable and modal expressions are due to the fact that both have a semantics built on scales. As a result, a good modal semantics will proceed in the same way that a good semantics for gradable adjectives does: by asking what the structure of the relevant scales are, and where the threshold value is constrained to fall for various simple and complex scalar expressions. I will also pose a number of problems for quantificational semantics for modality and show that they have straightforward explanations in terms of the scalar semantics for epistemic, deontic, and bouletic modality proposed here.

The next chapter lays the groundwork for this argument. I present an algebraic approach to the construction of scales building on the Representational Theory of Measurement (Krantz et al. 1971). The degree-based theories outlined in this chapter can be recast straightforwardly in this framework, and the algebraic approach clarifies a number of issues with respect to the analysis of modality and suggests new avenues of inquiry for both degree semantics and modal semantics.

Chapter 2

Measurement Theory, Gradability, and The Typology of Scales

This chapter discusses the issue of scale type in more detail, using the formal tools of the Representational Theory of Measurement (RTM, also known simply as Measurement Theory). RTM allows us to use qualitative (algebraic) and quantitative (measure function) characterizations of scales interchangeably, using what is effectively a type of supervaluation semantics for measure functions.

A number of linguists (e.g., Cresswell 1976; Klein 1980, 1982, 1991; Krifka 1989, 1990, 1998; Nerbonne 1995; Sassoon 2007, 2010; Bale 2006, 2008, 2011; van Rooij 2009, 2010) have suggested accounting for gradability and comparison in various domains using the resources of measurement theory rather than an apparatus taking degrees and scales composed of degrees as primitive. This approach is sometimes presented as a competitor of degree-based theories (e.g., by Klein 1980, 1982, but see Klein 1991 for a different perspective). However, the measurement-theoretic approach is not really an *alternative* to anyone's semantics of gradation; rather, it is a framework of considerable expressive power into which apparently diverse semantic proposals can be translated and compared. For example, both degree-based and delineation-based semantics for gradable adjectives can be expressed using measurement theory. However, degree-based approaches, and particularly measure function-based approaches, are the most straightforward to state using RTM tools, and I will focus on them here for this reason.

Using RTM as a formal underpinning for degree semantics carries a number of advantages for the purposes of this dissertation. First, RTM shares a crucial feature with the standard theory of modality due to Kratzer: both are built around binary relations, and the formal properties of binary relations and their relation to degree semantics have been studied in some depth in the case of RTM. This makes RTM a natural choice of representation framework for studying the predictions of Kratzer's theory, and for the overall project of unifying modal semantics and degree semantics.

Second, the process of constructing a degree semantics using RTM will force us to be very explicit about the mathematical properties attributed to the scales that we associate with gradable expressions in natural languages. The richer typology of scales that are naturally expressed using RTM is, I will argue, needed to express the full range of scales employed in natural languages. Several of the scales that are developed will also be crucial for the semantics of epistemic and deontic modals and desire verbs that I will develop in subsequent chapters, and their detailed formal properties to be developed in this chapter will be used there to explain a number of puzzling phenomena involving degree modification, disjunction, conditionals, and other interactions.

Third, RTM has interesting points of contact with algebraic semantics, both in the Boolean semantics tradition (Keenan & Faltz 1985; Winter 2001) and structured-domain semantics for the semantics for plurals and events (as in Link 1983, 1998; Bach 1986; Krifka 1989, a.o.). In particular, I will argue in this chapter that the concatenation operation in measurement theory can be identified as a restricted version of the join operation in algebraic semantics. This identification has important ramifications for the semantics of modality, since concatenation plays a role in the construction of many scales, and join realizes natural language disjunction in algebraic semantics.

2.1 Introduction to Measurement Theory

Measurement Theory was developed beginning in the late 19th century as a mathematical foundation for measurement in the physical and psychological sciences (e.g., Helmholtz 1887; Hölder 1901). Its modern incarnation, the Representational Theory of Measurement, stems from the foundational work on scale types by Stevens (1946), written with a strongly psychological focus: "Is it possible to measure human sensation?" Prior to this work it was often assumed that measurement had an intrinsic connection to numbers per se, and that it was senseless to speak of measurement unless, for example, some content could be given to the notion of addition applied to measurements in the domain in question. Since it was not obvious how to map measurement of sensations to an addition operation in many domains, quite a few psychologists had concluded that the concept of measurement did not make sense when applied to human perception.

Stevens' approach was essentially to work from the other direction: instead of starting with criteria for which operations are necessary for something to count as "measurement" and trying to impose these on the data of psychology, Stevens suggested looking at the kinds of data that were available and seeing what mathematical structures they could and could not support. Formally, this meant treating scales as algebraic structures consisting of one or more sets and (optionally) a n-ary relation and $m \geq 0$ further relations and operations. These algebraic structures and the qualitative relations that they encode are taken as basic, and numerical measurement involves asking which kinds of numerical representations faithfully preserve the structure of various types of scales; that is, what is the class of homomorphisms from a scale into the natural numbers, integers, rationals, or real numbers as the case may be. RTM was extended and further formalized by Scott & Suppes (1958); Suppes (1959); Suppes & Zinnes (1963); Narens (1985) and others, most authoritatively in Krantz et al. (1971). A good introduction to RTM is Roberts (1979).

The usefulness of this approach to measurement is illustrated by the contrast between measurements of length and width on the one hand and clock time on the other. Measurements of time can extend as far back as you like, and the choice of zero point is arbitrary; in contrast, all reasonable measurements systems for length have a fixed minimum (zero length) and will share a fixed zero point (the same). Differences of this type crop up routinely in physical and psychological measurement, and are reflected in intuitive judgments of the felicity of certain kinds of statements. It is unremarkable to say of one building that it is twice as wide or tall as another one, but difficult to make sense of the claim that one event is twice as late as another (to interpret this, we have to supply a third event with respect to which we are implicitly measuring lateness). In RTM, this difference is traced back to a qualitative difference between the scales which determines which kinds of quantitative statements are interpretable, as I will explain in §2.1.2.1.

2.1.1 Foundations

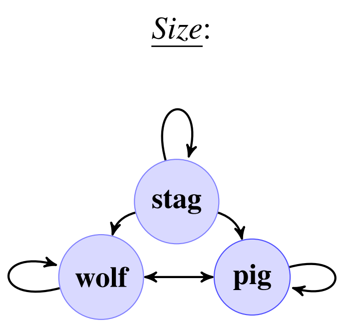
Imagine that you had to construct a measurement system from scratch, without the help of a system of numbers. How would you go about doing this?

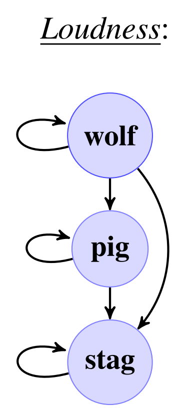
The first thing to decide, obviously, is what sort of property P you are measuring. The second step is to consider the relative ordering of pairs of objects of interest with respect to the ordering that you are trying to create. Given a domain of objects X which we are interested in, you can ask of each $x, y \in X$: Does x outrank y with respect to property P? Does y outrank x? Are they equal in P-ness? Are they incomparable?

This, at least, is the start of a rational reconstruction of the systems of measurement that human languages and human societies utilize. Starting with a **domain** X, we construct a **binary relation** $\succcurlyeq_P$ using a comparison procedure like the one just outlined. $\succcurlyeq_P$ is just the set of all pairs (x, y) such that x is equal to or greater than y with respect to property P. So, for example, with the tiny domain $X = \{$**stag, wolf, pig**$\}$ we might have:

(2.1) $\succcurlyeq_{size} = \{(\textbf{stag, stag}), (\textbf{stag, wolf}), (\textbf{stag, pig}), (\textbf{pig, pig}), (\textbf{pig, wolf}), (\textbf{wolf, wolf}), (\textbf{wolf, pig})\}$
$\succcurlyeq_{loudness} = \{(\textbf{wolf, wolf}), (\textbf{wolf, pig}), (\textbf{wolf, stag}), (\textbf{stag, stag}), (\textbf{pig, pig}), (\textbf{pig, stag})\}$

This can be represented more clearly as a directed graph with arrows playing the role of $\succcurlyeq$:

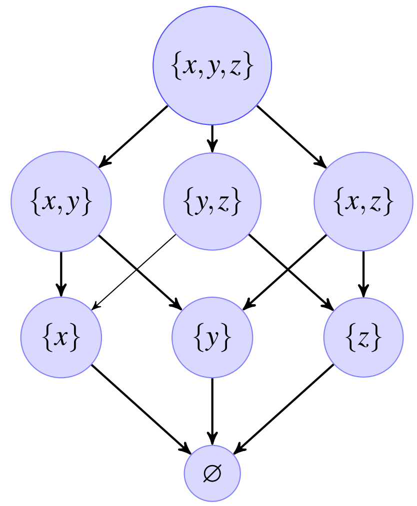




We can then inspect these relations to see if they have any other properties of interest. For example, the two binary relations in (2.1) are REFLEXIVE — everything is at least as great as itself with respect to size and loudness — and TRANSITIVE — if $(x, y) \in \succcurlyeq_P$ and $(y, z) \in \succcurlyeq_P$, then $(x, z) \in \succcurlyeq_P$. (I will mostly write $x \succcurlyeq_P y$ as an abbreviation for $(x, y) \in \succcurlyeq_P$.)

The relations in (2.1) are also both COMPLETE, meaning that any two objects in the domain X are comparable: $\forall x \forall y [x \succcurlyeq_P y \vee y \succcurlyeq x]$. Many binary relations are not complete, however, for instance the familiar subset relation: neither $\{x, y\} \subseteq \{y, z\}$ nor $\{y, z\} \subseteq \{x, y\}$. (To avoid clutter I suppress the reflexive and transitive arrows in this and later figures.)

Subset relation $\subseteq$ *with domain* $\{x,y,z\}$:



Another property of interest is ANTISYMMETRY, which is satisfied by a binary relation $\succcurlyeq_P$ if and only if $x \succcurlyeq_P y \wedge y \succcurlyeq_P x$ implies $x = y$. In the small domain of (2.1), $\succcurlyeq_{loudness}$ is antisymmetric (by accident of the small size of X, though), but $\succcurlyeq_{size}$ is not, since **pig** $\succcurlyeq_{size}$ **wolf** and **wolf** $\succcurlyeq_{size}$ **pig** — that is, these objects are judged to be equivalent in size — but **pig** $\neq$ **wolf**. I will often use $x \approx_P y$ as an abbreviation for $(x \succcurlyeq_P y) \wedge (y \succcurlyeq_P x)$, and I will use $x \succ_P y$ to abbreviate $(x \succcurlyeq_P y) \wedge \neg(y \succcurlyeq_P x)$.

Another important concept that can be derived from a binary order is the EQUIVALENCE CLASS:

(2.2) The equivalence class of x relative to a relation $\succcurlyeq_P$, written $[x]_P$, is the set $\{y \mid y \approx_P x\}$.

Some useful terms for frequently occurring types of relations are:

(2.3) a. **Pre-order** (a.k.a. **Quasi-Order**): transitive and reflexive.
b. **Partial Order**: transitive, reflexive, and antisymmetric.
c. **Weak Order**: transitive, reflexive, and complete.
d. **Simple Order**: transitive, reflexive, complete, and antisymmetric.

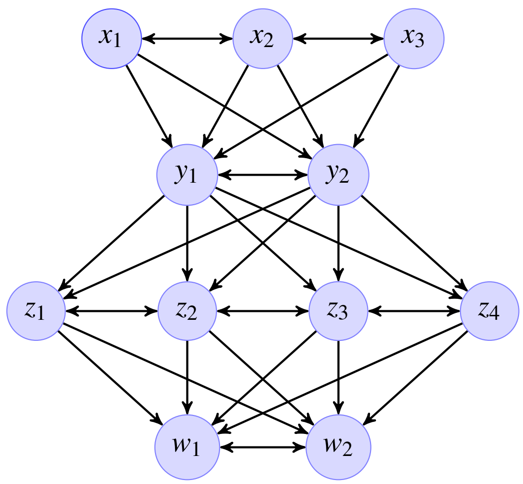
Note that a simple order $\succcurlyeq_P$ is a special case of a weak order where, since $\succcurlyeq_P$ is antisymmetric, $[x]_P$ is the unit set $\{x\}$ for any x. Similarly, a partial order is a special case of a pre-order where $[x]_P = \{x\}$ for every x.

In fact, every weak order is systematically related to a simple order in the following way. Let $(X/\approx_P)$ be the set of equivalence classes under the relation $\succcurlyeq_P$ with domain X, i.e. $\{Y \mid \exists x \in X : Y = [x]_P\}$. Then the relation $\succcurlyeq^*_P$ is the REDUCTION of $\succcurlyeq_P$ if and only if

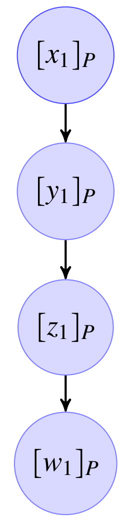
(2.4) a. $\succcurlyeq^*_P$ is a binary relation on $(X/\approx_P)$; and
b. $x \succcurlyeq_P y$ if and only if $[x]_P \succcurlyeq^*_P [y]_P$.

That is, if $\succcurlyeq_P$ is a binary relation on X, then its reduction $\succcurlyeq^*_P$ is the corresponding binary relation on the set of equivalence classes of X with respect to $\succcurlyeq_P$. If $\succcurlyeq_P$ is a weak order, then $\succcurlyeq^*_P$ will be a simple order.

Weak order $\succcurlyeq_P$:

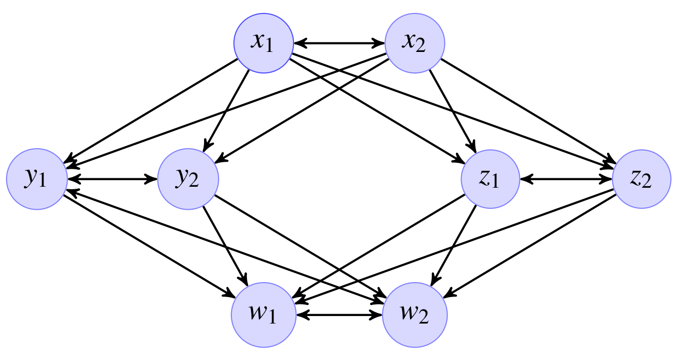


Simple order $\succcurlyeq^*_P$:

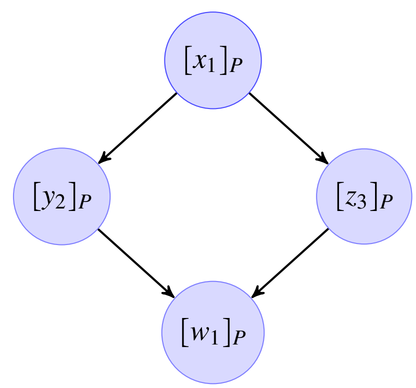


Similarly, $\succcurlyeq_Q$ in the figure below is a pre-order, since neither $y_1, y_2 \succcurlyeq_Q z_1, z_2$ nor $z_1, z_2 \succcurlyeq_Q y_1, y_2$. Its reduction $\succcurlyeq^*_Q$ is a partial order.

Pre-order $\succcurlyeq_Q$:



Partial Order $\succcurlyeq^*_Q$:



The binary relations $\succcurlyeq_{size}$ and $\succcurlyeq_{loudness}$ can also be thought of as part of algebraic STRUCTURES based on a domain X and a binary relation. For present purposes, we can think of a structure as a tuple $\langle X, \succcurlyeq_P, \ldots \rangle$, where X is any set, $\succcurlyeq_P$ is a n-ary relation on X (so, in the binary case, for all $(x,y) \in \succcurlyeq_P, x \in X$ and $y \in X$), and further members of the tuple represent other sets, relations on X, or (possibly partial) operations on X (such as concatenation, to be defined shortly). For example, the property **size** can be thought of extensionally as a structure $\langle X, \succcurlyeq_{size}, \ldots \rangle$, where X is some set of objects of which it makes sense to talk about their size, and the ellipsis represents additional constraints on this property that we may choose to add later.

With a larger domains of objects, we can start to ask more interesting questions about the structure $\langle X, \succcurlyeq_P, \ldots \rangle$. One question is whether we can fill in the ellipsis with a CONCATENATION operation on X with useful properties. Formally, concatenation is a ternary relation $\circ \subseteq X \times X \times X$

which obeys a set of axioms to be discussed in §2.2.1 below. z is required to be unique given x and y — that is, $\circ$ can also be thought of equivalently as a function $\circ : X \times X \to X$. I will usually abbreviate $(x,y,z) \in \circ$ as $x \circ y = z$.

Intuitively, concatentation can be thought of taking two objects and forming a complex object from them (although there is no need for them to be physical objects, and concatenation need not involve any procedure as in this metaphor). For example, if we are measuring the lengths of a number of rods x, y, and z we might ask not only whether $x \succcurlyeq_{length} y$, $x \succcurlyeq_{length} z$, and $y \succcurlyeq_{length} z$, but also how x compares in length to y and z placed end-to-end lengthwise, i.e. the concatenation $y \circ z$. Similarly if we are measuring the weights of the same objects using a balance, we can find out whether $x \succcurlyeq_{weight} y$ by considering whether the pan drops to the right when x is placed on the left pan and y is placed on the right; if it does not, then $x \succcurlyeq_{weight} y$. Equally, though, we can compare the weight of x to the concatenation of y and z by placing x on the left-hand side of the balance and both y and z on the right-hand side. If the pan drops right, then $(y \circ z) \succ x$.

Concatenation is important in a theory of measurement for natural language because, when concatenations are possible, we would like to know whether there is any interesting connection between the position of x and y in the relation $\succcurlyeq_P$ and the position of $x \circ y$. In particular, I will argue below that natural languages treat concatenation as used in RTM as a restricted version of the join operation. Joins are frequently encountered in natural languages: on relatively standard assumptions, it occurs in English as *or* in Boolean domains and as *and* in non-Boolean domains (as in the example of the balance in the last paragraph). As a result, we may expect different concatenation structures to trigger different interactions with *or* and *and* as appropriate. For many properties of interest — e.g., measurements of size, length, and weight — $\succcurlyeq_P$ will be ADDITIVE with respect to concatenation.

Although additivity is the most widely recognized effect of concatenation, I will argue below that certain expressions interact in a different way with concatenation, which I will call INTERMEDIATE. A property P is intermediate with respect to concatenation if $x \succcurlyeq_P y$ implies $x \succcurlyeq_P (x \circ y) \succcurlyeq_P y$. Although this type of concatenation has received very little attention in RTM (to my knowledge, it is only acknowledged by Luce & Narens (1985)) and none at all in natural language semantics, I will argue below that the familiar properties of temperature, danger, obligation, and desire interact in this way with concatenation. This is semantically important because additive properties are **upward monotonic**, while intermediate properties are **non-monotonic**.

2.1.2 Measure Functions, Interpretability, and the Typology of Scales in RTM

In the previous subsection we managed to introduce some basic concepts of measurement without mentioning numbers or degrees. Even though the idea of measurement without numbers or degrees may seem odd, the goal of RTM is to justify the use of these constructs in scientific practice, and for this purpose it would obviously be unwise to use them in the definitions. Rather, numerical measurements are justified by showing the existence of a homomorphism μ from a qualitative structure $\langle X, \succcurlyeq_P, \ldots \rangle$ into a structure making use of numbers such as $\langle \mathbb{R}, \geq, \ldots \rangle$.

μ is a **homomorphism** of a qualitative structure $\mathcal{S}_P = \langle X, \succcurlyeq_P \rangle$ into a numerical structure $\langle \mathbb{R}, \geq \rangle$ if and only if, for all $x, y \in X$,

- $\mu(x) \in \mathbb{R}$, $\mu(y) \in \mathbb{R}$, and
- If $x \succcurlyeq_P y$, then $\mu(x) \geq \mu(y)$.

If $\mathcal{S}_P$ also contains further relations or operations, then similar conditions apply. For instance, if $\mathcal{S}'_P = \langle X, \succcurlyeq, \circ \rangle$, where $\circ$ is a binary operation on X, then μ is a homomorphism from $\mathcal{S}'_P$ into $\langle \mathbb{R}, \geq, + \rangle$ if and only if in addition, for all $x, y, z \in X$,

- If $x \circ y = z$, then $\mu(x) + \mu(y) = \mu(z)$.

Intuitively, the requirement that μ be a homomorphism limits us to candidate μ which preserve all of the information contained in the source structure, while possibly adding more due to the structure inherent in the real numbers.

To eliminate the extra structure adhering to representations in the real numbers, we have to consider not just the information contained in one particular homomorphism μ, but the information that is common to all homomorphisms. That is, we will recapture the fact that some of our scales have a less rich structure than $\mathbb{R}$ by universally quantifying over homomorphisms from $\langle X, \succcurlyeq_P, \ldots \rangle$ into $\langle \mathbb{R}, \geq, \ldots \rangle$.

The empirical payoff is that we are able to represent contrasts between scales like temperature and clock time (where interval comparisons makes sense, but ratios comparisons do not) and scales like height (where both types of statements make sense).

(2.5) Sam grew from 2 feet to 3 feet, and Harry grew from 4 feet to 6 feet.
 a. ✓ So, Harry grew twice as much as Sam did.
 b. ✓ So, Harry is now twice as tall as Sam is.

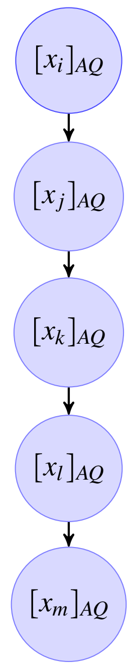
(2.6) I ran from 2PM to 3PM, and you ran from 4PM to 6PM.
 a. ✓ So, you ran for twice as long as I did.
 b. # So, you started running twice as late as I did.

As we will see, this difference in interpretability is explained by a qualitative difference in the underlying structure of the scales.

2.1.2.1 Ordinal Scales and Interpretability

For example, consider a structure $\langle X, \succcurlyeq_{AQ} \rangle$ where X is the set of cities in the United States with population over 1,000,000 and $x \succcurlyeq_{AQ} y$ is interpreted as "x has air quality as least as good as y". (The example is from Roberts 1979; this is apparently a measurement system which was actually employed in some locales in the 1970's.) Assume that $\succcurlyeq_{AQ}$ is a weak order whose reduction $\succcurlyeq^*_{AQ}$ is a simple order on the following equivalence classes:

$\underline{\succcurlyeq^*_{AQ}:}$



A homomorphism from $\langle X, \succcurlyeq_{AQ} \rangle$ into $\langle \mathbb{R}, \geq \rangle$ is any function μ with domain X and range $\mathbb{R}$ where $x \succcurlyeq_{AQ} y$ if and only if $\mu(x) \geq \mu(y)$. So, for example, the following are all homomorphisms (assuming that all cities in an equivalence class are mapped to the same number).

$$\mu_1 = \left\{ \begin{array}{lc} x_i \longrightarrow & 5 \\ x_j \longrightarrow & 4 \\ x_k \longrightarrow & 3 \\ x_l \longrightarrow & 2 \\ x_m \longrightarrow & 1 \\ \quad \ldots & \end{array} \right\} \qquad \mu_2 = \left\{ \begin{array}{lc} x_i \longrightarrow & 10 \\ x_j \longrightarrow & 6 \\ x_k \longrightarrow & 2 \\ x_l \longrightarrow & 1 \\ x_m \longrightarrow & 0 \\ \quad \ldots & \end{array} \right\} \qquad \mu_3 = \left\{ \begin{array}{lc} x_i \longrightarrow & 2{,}048{,}348 \\ x_j \longrightarrow & 194 \\ x_k \longrightarrow & 193 \\ x_l \longrightarrow & 22 \\ x_m \longrightarrow & -438 \\ \quad \ldots & \end{array} \right\}$$

If μ is a homomorphism from a structure $\langle X, \ldots \rangle$ into $\langle \mathbb{R}, \ldots \rangle$, we will say that μ is an **admissible measure function** on $\langle X, \ldots \rangle$. (The concept of admissibility will appear frequently in the rest of the dissertation.)

$\langle X, \succcurlyeq_{AQ} \rangle$ is an example of an ORDINAL SCALE, one of the weakest scale types standardly employed in RTM:

(2.7) If a structure $\langle X, \succcurlyeq_P \rangle$ is an **ordinal scale** then, for all admissible μ, $x \succcurlyeq_P y \equiv \mu(x) \geq \mu(y)$.

Every weak order has at least as much structure as an ordinal scale. In the case of the weak order $\succcurlyeq_{AQ}$, for example, (2.7) is clearly satisfied: for example, since $x_j \succ_{AQ} x_k$ we have $\mu_1(x_j) = 4 > \mu_1(x_l) = 2$, $\mu_3(x_j) = 194 > \mu_3(x_l) = 22$, etc.[1]

1 Note that stronger scale types, such as ratio scales and interval scales to be defined below, also have this property. I use an "if ... then" statement here in order to avoid overlap between scale types, but this is just a matter of definition: we could equally well define the scale types so that all interval and ratio scales are also ordinal scales, for example.

Another way to characterize an ordinal scale is in terms of transformations among the admissible measure functions: any monotone increasing transformation of an admissible μ is also an admissible μ.

(2.8) If a structure $\langle X, \succcurlyeq_P \rangle$ is an **ordinal scale** then, for all admissible measure functions μ and all order-preserving (monotone increasing) functions $f : \mathbb{R} \to \mathbb{R}$:
$\mu'(x) = f(\mu(x))$ is also an admissible measure function.

(2.8) captures the fact, already apparent from the air quality example, that the relative distance between the measures of objects is not important in an ordinal scale, but only the relative ordering of the measures assigned to objects. (See Krantz et al. (1971); Roberts (1979) for proofs of the equivalence of the conditions in (2.7) and (2.8).)

One consequence of the relatively weak structure of an ordinal scale is that many quantitative statements that can be framed using the numbers assigned to objects do not have a stable truth-value across admissible μ. For example, consider the statements:

(2.9) a. x_j has better air than x_k does.
b. x_m has better air than x_i does.
c. x_j has air twice as good as x_l does.

Are these statements true or false? Well, if we looked only at μ_1, the first would appear to be true, the second false, and the third true. If we consider μ_2 and μ_3, however, (2.9a) and (2.9b) remain true and false respectively, but (2.9c) comes out false. Since all three measure functions are of equal status, a natural move is to declare (2.9a) to be true and (2.9b) false, but to conclude that (2.9c) does not have a truth-value relative to this structure.

In RTM this type of situation is usually described by saying that (2.9a) and (2.9b) are "meaningful" while (2.9c) is "meaningless". Since this term bears a good deal of weight already both in ordinary usage and formal semantics, I will formulate this notion instead as a constraint on semantic interpretability:

(2.10) A statement S is semantically **interpretable** only if its truth-value remains the same under all admissible μ.[2]

Since all homomorphisms from $\langle X, \succcurlyeq_P \rangle$ into $\langle \mathbb{R}, \geq \rangle$ are admissible, the effect of (2.10) is that statements which add extra quantitative information beyond what is contained in $\langle X, \succcurlyeq_P \rangle$ are ignored. That is, if a measure function shows some patterned behavior in $\mathbb{R}$, this is ignored unless the same pattern is also observed in the underlying qualitative structure. This makes it possible to use real numbers, which have a very rich structure, to represent poorer structures accurately.

2 There is a clear connection between the measurement-theoretic notion of "meaningfulness" and a three-valued logic based on supervaluations, explored — minus the label "supervaluation" — in Suppes (1959). Interestingly, this article predates considerably the source to which this idea is usually attributed (van Fraassen 1966, 1968). Given that three-valued logics are often used to treat presupposition, we might even exploit this connection by stipulating that a statement presupposes its own "meaningfulness", in the technical measurement-theoretic sense. I will not pursue the connection with presupposition here, however.

2.1.2.2 Ratio Scales

Treating (2.9c) as undefined seems fine, but there are other similar statements which should clearly get truth-values, for example:

(2.11) a. Sam is twice as tall as his little brother.
b. I ran 4.3 times as far as you did.
c. It is three times as likely to rain as it is to snow.

What sorts of structures are needed to make these statements come out as interpretable? It turns out that a sufficient condition for statements like *x is n times as P as y* to have stable truth-conditions across all admissible μ is that the scale in question be a RATIO SCALE. Ratio scales can be characterized easily in terms of admissible measure functions:

(2.12) If a structure $\langle X, \succcurlyeq_P, \circ, ...\rangle$ is a **ratio scale** then, for all admissible μ and all $x \in X$: $\mu'(x) = n \times \mu(x)$ is also admissible, where $n \in \mathbb{R}^+$ (the positive real numbers).

Ratio scales are those where all and only admissible transformations are those which involve multiplying each value by the same positive real number. Familiar examples of ratio scales include measurements of extent (length, width, height, etc.), mass, and weight. For example, measurements of extent in feet and meters can be converted by using the transformation

$$\text{Length in feet} = 0.3048 \times \text{length in meters.}$$

and its converse

$$\text{Length in meters} = \frac{1}{0.3048} \times \text{length in feet} \approxeq 3.2808 \times \text{length in feet.}$$

Similarly, in ordinary usage measurements in pounds and kilograms can be converted without loss of information by the transformation

$$\text{Weight in kilograms} = 2.2 \times \text{weight in pounds.}$$

$$\text{Weight in pounds} = \frac{1}{2.2} \times \text{weight in kilograms} \approxeq 0.455 \times \text{weight in kilograms.}$$ [3]

If μ and μ' are both admissible measure functions for some ratio scale $\langle X, \succcurlyeq_P, \circ, ...\rangle$, then the ratio $\frac{\mu(x)}{\mu(y)}$ is guaranteed to be equal to the ratio $\frac{\mu'(x)}{\mu'(y)}$. This is because $\mu'(x) = n \times \mu(x)$ for some $n > 0$, and so

$$\frac{\mu'(x)}{\mu'(y)} = \frac{n \times \mu(x)}{n \times \mu(y)} = \frac{\mu(x)}{\mu(y)}$$

As a result, statements involving ratios like (2.11) are predicted to be interpretable for ratio scales, since they will be true or false in every admissible μ.

Another important property of ratio scales is that they are additive with respect to concatenation: that is, as long as two objects do not overlap, then the measure assigned to their concatenation is the sum of their individual measures.

3 Obviously in scientific usage kilograms are a measure of mass rather than weight, and so the transformation would have to take into account gravity; but this is probably not linguistically relevant.

(2.13) A scale $\langle X, \succcurlyeq_P, \circ, \ldots \rangle$ is **additive** with respect to concatenations iff, for all non-overlapping x, y and all admissible μ, $\mu(x \circ y) = \mu(x) + \mu(y)$.

See Roberts & Luce (1968); Krantz et al. (1971) for a proof that ratio scales are additive.

Two properties of ratio scales are worth noting here. First, it follows from additivity that $\mu(x \circ y) > \mu(x)$ and $\mu(x \circ y) > \mu(y)$, and so that $(x \circ y) \succ_P x$ and $(x \circ y) \succ_P y$. This contrasts with some familiar types of interval scales, as we will see in a moment. Second, ratio scales have a fixed minimum corresponding to $\mu(x) = 0$ under all admissible μ.

In keeping with the reductive spirit of RTM, it is possible to give qualitative axioms which define ratio scales without reference to numerical measures, and to show that they are equivalent to the characterization in (2.12). However, the axiomatization would not be very enlightening here without an excessive amount of discussion, and is relegated to a footnote.[4]

The characterization of ratio scales as "stronger" than ordinal scales, or as "giving more information", can now be made precise as follows: a scale $\mathcal{S}$ is stronger than a scale $\mathcal{S}'$ just in case the permissible transformations of admissible measure functions on $\mathcal{S}$ are a proper subset of permissible transformations of admissible measure functions on $\mathcal{S}'$. The permissible transformations $\mu'(x) = n \times \mu(x), n > 0$ on a ratio scale are all monotone increasing, and so all such transformations are permissible for ordinal scales as well. However, many other monotone increasing transformations are not permissible for ratio scales, so that the latter are a stronger scale type by this definition.

2.1.2.3 Interval Scales

There are certain properties of interest whose scales seem to be stronger than an ordinal scale, but weaker than a ratio scale. Temperature and clock time are two familiar examples: if it is 10 degrees Celsius in Boston and 30 degrees Celsius in Atlanta, it is not natural to describe this situation using (2.14).

---

4 One way to axiomatize a ratio scale is (Roberts 1979):

- $\succcurlyeq_P$ is a weak order (complete, transitive, and reflexive);
- $\circ$ is
  - associative: $\forall abc \in X[a \circ (b \circ c)] \approx_P [(a \circ b) \circ c]$,
  - monotonic: $\forall abc \in X[a \succcurlyeq_P b \equiv (a \circ c) \succcurlyeq_P (b \circ c) \equiv (c \circ a) \succcurlyeq_P (c \circ b)]$,
  - Archimedean: $\forall abcd \in X[a \succ_P b \rightarrow \exists n > 0 : (na \circ c) \succcurlyeq_P (nb \circ d)]$.

See Roberts & Luce (1968); Krantz et al. (1971) for the proof that these are sufficient conditions for constituting a ratio scale.

Note that *na* in the Archimedean condition is not numerical multiplication, but an abbreviation for "the concatenation of *n* non-overlapping objects which are in the $\succcurlyeq$-equivalence class of *a*".

The Archimedean condition given here implies that *X* is infinite and that $\succcurlyeq_P$ has no maximal element and no *x* s.t. $\mu(x) = 0$. These are not really necessary properties of ratio scales, however: Krantz et al. (1971: ch.3) give an alternative formulation of the Archimedean condition which is compatible with finiteness and upper- and lower-boundedness, which will also be discussed in §2.2.1 below. This characterization is more appropriate for natural language, since we have both ratio scales with elements occupying their minimum (e.g. *expensive*) and ones with elements occupying their maximum (*open, closed, full, empty*).

(2.14) # It is three times as hot in Atlanta as it is in Boston.

The formal explanation offered by RTM of the oddity of (2.14) is, of course, that translating the temperature measurements into another equally valid measurement system for temperature such as Fahrenheit could make (2.14) false, violating the condition on interpretability in (2.10). Specifically, Fahrenheit and Celsius measurements are related by the transformation

$$\text{Temperature in Fahrenheit} = \frac{9}{5} \times \text{temperature in Celsius} + 32$$

Under this transformation, the temperatures in Atlanta and Boston do not have the ratio $\frac{30}{10} = 3$, but $\frac{86}{50} = 1.72$. Since the truth-value of (2.14) is not stable under this transformation, we conclude that this statement and other statements involving non-trivial ratios are uninterpretable.

However, we do not want temperature to have as little structure as an ordinal scale. Not just any monotone increasing transformation of Celsius would give us a usable measurement system for temperature; we need one that preserves information about differences, as in:

(2.15) It is 30 Celsius in Atlanta, 10 Celsius in Boston, 35 Celsius in Rio de Janeiro, and 25 Celsius in Rome. So, Atlanta is hotter than Boston by twice as much as Rio is hotter than Rome.

Even though the absolute ratio statement (2.14) does not keep its truth-value in the transformation from one admissible measure function (Celsius) to another (Fahrenheit), the **difference ratio** in (2.15) does: the Celsius ratio $\frac{30-10}{35-25} = 2$ is equal to the Fahrenheit ratio $\frac{86-50}{95-77} = 2$.

We cannot associate temperature with an ordinal scale if we want to explain the fact that the relative size of intervals on the temperature scale are stable quantities across admissible μ: allowing all monotone increasing transformations would destroy this information. Similar considerations hold, for example, for clock time, where the numbers assigned to points in time are not interpretable, nor are their ratios, but the relative sizes of intervals are interpretable quantities:

(2.16) I ran from 3PM to 4PM, and you ran from 6PM to 8PM.

a. # So, you started running twice as late as I did.

b. So, you ran for twice as long as I did.

In order to capture these features using RTM, temperature and time are standardly associated with INTERVAL SCALES, which I will also characterize using the class of admissible transformations.

(2.17) If $\mathcal{S}_P$ is an **interval scale** then, whenever μ is admissible for $\mathcal{S}_P$, for all μ':
If $\mu'(x) = \alpha \times \mu(x) + \beta$ for any $\alpha \in \mathbb{R}^+$ and $\beta \in \mathbb{R}$, then μ' is also admissible for $\mathcal{S}_P$.

The conversion from Celsius to Fahrenheit given above is an example of such a transformation, setting $\alpha = \frac{9}{5}$ and $\beta = 32$. Another example of an interval scale which we will make considerable use of in later chapters is **expected utility**. Algebraic manipulation shows that, with this class of transformations, ratios of differences will always be interpretable with interval scales, but absolute ratios will be interpretable only in the trivial case where $\mu(x) = \mu(y)$.

Equivalently, we can think of interval scales as structures $\langle X, Y, \succcurlyeq_P \rangle$, where $Y \subseteq X \times X$ is a set of pairs of objects in X, and $\succcurlyeq_P$ is a binary relation on Y. The relation

$$(a,b) \succcurlyeq_P (c,d)$$

can be read "*a* exceeds *b* with respect to property *P* by more than *c* exceeds *d*". $\succcurlyeq_P$ is required to be a weak order and obey four further axioms.[5] We then define the class of admissible measure functions μ in terms of this structure by

$$(a,b) \succcurlyeq_P (c,d) \quad \text{if and only if} \quad [\mu(a) - \mu(b)] \geq [\mu(c) - \mu(d)]$$

See Krantz et al. (1971) for the proof that this condition picks out the same set of measure functions that were characterized in (2.17).

Naturally, we want to be able to make simple comparisons between individual objects in terms of interval-scale properties like temperature and time. We can do this in two ways. Using measure functions, we can say that x has greater temperature than y iff $\mu(x) > \mu(y)$ for all admissible μ. An equivalent characterization without numbers is:

$$x \textit{ is warmer than } y \text{ if and only if } \exists z : (x,z) \succ_{temp} (y,z)$$

That is, if you pick some z which is cooler than both x and y, then x is warmer than y if and only if the difference in temperature between x and z is greater than the difference between y and z.

These three scale types — ordinal, ratio, and interval — are the three standard scale types in RTM that are most relevant for us here. In the following sections we will explore some ways of enriching the representations to connect them directly with standard assumptions in natural language semantics, focusing on concatenation and boundedness.

2.2 Measurement Theory and Natural Language Scales

2.2.1 Positivity, Intermediacy, and Concatenation

Of the three scale types that we saw in the previous section, only one was associated with a concatenation operation: the ratio scale. As Luce & Narens (1985) note, Krantz et al. (1971) and other measurement theorists have generally assumed a very restrictive definition of concatenation which essentially limited concatenation to infinite additive structures (e.g., infinite ratio scales). The standard assumptions are (Krantz et al. 1971: 72ff.):

(2.18) a. **Closure**: $\forall x \forall y \exists z : x \circ y = z$
b. **Associativity**: $\forall x \forall y \forall z : (x \circ y) \circ z \approx x \circ (y \circ z)$
c. **Monotonicity**: $\forall x \forall y \forall z : x \succcurlyeq y \equiv (x \circ z) \succcurlyeq (y \circ z) \equiv (z \circ x) \succcurlyeq (z \circ y)$
d. **Positivity**: $\forall x \forall y : x \circ y \succ x$
e. **Archimedean**: If $a \succ b$ then, for all c and d, there is a positive integer n s.t. $na \circ c \succcurlyeq nb \circ d$, where na is defined inductively as: $1a = a$, $(n+1)a = na \circ a$.

5 Again following Roberts's (1979) presentation,
Axiom 2: $\forall abcd \in X : (a,b) \succcurlyeq_P (c,d) \rightarrow (d,c) \succcurlyeq_P (b,a)$
Axiom 3 (weak monotonicity): $\forall abca'b'c' \in X : [(a,b) \succcurlyeq_P (a',b') \wedge (b,c) \succcurlyeq_P (b',c')] \rightarrow (a,c) \succcurlyeq_P (a',c')$
Axiom 4 (solvability): $\forall abcd \in X : [(a,b) \succcurlyeq_P (c,d) \wedge (c,d) \succcurlyeq_P (x,x)] \rightarrow \exists ef \in X[(a,e) \succcurlyeq_P (c,d) \wedge (f,b) \succcurlyeq_P (c,d)]$
Axiom 5 (Archimedean): Every strictly bounded standard sequence is finite. [This is a modification of the Archimedean axiom in footnote 3 for domains without a zero point; see (Krantz et al. 1971: 83-85), Roberts (1979: 137-138) for formal details and discussion.]

Most clearly problematic for are purposes are **positivity**, the assumption that the concatenation of two objects is strictly larger than either; and **closure**, the assumption that the domain X is closed under $\circ$. A third assumption, implicit in the Archimedean axiom and standard in RTM, is that the concatenation of an object with itself can form a third object: $(x \circ x)$ does not necessarily equal x, and indeed, given positivity, cannot. (The obvious conceptual problems with self-concatenation are usually finessed by saying that this procedure corresponds to concatenation of an object with an exact duplicate of itself; cf. also Klein 1991.)

Together, these conditions entail that any non-trivial structure with a concatenation operation is infinite and is not upper-bounded. This is because, by closure and non-triviality, $x \circ y$ exists; by positivity, $x \circ y > x$; irreflexivity of $>$ implies $x \circ y \neq x$. By closure, $(x \circ y) \circ x$ exists; positivity implies $(x \circ y) \circ x > (x \circ y)$; irreflexivity of $>$ implies $(x \circ y) \circ x \neq (x \circ y)$; closure implies $((x \circ y) \circ x) \circ x$ exists; etc. This procedure can be iterated forever to create bigger and bigger objects.

The requirement of an infinite domain may well be problematic. Non-upper-boundedness definitely is, because — as we saw in chapter 1 — Kennedy & McNally (2005); Kennedy (2007) have shown that many scales in natural language are upper-bounded (e.g., the scales associated with *full, safe, clean* and telic verbs). We would like to be able to associate these scales with a concatenation operation if appropriate, but positivity and closure of the concatenation operation appear to make this impossible.

A further problem associated with the assumption of positivity is that there are properties for which concatenation is intuitively a meaningful operation, but additivity (and positivity more generally) would license clearly incorrect entailments. Suppose you have two bowls of soup x and y, and you pour one into the other to form a bowl z. If we ask about the volume of soup in z, the answer is straightforward: z is equal to $(x \circ y)$, the concatenation of x and y, and its volume is just the volume of x plus the volume of y. This clearly satisfies the assumptions in (2.18), in particular positivity: if we know the volume of x and y we can infer that the volume of z is greater than the volume of either x or y.

If we ask about the temperature of z instead, however, things are more difficult. If temperature had a standard concatenation operation, then we would expect the following inference to be valid:

(2.19) a. This bowl of soup is 40 degrees Celsius. That one is 20 degrees Celsius.
b. # So, if we pour one into the other the result will be a bowl which is more than 40 degrees Celsius.

This reasoning is clearly invalid: not only is temperature not additive with respect to join, it is not even positive.

Another example where positivity fails is the property of danger. For properties P for which concatenation is positive, we can validly infer that $x = (y_1 \circ y_2 ... \circ y_n)$ is P to a greater degree than any of its proper parts: $(x >_P y_1) \wedge (x >_P y_2) \wedge ... \wedge (x >_P y_n)$. For additive properties such as size, this inference seems trivial:

(2.20) a. Fulton County, Georgia has an area of 535 square miles.
b. Adjacent Cobb County, Georgia has an area of 345 square miles.
c. So, Fulton and Cobb Counties taken together have an area greater than 535 square miles.

However, the inference is invalid when we consider danger rather than size:

(2.21) a. Fulton County is extremely dangerous.
b. Cobb County is quite safe (only slightly dangerous).
c. # So, Fulton and Cobb Counties taken together are extremely dangerous or worse.

The usual measurement-theoretic response to facts like these would be to conclude that temperature and danger are associated with structures that do not contain a concatenation operation. I think that this would be too hasty, however. By virtue of our understanding of the properties **temperature** and **danger**, we do have clear intuitions about how the temperature or danger of a complex object relates to the temperature or danger of its component parts. In both cases, concatenation intuitively produces an object with an INTERMEDIATE degree of the property in question:

(2.22) a. This bowl of soup is 40 degrees Celsius. That one is 20 degrees Celsius.
b. So, if we pour one into the other the result will be a bowl which is somewhere between 20 and 40 degrees Celsius.

(2.23) a. Fulton County is extremely dangerous.
b. Cobb County is quite safe (only slightly dangerous).
c. So, Fulton and Cobb Counties taken together are somewhere between quite safe and extremely dangerous (e.g., moderately dangerous).

Basically, it looks as if instead of obeying the positivity assumption, these properties respond to concatenation as in (2.24):

(2.24) **Intermediacy**: If $x \circ y$ is defined, then $x \succcurlyeq (x \circ y) \succcurlyeq y$.

If temperature and danger were associated with structures with no concatenation operation, it would not be possible to even make sense of this condition. Further, the standard interpretation of concatenation is not compatible with intermediacy.[6]

In the rest of this section I will argue that all scale types used in natural language are able to make use of a concatenation operation. Variability in how scales respond to concatenation stems from what further axioms a structure obeys, and in particular whether the structure is positive or intermediate with respect to concatenation. I will connect this claim with standard assumptions about the algebraic structure of various domains in natural language semantics, showing that we can essentially treat concatenation as the operation **join**. The claims that all scales are able to make use of concatenation structures connects with the fact that domains come equipped with a join operation; this will allow us to derive predictions about the interaction of different scale types with expressions which make use of the join operation, notably *and* and *or*.

2.2.2 Concatenation and Join

Many domains standardly used in natural language semantics are Boolean, i.e. have a type ending in t. It is well-known that these domains have a structure which is isomorphic to a BOOLEAN ALGEBRA:

6 Luce & Narens (1985) call this property "intern".

(2.25) A **Boolean algebra** is a structure $\langle X, \vee, \wedge, \neg, \bot, \top \rangle$, which obeys the following axioms as well as their inverses:

a. **Associativity**: $a \vee (b \vee c) = (a \vee b) \vee c$

b. **Commutativity**: $a \vee b = b \vee a$

c. **Absorption**: $a \vee (a \wedge b) = a$

d. **Distributivity**: $a \vee (b \wedge c) = (a \vee b) \wedge (a \vee c)$

e. **Complements**: $a \vee \neg a = \top$

(The "inverses" of these axioms are the formulas that you get if you interchange $\vee$ and $\wedge$ throughout and change $\top$ to $\bot$ in (2.25e).)

The fact that Boolean domains have this structure is closely related to the fact that expressions whose type ends in t are in a 1-to-1 relationship with sets: e.g., the type $\langle e,t \rangle$ expression $\lambda x_e [P(x)]$ is the characteristic function of the set $\{x \mid P(x)\}$. The possible expressions of type $\langle e,t \rangle$ form an algebra (isomorphic to) $\langle \mathcal{P}(\mathbf{D}_e), \cup, \cap, -, \varnothing, \mathbf{D}_e \rangle$, where meet $\wedge$ is identified with set intersection $\cap$, join $\vee$ with set union $\cup$, and complement $\neg$ with set complement $-$.

Not all natural language domains have this structure, however. For example, the possible expressions of type e have no inherent structure in Montague's (1973) theory. In the Boolean semantics of Keenan & Faltz (1985), this is taken as a point in favor of eliminating type e from the object language altogether. Although there is a set of individuals X whose powerset algebra forms the domain of type $\langle e,t \rangle$ expressions, expressions usually assigned type e such as *Barack Obama* or *the Queen of England* do not receive interpretations in type e, but in $\langle \langle e,t \rangle, t \rangle$: for instance, $[\![\textit{Barack Obama}]\!]^{\mathcal{M},w,g} = \{P \mid P(\mathbf{Obama})\}$, the set of properties that the real-world individual associated with the name "Barack Obama" has.

Once of the nice consequences of this approach is that *or* can be interpreted as the join operation, and *and* with the meet operation, in any Boolean type, including the type $\langle \langle e,t \rangle, t \rangle$ in which individuals denote in Keenan & Faltz's (1985) theory.

(2.26) a. $[\![\textit{run and jump}]\!]^{\mathcal{M},w,g} = \{x \mid \mathbf{run}(x)\} \cap \{y : \mathbf{jump}(y)\}$

b. $[\![\textit{Barack and Michelle}]\!]^{\mathcal{M},w,g} = \{P \mid P(\mathbf{Barack})\} \cap \{Q : Q(\mathbf{Michelle})\}$

From here we get the equivalence *Barack and Michelle like carrots* $\leftrightarrow$ *Barack likes carrots and Michelle likes carrots*: the property denoted by *like carrots* just has to be in the intersection of the set of properties that Barack has and the set of properties that Michelle has.

However, Link (1983) points out that treating *and* as denoting meet in all domains equally leads — at least in the simple for sketched here — to incorrect predictions about collective predicates such as intransitive *meet*, where the equivalence between *P(x and y)* and *P(x) and P(y)* does not hold.

(2.27) a. Barack and Michelle met in 1989.

b. Barack met in 1989 and Michelle met in 1989.

Not only does (2.27b) not mean the same as (2.27a), it's not clear that it is even intelligible. Link argues that we should abandon the assumption that type e is unstructured, treating it instead as having a part-whole structure: the individuals **Barack** and **Michelle** are proper parts of the compound individual (**Barack** $\sqcup_i$ **Michelle**), the INDIVIDUAL JOIN (i-join) of Barack and Michelle. We can then translate (2.27a) as saying that the compound individual (**Barack** $\sqcup_i$ **Michelle**) participated in a meeting event which occurred in 1989. Distributive predicates such as *like carrots*, on the other hand, have the special property that they hold of a compound individual if and only they hold of all of the atoms which make up the compound individual, so that the equivalence *P(x and y) iff P(x) and P(y)* holds only in this special case.

The structure that Link ascribes to the domain of (countable) individuals treats it as an UNBOUNDED JOIN SEMILATTICE. An unbounded join semilattice resembles a Boolean algebra without the bottom element $\bot$ (i.e. $\varnothing$, in the case of a powerset algebra). As the name suggests, these structures are closed under the join operation, but they are not closed under the operations meet and complement. They are characterized as follows (where $x \sqsubseteq y$ abbreviates $x \sqcup y = y$):

(2.28) An **unbounded join semilattice** is a structure $\langle X, \sqcup \rangle$, where X is a non-empty set, and $\sqcup$ obeys

a. **Closure**: $\forall x \forall y \exists z : x \sqcup y = z$

b. **Commutativity**: $\forall x \forall y : x \sqcup y = y \sqcup x$

c. **Associativity**: $\forall x \forall y \forall z : x \sqcup (y \sqcup z) = (x \sqcup y) \sqcup z$

d. **Idempotency**: $\forall x : x \sqcup x = x$

e. **No Bottom Element**: $\neg \exists x \forall y : x \sqsubseteq y$

Mass and count nouns are interpreted similarly, except that the count domain has atoms (elements x s.t. $y \sqsubseteq x$ implies $y = x$) while the domain of mass nouns does not.

What is interesting about Link's proposal for the purposes of a measurement-theoretic semantics of degree is this. If Link is correct, the domain of individuals also comes equipped with a join operation. Furthermore, there are close intuitive and formal correspondences between the join operation in structured domains and the concatenation operation in RTM. On the intuitive level, it is obvious that we want the weight of John and the weight of Mary to have a systematic connection with the weight of the complex individual (**John** $\sqcup_i$ **Mary**): it should be the sum of their individual weights. This is, of course, exactly what the axiomatization of concatenation for additive structures yields.

The intuitive connection between concatenation and join was noted by Krifka (1989: 79), who suggests in his RTM-inspired treatment of amount expressions that "concatenation ... can be defined as join ... restricted to non-overlapping individuals". Krantz et al. (1971: 208), discussing the axiomatization of qualitative probability, also point out the intuitive connection between concatenation and disjunction, which is join in Boolean domains, without drawing any formal connection. In fact, as noted above, it is not possible on assumptions standard in RTM to make this equation: concatenation is typically assumed to be positive and closed ((2.18d), (2.18a)) and these properties are incompatible with the idempotency of join ((2.28d), and entailed by the axioms in (2.25)). This is because $(x \circ x) \succ x$ by positivity and $\succ$ is irreflexive. Positivity and closure also together entail the infinity and unboundedness of the domain, as noted above.

The solution, I suggest, is to generalize the definition of concatenation so that it is not only applicable to additive structures such as ratio scales. This will allow us to treat concatenation as restricted join, as Krifka suggests, and also to capture intermediate concatenation as discussed in the previous section.

(2.29) A **concatenation structure** is a structure $\langle X, \succcurlyeq_P, \circ \rangle$, where $\succcurlyeq_P$ is a weak order on X and $\circ$ is a partial binary operation on X which obeys

a. **Disjoint Closure**: $x \circ y$ is defined iff $x, y \in X$ and x and y do not overlap.

b. **Associativity**: $\forall x \forall y \forall z : (x \circ y) \circ z \approx_P x \circ (y \circ z)$ (when defined)

c. **Monotonicity**: $\forall x \forall y \forall z : x \succcurlyeq_P y \equiv (x \circ z) \succcurlyeq_P (y \circ z) \equiv (z \circ x) \succcurlyeq_P (z \circ y)$ (when defined)

d. **Archimedean**: Every strictly bounded standard sequence is finite.[7]

This weakening of the concatenation axioms follows the discussion in Luce & Narens (1985: 10-11) generally, but does not abandon as many of the standard assumptions as they do. A few notes: first, since positivity and full closure are not assumed and concatenation is restricted to non-overlapping objects, concatenation structures will be infinite only if the domain is. Second, idempotency, although part of the definition of join, is not a property of concatenation on this definition because $x \circ x$ is never defined (2.29a).

We can now state the relationship between concatenation and join (for the purpose of a degree semantics for English) as being simply:

(2.30) a. If $\langle X, \succcurlyeq_P, \circ \rangle$ is a concatenation structure, then $x \circ y$ is defined if and only if $x, y \in X$ and x and y are non-overlapping $(\neg \exists z : [z \vee x = x] \wedge [z \vee y = y])$.

b. If defined, $x \circ y = x \vee y$, where $\vee$ is the join operation on the domain of x and y.

$\vee$ here corresponds to set union $\cup$ in Boolean domains and individual join $\sqcup$ in the structured domains of individuals and masses. Note that the fact that concatenation is only partially closed is not a serious restriction because (as Krifka 1989 points out) when two objects x and y overlap partially, it is always possible to find two non-overlapping objects x and $y' = \mathbf{max}(\{z : z \leq y \wedge \neg(z \leq x)\})$.[8] These can then be concatenated, and their concatenation will be the same as the join of x and y.

(The reduction of) a concatenation structure in this sense satisfies the axioms for being an ordered local semigroup (Krantz et al. 1971: 44-45).[9] Krantz et al. prove that, if $\langle X, \succcurlyeq, \circ \rangle$ is an

7 This axiom needs a note of explanation. A standard sequence is a sequence of objects $a_1, a_2, ..., a_n$ such that the distance between each a_{m-1} and a_m is equal. The Archimedean property is needed here in order to rule out infinitesimal quantities, which would destroy additivity for ratio scales. This formulation of the Archimedean property accomplishes this without entailing non-upper-boundedness by simply requiring that, for any finite upper bound you choose, you cannot extend a standard sequence forever without eventually either exceeding that bound or running out of objects to add to the sequence. See Krantz et al. (1971: 83-84) for discussion.

8 $\leq$ means "part of" or "subset of" as appropriate here, and $\mathbf{max}(A) =_{df} \iota x[x \in A \wedge \forall y \in A : x \geq y]$.

9 Let B be the set of all pairs for which $\circ$ is defined, i.e. $B = \{(x,y) \mid \exists z : (x \circ y) = z\}$. An **ordered local semigroup** is a structure $\langle X, \succcurlyeq, B, \circ \rangle$ where $\succcurlyeq$ is a simple order and, for all $a, b, c, d \in X$,

Ax1. $(a,b) \in B \wedge a \succcurlyeq c \wedge b \succcurlyeq d \rightarrow (c,d) \in B.$

Ax2. $(c,a) \in B \wedge a \succcurlyeq b \rightarrow (c \circ a) \succcurlyeq (c \circ b).$

Ax3. $(a,c) \in B \wedge a \succcurlyeq b \rightarrow (a \circ c) \succcurlyeq (b \circ c).$

Ax4. $(a,b) \in B \wedge (a \circ b, c) \in B \equiv (b,c) \in B \wedge (a, b \circ c) \in B.$

Ax5. When both conditions in Ax4 hold, $(a \circ b) \circ c = a \circ (b \circ c).$

The reader should be able to convince him- or herself that Ax1-5 are satisfied by the interpretation of $\circ$ as disjoint join.

ordered local semigroup which is Archimedean, regular, and positive with respect to concatenation, then $\langle X, \succcurlyeq, \circ \rangle$ is a ratio scale (all admissible μ, μ' are additive and related by a transformation $\mu(x) = \alpha \times \mu'(x)$ for some $\alpha > 0$).

Since concatenation structures are required to be Archimedean by the definition in (2.29), the only properties of ratio scales missing from this definition are positivity and regularity. So we can define a ratio scale simply as a concatenation structure which is positive and regular:

(2.31) A **ratio scale** is a concatenation structure $\langle X, \succcurlyeq_P, \circ \rangle$ which also satisfies
  a. *Positivity*: $\forall x \forall y$: if $x \circ y$ is defined, then $x \circ y \succ_P x$.
  b. *Regularity*: $\forall x \forall y$: if $x \succ_P y$, then $\exists z : x \succcurlyeq_P (y \circ z)$.

Regularity is a solvability assumption which is needed to ensure that the domain is rich enough; positivity is the substantive axiom here.

However, since positivity is an extra feature of ratio scales rather than an inherent property of concatenation, we can now consider other ways that concatenation might interact with the order $\succcurlyeq$. In other words, (2.29) does not require $\mu(x \circ y)$ to equal $\mu(x) + \mu(y)$; we can imagine many other relationships that $\mu(x \circ y)$, the measure of the join of x and y, might have to $\mu(x)$ and $\mu(y)$, and see whether it is possible to formulate scales that have these properties. To the extent that these are attested, we have a new parameter of variation in scale type, in addition to the better-known boundedness parameters.

In fact, once we decide to treat concatenation as restricted join, it turns out that a variety of proposals have already been made as to this relationship. These include:

(2.32)
  a. **Additive**: $\mu(x \circ y) = \mu(x) + \mu(y)$
  b. **Superadditive**: $\mu(x \circ y) > \mu(x) + \mu(y)$
  c. **Subadditive**: $\mu(x) + \mu(y) > \mu(x \circ y) \geq \mu(x)$
  d. **Maximal**: If $x \succcurlyeq y$, then $(x \circ y) \approx x$
  e. **Intermediate**: If $x \succcurlyeq y$, then $(x) \succcurlyeq (x \circ y) \succcurlyeq y)$
  f. **Minimal**: If $x \succcurlyeq y$, then $(x \circ y) \approx y$
  g. **Subtractive**: If $x \succcurlyeq y$, then $\mu_P(x \circ y) = \mu_P(x) - \mu_{P^{-1}}(y)$, where $\mu_{P^{-1}}$ is a measure function associated with the antonym of μ_P with shared units.
  h. **Atomic Only**: $\succcurlyeq_p$ contains no concatenations, i.e. $x \succcurlyeq_P y$ implies that x, y are atomic.

All of these possibilities have either been proposed to account for natural language phenomena (implicitly or explicitly), or will be proposed in this dissertation.

For example, additive scales have already gotten a good deal of discussion, and are clearly relevant for many purposes. Super- and sub-additivity have been discussed in some detail in the psychological literature on probability judgment (e.g., Tversky & Koehler 1994; Macchi, Osherson & Krantz 1999). Subtractive scales have not been noticed previously, but they appear to be the correct characterization of the scales associated with the antonyms of additive adjectives, such as *short* and *light*. Also, I argued in the previous section that temperature and danger are intermediate with respect to concatenation; in chapter 6 I will add desire and obligation to the list.

Maximality with respect to concatenation will be relevant throughout the chapters on modals, since it is a property of Kratzer's semantics for modality when we restrict attention to connected parts of the ordering (although I will argue that it is not actually a correct characterization of the scales associated with either epistemic or deontic modals). For the same reason, Kratzer's semantics appears to predict that the scales associated with the antonyms of positive modal adjectives (*unlikely, bad*) should be minimal with respect to concatenation when the relevant propositions are comparable.

Finally, many gradable properties seem to be Atomic Only, i.e. distributive: if John and Mary are happier than Sue, this cannot mean, for instance, that the sum or the average of their happiness exceeds Sue's, but only that each of them is individually happier than Sue. (Note that a scale which is Atomic Only (2.32h) is technically not a concatenation structure, since axioms (2.29b),(2.29c) require concatenations not only to exist but also to be ordered in specific ways.)

The main results of this section, then are these. First, although the interpretation of concatenation as restricted join is not immediately compatible with standard RTM, it is possible to modify the axioms in such a way that this identification can be made. This allows us to draw systematic connections between RTM and the inherent structure of the domains of denotation of natural language expressions, which I will exploit repeatedly in the chapters on modality. (Note that this is not a claim about the meaning of concatenation in RTM generally, but just about the best interpretation of concatenation in a degree semantics for English built on RTM. The interpretation of concatenation as join is too strong for some purposes to which RTM has been put.)

Second, scales may interact in any of a number of ways with concatenation/join. The two most important for our purposes here are scales which are additive with respect to concatenation and those which are intermediate. The former is exemplified by properties such as length, height, weight, and volume, and the latter by danger and temperature. This distinction will be very important in coming chapters: I will argue that epistemic modals are associated with additive scales, so that the likelihood of ϕ *or* ψ is related additively to the individual likelihoods of ϕ and ψ when these are disjoint, while deontic modals and desire verbs are associated with intermediate scales, so that e.g. the desirability of ϕ *or* ψ is intermediate between the desirability of ϕ and the desirability of ψ.

Finally, we can add a characterization of intermediate interval scales which complements the definition of non-concatenative interval scales given in (2.17) above.

(2.33) An **intermediate interval scale** is a structure $\langle X, Y, \succcurlyeq_P, \circ\rangle$ where

a. $\langle X, Y, \succcurlyeq_P\rangle$ is an interval scale;

b. $\langle X, \succcurlyeq^i_P, \circ\rangle$ is a concatenation structure, where $x \succcurlyeq^i_P y =_{df} \{(x,y) \mid \exists z : (x,z) \succcurlyeq (y,z)\}$;

c. Whenever $x \succcurlyeq^i_P y$ and $x \circ y$ is defined, $x \succcurlyeq^i_P (x \circ y) \succcurlyeq^i_P y$.

(2.33) will ensure, for example, that the temperature of the concatenation of two bowls of soup *a* and *b* will be somewhere between the temperature of *a* and the temperature of *b*, inclusive.

2.2.3 Boundedness

The boundedness of scales has not traditionally been a topic of major interest in RTM, but, as we saw already in chapter 1, it is an issue of considerable import in the semantics of gradability (Rotstein & Winter 2004; Kennedy & McNally 2005; Kennedy 2007). To get a usable characterization of boundedness, we need to separate INHERENT boundedness from ACCIDENTAL boundedness. One prominent kind of accidental boundedness is the upper- and lower-boundedness exhibited by any property in a finite domain. In a finite model, the ratio scale $\mathcal{S}_{tall} = \langle X, \succcurlyeq, \circ \rangle$ will have an upper bound (the height of the tallest object in X) and a non-zero lower bound (the height of the shortest object in X). This is not enough to make *tall* behave as if it had a fully closed scale, judging by the standard tests:

(2.34) a. This glass is almost full.
b. # Sam is almost tall.

(2.35) a. This glass is half full.
b. # Sam is half tall.

It is not possible, for example, to interpret (2.35b) as meaning "Sam is half as tall as the tallest person around". Accidental boundedness is not very interesting for our purposes; the interpretation and acceptability of degree modifiers and the like does not appear to be responsive to properties that hold of domains simply due to contingent features of the model.

Inherent boundedness, on the other hand, is due to some structural feature of a domain. For example, a Boolean algebra is inherently upper- and lower-bounded, due to the presence of a bottom element $\bot$ and a top element $\top$ in its structural definition.

In some such cases it is necessary to include the bounds explicitly; in other cases boundedness properties are entailed by other characteristics of a scale. For instance, as noted above, ratio scales such as the one underlying *tall* are inherently lower-bounded. This is due to the Archimedean property, which entails that any object in the domain is n times as tall as any other object for some $n > 0$. If there were not a lower bound at zero, then this property would not hold. As a result, we do not need to include a bottom element $\bot_{tall}$ explicitly in the structure underlying *tall*; indeed, we should not, since nothing can have zero height. On the other hand, it seems clear that *tall* does not have an inherent top element, for reasons already discussed.

Although ratio scales are inherently lower-bounded, boundedness and scale type (in the RTM sense) are at least partially independent parameters of variation. For instance, some lower-bounded scales do not seem to be ratio scales, and some ratio scales are inherently upper-bounded. For an example of the first, consider again *dangerous*. As we saw in (21,23), danger is not additive; this is enough to exclude it as a candidate for denoting on a ratio scale. On the other hand, as discussed in some detail by Kennedy & McNally (2005); Kennedy (2007), *dangerous* passes the usual tests for the presence of a lower bound, as well as the absence of an upper bound.

(2.36) Presence of lower bound:
a. This neighborhood is slightly dangerous. (*slightly*-modification)
b. This neighborhood is completely/almost safe. (Upper-boundedness of antonym)

(2.37) Absence of upper bound:
a. # This neighborhood is half dangerous. (Proportional modification with degree meaning)
b. # This neighborhood is slightly safe. (Non-lower-boundedness of antonym)

A further difference between *tall* and *dangerous* is that *dangerous* has a non-empty zero: a neighborhood can have no degree of danger, but no physical object can have zero height. So we have at least two different ways to get a scale with a lower bound but no upper bound, which differ in whether any object in their domain can occupy their minimum.

(2.38) *Tall* is associated with a ratio scale $\langle X, \succcurlyeq_{height}, \circ \rangle$, where X is a set of objects of which height can sensibly be predicated (e.g., physical objects).

(2.39) *Dangerous* is associated with an intermediate interval scale $\langle X, Y, \succcurlyeq_{danger}, \bot_{danger} \rangle$, where
a. X is a set of objects of which danger can sensibly be predicated;
b. $\langle X, Y, \succcurlyeq_{danger} \rangle$ is an interval scale (cf. (2.17));
c. $\langle X, \succcurlyeq^{i}_{danger}, \circ \rangle$ is a concatenation structure (cf. (2.29));
d. Whenever $x \succcurlyeq^{i}_{P} y$ and $x \circ y$ is defined, $x \succcurlyeq^{i}_{P} (x \circ y) \succcurlyeq^{i}_{P} y$;
e. $\forall y \in X : y \succcurlyeq_{danger} \bot_{danger}$.

Tall and *dangerous* illustrate the partial independence of boundedness and scale type in the RTM sense. In general, there does not seem to be any barrier to defining upper- and/or lower-bounded interval scales, with or without a concatenation relation. It remains to be seen whether natural languages utilize all of the possibilities made available by these parameters. For example, if *dangerous* is associated with a lower-bounded interval scale as I suggested, its antonym *safe* is presumably associated with an upper-bounded interval scale. It is not clear whether there are fully closed interval scales; one possible candidate is measurement of similarity and difference, but I am not sure whether this is the correct characterization of these scales.[10]

Another useful example of the independence of boundedness and scale type is the existence of ratio scales which, unlike *tall*, have an inherent upper bound. The usual example (e.g., Krantz et al. 1971) is numerical probability, which is additive with respect to concatenation (disjoint union) but has an upper limit with probability 1. (We will come back to this point in some detail in the next chapter.) Other familiar examples include several of the fully closed scales of Kennedy & McNally (2005), e.g. the properties of **fullness** and **closure**. These properties are clearly additive: for example, if you fill a glass 20% of its total volume and then fill it 40% of its total volume, you have just filled it 60%; or if you close a door half of the arc from one side to the other and then close it halfway again, it will be completely closed. At the same time, they are standard examples of upper- and lower-bounded properties (Kennedy & McNally 2005):

(2.40) a. The door is completely/almost open/closed.
(Upper-boundedness of adjective and antonym)
b. The glass is completely/almost full/empty. (ibid.)

These facts suggest that *full* and *closed* are associated with ratio scales that have an inherent upper bound in addition to the lower bound. Interestingly, however, these scales differ from heights in that objects in the domain **can** have a minimal degree of the property in question. We can capture the properties of fully closed ratio scales as in (2.41):

10 There are further interactions between boundedness and scale type, of course. One such interaction which is especially interesting for natural language is the fact that, although a non-lower-bounded interval scale cannot be positive, a lower-bounded interval scale can; if it is, though, it is equivalent to a ratio scale. Since *dangerous* (interval) and *expensive* (ratio) both have bottom elements which can be occupied by some element in their domain, the only structural difference in their scales appears to be whether they are positive or intermediate with respect to concatenations.

(2.41) A **fully closed ratio scale** is a concatenation structure $\langle X, \succcurlyeq_P, \circ, \bot_P, \top_P \rangle$ where

a. $\forall x : \top_P \succcurlyeq_P x \succcurlyeq_P \bot_P$;

b. $\forall x \forall y$: if $x \circ y$ is defined and $y \neq \bot_P$, then $x \circ y \succ_P x$.

Fully closed ratio scales are additive, and so there will be a unique equivalence class of individuals $[y]_P$ such that $\mu(\top_P) - \mu(y) = \mu(y) - \mu(\bot_P)$. The exact number assigned here is not a stable quantity, but the ratio of this point to the measure of $\top_P$, $\frac{\mu(y)}{\mu(\top_P)}$, is fixed as usual for any two points on a ratio scale; thus statements such as *The glass is half full* are predicted to be interpretable.

This statement would not be interpretable if the scale were not additive, e.g. if it were a fully closed interval scale (with no concatenation relation or one that is intermediate). In this case, admissible μ would disagree on what elements occupied the halfway point between $\top_P$ and $\bot_P$. This means that, if *similar* is a fully closed interval scale as I speculated above, we expect the infelicity of *x and y are half similar* despite the acceptability of *x and y are completely similar* (similar in no relevant respect) and *x and y are completely different* (different in all relevant respects).

It is worth noting here (in preparation for the discussion of epistemic modals in the next chapter) that if we consider the class of homomorphisms from an additive structure $\langle X, \succcurlyeq_P, \circ, \bot_P, \top_P \rangle$ into $\langle \mathbb{R}, \geq, +, 0, n \rangle$ for some particular $n > 0$, the choice of μ is determined uniquely, and there are no non-trivial admissible transformations. Some authors consider this a different scale type from ratio scales as a result, an **absolute** scale (e.g. Ellis 1966). In the case of standard probability, for example, it is customary to take $n = 1$ here, and no reassignment of values is possible. However, the difference between ratio and absolute scales seems to be mostly a matter of definition, depending on whether we pick n in advance for the target structure $\langle \mathbb{R}, \geq, +, 0, n \rangle$ or consider the class of homomorphisms into all structures of the form $\langle \mathbb{R}, \geq, +, 0, n \rangle$ (with $n > 0$). For our purposes, we can consider fully closed ratio scales to be true ratio scales with extra structure imposed by the presence of the top element.

2.3 Summary and Conclusion

This chapter gave an overview of the Representational Theory of Measurement, which provides a method for constructing measure functions from ordered qualitative structures. As we will soon see, this is directly applicable to Kratzer's theory of modality, and illuminates the logical properties of this theory and its prospects as a basis for a semantics of gradable modals.

Measurement theory also invites us to consider an expanded range of possible scale types: in addition to the boundedness properties familiar from Kennedy & McNally (2005), we have a distinction between ordinal, interval, and ratio scales, and a distinction between scales which are additive with respect to concatenation and those which are intermediate. All of these distinctions will crop up repeatedly in the discussion of scalar modals.

In the next chapter we turn to the discussion of modality in earnest. Using tools developed in this chapter and new data involving degree modification and disjunction, I show that the scales underlying epistemic modals cannot be the ones that Kratzer's theory gives us, but must instead be one of the scale types that we have seen in this chapter: a fully closed ratio scale, like those associated with the adjective pairs *full/empty* and *open/closed*. In later chapters we will see that deontic modals and desire verbs are naturally associated with another type of scale discussed here: the intermediate interval scale.

CHAPTER 3

The Structure of Epistemic Modality

3.1 Chapter Overview

This chapter begins the discussion of modality with an examination of epistemic modals. §3.2 considers the standard theory of (epistemic) modality due to Kratzer (1981, 1991), both on its own terms and in light of measurement theory. As with the discussion of degree semantics in the previous chapter, recasting this theory using RTM does not change it in any essential way, but it does bring certain issues into sharp relief which have not been a major topic of investigation in previous work on modality.

Given the widespread acceptance of Kratzer's semantics for modality, there has been surprisingly little work examining the logical properties of the theory (some notable exceptions are Costa & Taysom 2005; Yalcin 2010; Swanson 2011, and, indirectly, Lewis 1981; Halpern 1997, 2003). In §§3.2-3.3 I show that, as a theory of epistemic modality, this semantics has severe logical problems of at least three kinds:

- The theory validates inferences involving epistemic comparatives and disjunction that are clearly invalid, notably:

  (3.1) a. ϕ is at least as likely as ψ.
  b. ϕ is at least as likely as χ.
  c. ∴ ϕ is at least as likely as $(\psi \vee \chi)$.

  This is connected with the fact that all modalities in Kratzer's theory are **maximal** with respect to concatenation/disjunction of comparable propositions, as defined in chapter 2. The problem is not alleviated by modifying the definition of comparative possibility (cf. §3.7).

- Kratzer's theory fails to provide interpretable truth-conditions for sentences where ratio and proportional modifiers are used with epistemic adjectives, such as ϕ *is twice as likely as* ψ and ϕ *is 90% certain*, but these sentence-types are widely used and clearly meaningful.

- The ordering of propositions induced by Kratzer's method using a modal base and an ordering source is too weak to support a theory of epistemic comparatives. In particular, it predicts that far too many pairs of propositions will be incomparable, and so many epistemic comparatives and equatives undefined, in models of even modest size.

These problems are due to built-in logical features of the theory, and bring into serious doubt whether the standard theory can provide a suitable logic for epistemic modals.

In §3.4 we will seek a semantics that avoids these problems, using the gradable epistemic modals *possible*, *probable*, *likely*, and *certain* to diagnose the structure of the scales associated with (at least) epistemic adjectives. Using a number of tests for boundedness properties borrowed from Rotstein & Winter (2004); Kennedy & McNally (2005); Kennedy (2007), I show that these adjectives are interpreted with respect to a scale which is both upper- and lower-bounded (i.e., fully closed) and positive with respect to concatenation. I also argue, contra Portner (2009), that there is no compelling reason from the theory of gradability to associate these adjectives with different scales, and in fact there is good reason not to; data from degree modification, entailments, and implicatures support a straightforward approach with these items picking out different points on the same scale.

§3.5 brings these observations together into a single proposal which builds on the typology of scales developed in chapter 2. Structures which are upper- and lower-bounded and positive with respect to concatenation correspond are **fully closed ratio scales**, which we saw in the last chapter when looking at *full/empty* and *open/closed*. Interestingly, it turns out that this class of structures is very close to those picked out by axiomatizations of qualitative probability due to Savage (1954); Krantz et al. (1971); Fine (1973) and many others. A result due to Narens (2007) shows that, if the scale associated with epistemic modals is fully closed and has certain structural properties which are standard in formal semantics, it is provably equivalent to a familiar type of numerical probability. Suprisingly, then, the degree modification facts seem to compel us to the conclusion that the right scale for adjectival epistemic modals is standard probability or something very closely related.

§3.7 discusses some modifications to the theory proposed by Kratzer (2012) and considers whether the empirical problems discussed in §§3.2-3.3 are alleviated; I show that the objections are not resolved by tweaking the definition of comparative possibility in the way that she suggests. §3.8 considers whether probability also underlies the semantics of the auxiliary modals *must*, *should*, *might*, and *ought* in their epistemic interpretation, or whether it is better to adopt a hybrid theory where the adjectives have a probabilistic scalar semantics but the interpretation of the auxiliaries is given by Kratzer's theory. I show that, on either version of comparative possibility, the hybrid theory makes demonstrably incorrect predictions about the logical relationship between the epistemic adjectives and auxiliaries. The problems are resolved, however, if the auxiliaries too have a probabilistic semantics; this indicates that they are scalar, even though they are not gradable. In the final sections I attend to some further details, including the interaction of the scalar theory with conditionals and an objection to the equation of possibility with non-zero probability due to Yalcin (2007) and others. I also suggest an information-theoretic semantics for question-embedding *certain* which explains the uniqueness entailments of this construction and accounts for an otherwise puzzling dissociation between certainty and probability.

Some of the material in this chapter appeared in Lassiter (2010a), in particular section §3.4. Like that paper, this chapter is indebted to Seth Yalcin, who first suggested the connection between epistemic modality and the theory of gradability developed here, and also argued that these considerations might favor the use of numerical probability. His conclusions with respect to these questions are generally consonant with my own, cf. the recently published Yalcin (2010).

3.2 Kratzer's Theory and Degree Semantics

3.2.1 Review of the Theory

Recall from §1.5.3 that Kratzer's influential semantics for modality gives truth-conditions to modal sentences using binary orders, in a way that is generally reminiscent of the measurement-theoretic approach to degree semantics discussed in chapter 2. In this section we will consider the predictions of Kratzer's theory more closely, focusing for the moment on epistemic modals.

In Kratzer's theory, modal sentences get their truth-conditions from the interaction of three factors:

- The MODAL BASE **f**, which is a function from worlds to sets of propositions;
- The ORDERING SOURCE **g**, which is also a function from worlds to sets of propositions;
- The lexical semantics of the modal in question, which determines what use it makes of **f** and **g**.

In many contexts, **f** and **g** are left as free parameters to be determined pragmatically, possibly subject to constraints from the lexical semantics of particular expressions. However, **g** is sometimes given explicitly by a phrase like "In view of what we know, ..." or "In view of what the law provides, ...". Different modal "flavors" are modeled by varying the particular choice of **f** and **g** in a context.

As in standard modal logic, the truth-conditions of sentences containing strong modals (*must, should, necessarily, obligatorily,* ...) and weak modals (*may, might, possibly, permissibly,* ...) are given in terms of universal and existential quantification respectively. The main difference, as discussed in §1.5.3, is that the set of worlds quantified over is determined in a more complex fashion by the interaction of **f** and **g**. Rather than being made true by universal quantification over an unstructured set of worlds given by context, a sentence like *Paul must be at home* is true if and only if it is true in all of the worlds which are maximal with respect to a binary relation $\succcurlyeq_{\mathbf{g}(w)}$, defined as

(3.2) $u \succcurlyeq_{\mathbf{g}(w)} v$ if and only if $\{p : p \in \mathbf{g}(w) \wedge u \in p\} \supseteq \{p : p \in \mathbf{g}(w) \wedge v \in p\}$

or equivalently:

(3.3) $\succcurlyeq_{\mathbf{g}(w)} =_{df} \{(u,v) \mid \forall p \in \mathbf{g}(w) : v \in p \rightarrow u \in p\}$

That is, $u \succcurlyeq_{\mathbf{g}(w)} v$ holds if and only if u satisfies every proposition in the ordering source that v does, and possibly more.

Recall from chapter 1 the fact that the relation $\succcurlyeq_{\mathbf{g}(w)}$ is not generally connected. In fact, as long as the propositions in $\mathbf{g}(w)$ are consistent and there are enough worlds to instantiate them all, $\succcurlyeq_{\mathbf{g}(w)}$ is a **pre-order**: transitive and reflexive, but not connected or antisymmetric. This is because $\succcurlyeq_{\mathbf{g}(w)}$ is defined in terms of the subset relation on propositions in $\mathbf{g}(w)$, which is also generally a pre-order. Consider an example similar to the one that we saw in chapter 1, with an ordering source appropriate for epistemic modals this time (a set of expectations, tailored for a conversation about a very predictable person named Bill).

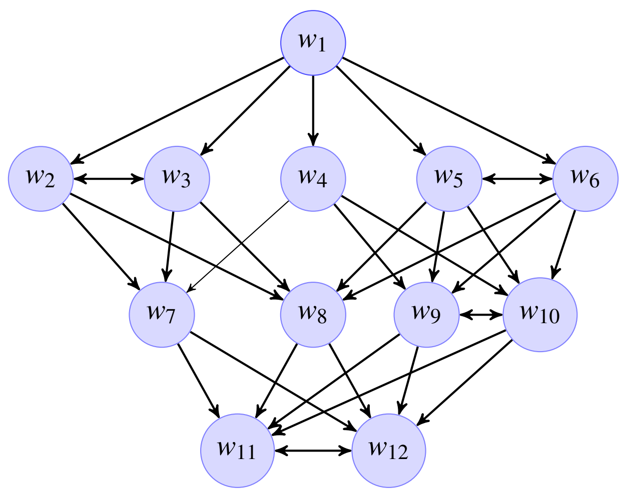
(3.4) Set of expectations $\mathbf{g}(w)$:
E1. Bill is at home by 6PM.

E2. Bill drives his car.
E3. Bill has macaroni for dinner.

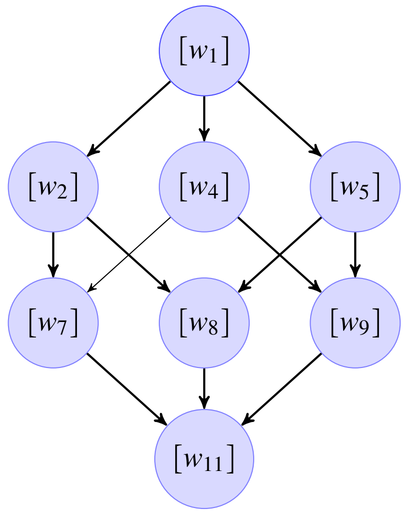
(3.5) w_1: E1-E3 are obeyed.
w_2, w_3: Only E1 violated.
w_4: Only E2 violated.
w_5, w_6: Only E3 violated.
w_7: Only E1 and E2 violated.
w_8: Only E1 and E3 violated.
w_9, w_{10}: Only E2 and E3 violated.
w_{11}, w_{12}: E1, E2, and E3 all violated.

The pre-order $\succcurlyeq_{\mathbf{g}(w)}$ is (as with pre-orders generally) associated with a reduction $\succcurlyeq^*_{\mathbf{g}(w)}$ to equivalence classes which is a partial order. In fact, if there are enough worlds in the modal base, $\succcurlyeq^*_{\mathbf{g}(w)}$ has the structure of a Boolean algebra:

Relation $\succcurlyeq_{\mathbf{g}(w)}$ *with* $\bigcap \mathbf{f}(w) = \{w_1, w_2, ..., w_{12}\}$:



Reduction $\succcurlyeq^*_{\mathbf{g}(w)}$ *to equivalence classes*:



If the modal base $\bigcap \mathbf{f}(w)$ does not happen to contain worlds instantiating all of these possibilities, then the ordering $\succcurlyeq_{\mathbf{g}(w)}$ might have more structure than this. $\succcurlyeq_{\mathbf{g}(w)}$ might even be connected in an extreme case, if by accident the modal base is very limited.[1] For instance, if in the above model $\bigcap \mathbf{f}(w)$ contained only w_1, w_2, w_3, w_7, and w_{10}, then $\succcurlyeq_{\mathbf{g}(w)}$ would be a weak order (but accidentally!), and $\succcurlyeq^*_{\mathbf{g}(w)}$ would be a simple order. In this case, every world would violate a super- or subset of the expectations that every other does does. There is not, however, any reason to think that realistic models will have this property; in general, $\succcurlyeq_{\mathbf{g}(w)}$ will have a weaker structure.

Must and *might* are now defined as functions which take a propositional argument ϕ and return True if and only if ϕ is true in every/some $\succcurlyeq_{\mathbf{g}(w)}$-maximal world. (As discussed in chapter 1, this definition is only appropriate if $\mathbf{g}(w)$ is finite; this is probably not a major limitation, and in any case it would not affect the main point here if we were to use Kratzer's more complicated definitions which work for infinite $\mathbf{g}(w)$.)

(3.6) $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) =_{df} \{v \mid v \in \bigcap \mathbf{f}(w) \wedge \neg \exists v' \in \bigcap \mathbf{f}(w) : v' \succ_{\mathbf{g}(w)} v\}$

(3.7) a. $[\![\textit{must } \phi]\!]^{\mathcal{M},w,g} = 1$ iff $\forall u : u \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) \rightarrow u \in \phi$.
b. $[\![\textit{might } \phi]\!]^{\mathcal{M},w,g} = 1$ iff $\exists u : u \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) \wedge u \in \phi$.

Remember that the worlds in $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w))$ are not necessarily comparable *to each other*; they are the worlds which are maximal in *some* branch of $\mathbf{g}(w)$, and there may well be multiple branches since $\mathbf{g}(w)$ is not a connected order. The worlds in $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w))$ will only be global maxima if the propositions in the ordering source are consistent and there are some worlds in the modal base which satisfy them all. It holds only in this special case that $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) = \bigcap \mathbf{f}(w) \cap \bigcap \mathbf{g}(w)$ ("the best of all accessible worlds", though not necessarily the best of all possible worlds).

1 Or all the propositions in the ordering source are nested, i.e. no two are logically independent. I'm ignoring bizarre models of this sort here.

The most important features of Kratzer's theory for our purposes, and the ones which will cause significant problems, relate to the treatment of gradability and comparison in the modal domain. Kratzer gives the truth-conditions of modal comparatives and several other operators — including *probable* — in terms of a binary relation on propositions $\succcurlyeq^s_{\mathbf{g}(w)}$ which is defined in terms of the binary relation $\succcurlyeq_{\mathbf{g}(w)}$ on worlds:[2]

(3.8) $\succcurlyeq^s_{\mathbf{g}(w)} =_{df} \{(\phi, \psi) \mid \forall u \in \psi \; \exists v \in \phi : v \succcurlyeq_{\mathbf{g}(w)} u\}$, where $u, v \in \bigcap \mathbf{f}(w)$.

So, for example, *It is at least as likely to rain as it is to snow* is true if and only if, for every world u in which it snows, there is some world v in which it rains such that v satisfies every proposition in $\mathbf{g}(w)$ that u does.

3.2.2 From Kratzer's Semantics to Measure Functions

As Yalcin (2007, 2010); Portner (2009) point out, and as we discussed in chapter 1, the fact that certain epistemic modals can form comparatives (3.9) and accept degree modification (3.10) suggests that we need a semantics for epistemic modals that is closely connected with our semantics for gradability and comparison.

(3.9) It is more likely that the Yankees will win than it is that the Blue Jays will.

(3.10) It is very probable/somewhat likely/almost certain/nearly impossible that the Yankees will win.

Obviously, the semantics described in the last section does not give us this: complex epistemic modals are not given a compositional interpretation, and there is no mention of degrees or anything which would play their role. Nevertheless, given that this theory is widely accepted and has been quite successful in dealing with various other issues, we might want to try to extend it to a degree semantics (cf. Portner 2009: 77-79).

Indeed, the fact that Kratzer's semantics is built around binary orders, like the various RTM-style scales that we have seen, makes it easy to extract a degree semantics from Kratzer's theory: we just define a class of scales $\mathcal{K}$ (for $\mathcal{K}$RATZER-STRUCTURES) which make use of the relations $\succcurlyeq_{\mathbf{g}(w)}$ and $\succcurlyeq^s_{\mathbf{g}(w)}$ just discussed.

(3.11) A Kratzer-Structure $\mathcal{K}$ is a tuple $\langle w, W, \Phi, \mathbf{f}(w), \mathbf{g}(w), \succcurlyeq_{\mathbf{g}(w)}, \succcurlyeq^s_{\mathbf{g}(w)} \rangle$, where

a. w is the world of evaluation;
b. W is a set of possible worlds;
c. $\Phi \subseteq \mathcal{P}(W)$ is an algebra of propositions (closed under union and complement);
d. $\mathbf{f}(w)$ and $\mathbf{g}(w)$ are sets of propositions (the modal base and ordering source);
e. $\succcurlyeq_{\mathbf{g}(w)}$ is the pre-order on W defined in (3.2),(3.3);

2 As I mentioned in chapter 1, $\succcurlyeq^s_{\mathbf{g}(w)}$ is notation that I've borrowed from Halpern (1997), who presents a logic extremely similar to Kratzer's but developed independently (both were inspired by David Lewis' work on counterfactuals and comparative possibility, cf. Lewis 1973, 1981). Kratzer doesn't give this relation a name, but it plays an important role in her theory and will be used frequently enough that it is useful to have an abbreviation.

f. $\succcurlyeq^s_{\mathbf{g}(w)}$ is the pre-order defined in (3.8), with domain Φ.

As in the RTM approach to degree semantics in chapter 2, we can identify the set of modal degrees with the set of equivalence classes of propositions in Φ under the $\succcurlyeq^s_{\mathbf{g}(w)}$ relation. Since the latter is a pre-order, the set of degrees will be a partial order, as we saw in the last subsection. The partiality of $\succcurlyeq_{\mathbf{g}(w)}$ is an important feature of Kratzer's semantics, and leads to a prediction of partiality in the order on propositions as well.

There are several further questions that the discussion of RTM in the last chapter leads us to ask about the structure of this scale. First, assuming as I did in (3.11c) that the set of propositions Φ is closed under union, we can ask how the order $\succcurlyeq^s_{\mathbf{g}(w)}$ interacts with disjunction. As I argued in the last chapter, for the purposes of natural language semantics concatenation can be identified with the join operation, which is union/disjunction in this domain. It turns out that there is no simple way to define a concatenation operation that captures the effects of disjunction in Kratzer's semantics, due to the fact that $\succcurlyeq^s_{\mathbf{g}(w)}$ is not connected; the typology that we developed in chapter 2, (3.32) assumed that scales were weak orders. As it turns out, though, the interaction of $\succcurlyeq^s_{\mathbf{g}(w)}$ with disjunction is quite complicated and produces some counter-intuitive results; some of these will be discussed in §3.3.1 in particular.

Second, we would like to know whether $\succcurlyeq^s_{\mathbf{g}(w)}$ is inherently upper- and/or lower-bounded. The answer to this question is slightly complicated. As long as Φ contains enough propositions and the propositions in $\mathbf{g}(w)$ are consistent, the binary order $\succcurlyeq^s_{\mathbf{g}(w)}$ is upper-bounded by $\bigcap \mathbf{g}(w)$ — the set of worlds which fulfill all of the expectations in $\mathbf{g}(w)$ — and lower-bounded by $\varnothing$, which is guaranteed to fulfill none of them.[3] In models in which $\mathbf{g}(w)$ is not consistent, however, $\succcurlyeq^s_{\mathbf{g}(w)}$ will not have a unique upper bound. Instead, it will have as many upper bounds as there are $\succcurlyeq^s_{\mathbf{g}(w)}$-equivalence classes in the set of $\succcurlyeq^s_{\mathbf{g}(w)}$-undominated propositions.

Because $\succcurlyeq^s_{\mathbf{g}(w)}$ is only a pre-order, we cannot identify it with any of the standard RTM scale types: ratio, interval, and ordinal scales require at least a weak order. This is not a problem per se, since it may be that some gradable adjectives make use of partially ordered scales anyway; standard examples are *clever* and *big*. What does follow from this fact is just that when we define admissible measure functions for $\mathcal{K}$-structures, a larger class of transformations will be admissible than in the case even of ordinal scales (which, remember, allow all monotone increasing transformations).

(3.12) A measure function $\mu_{\mathcal{K}}$ is **admissible** for a Kratzer-structure $\mathcal{K}$ iff, for all $\phi, \psi \in \Phi$: $\phi \succcurlyeq^s_{\mathbf{g}(w)} \psi$ implies $\mu_{\mathcal{K}}(\phi) \geq \mu_{\mathcal{K}}(\psi)$.

(3.13) **Fact**: If ϕ and ψ are incomparable — i.e. neither $\phi \succcurlyeq^s_{\mathbf{g}(w)} \psi$ nor $\psi \succcurlyeq^s_{\mathbf{g}(w)} \phi$ — then there are admissible $\mu_{\mathcal{K}}, \mu'_{\mathcal{K}}$ such that $\mu_{\mathcal{K}}(\phi) > \mu_{\mathcal{K}}(\psi)$ and $\mu'_{\mathcal{K}}(\phi) < \mu'_{\mathcal{K}}(\psi)$.

(The proof is straightforward and is omitted here.)

3 "Enough" means that Φ contains at least one proposition containing worlds where all of the propositions in $\mathbf{g}(w)$ are true — i.e., some ϕ such that $\phi \approx_{\mathbf{g}(w)} \bigcap \mathbf{g}(w)$; and at least one proposition where none of the propositions in $\mathbf{g}(w)$ are true — some ψ such that $\psi \approx_{\mathbf{g}(w)} \varnothing$.

(3.14) **Corollary of (3.12) and (3.13)**: If $\mu_{\mathcal{K}}$ is a $\mathcal{K}$-admissible measure function, then $f(\mu_{\mathcal{K}})$ is admissible as well for all monotone increasing (order-preserving) f, and many non-order-preserving f as well.

The fact that $f(\mu_{\mathcal{K}}(x))$ is admissible for all monotone increasing f follows from the proof given for ordinal scales in Krantz et al. (1971); $f(\mu_{\mathcal{K}})$ is also admissible for non-increasing f, as long as it does not reverse the order of any connected parts of $\succcurlyeq_{\mathbf{g}(w)}$ in $\mathcal{K}$.

Note, however, that many modal expressions in Kratzer's semantics — in particular *must* and *might* — are not given meanings in terms of the binary relation on propositions $\succcurlyeq^{s}_{\mathbf{g}(w)}$, but in terms of first-order quantification over the members of a set whose identity is determined in a different way by the binary relation $\succcurlyeq_{\mathbf{g}(w)}$ on worlds. This means that, even though we can define a reasonable notion of degree using Kratzer's theory, degrees so defined do not play any role in the semantics of sentences with *must, might, ought*, etc.

3.3 Logical and Empirical Problems with Kratzer's Theory

The three problems for Kratzer's theory that I will discuss here all involve the interaction of the comparative possibility relation with disjunction. If disjunction translates as the join operation, and concatenation is restricted join as I argued in chapter 2, another way to treat this issue is to ask how a concatenation operation would behave if it were added to a Kratzer-structure $\mathcal{K}$. It turns out that the interaction of $\succcurlyeq^{s}_{\mathbf{g}(w)}$ with concatenation does not generally display any simple pattern, except in one particularly interesting case: $\mathcal{K}$-structures are **maximal** with respect to concatenations when the propositions concatenated are $\succcurlyeq^{s}_{\mathbf{g}(w)}$-comparable. This property will be the source of a number of problems, and the target of revision when I come to my counter-proposal later. I'll argue that concatenation should be positive with epistemic modals, no matter what.

3.3.1 Problem 1: Epistemic Comparatives and Equatives with Disjunction

The first problem is also the most clearly problematic: Kratzer's semantics predicts the validity of a class of inferences involving equatives and disjunction which are clearly invalid.

(3.15) **The Disjunctive Inference**

a. ϕ is at least as likely as ψ.

b. ϕ is at least as likely as χ.

c. ∴ ϕ is at least as likely as $(\psi \vee \chi)$.

Proof of (3.15). In Kratzer's theory (3.15a) is true iff, for every ψ-world v, there is a ϕ-world u such that $u \succcurlyeq_{\mathbf{g}(w)} v$. Likewise, (3.15b) is true iff, for every χ-world v' there is a ϕ-world u' such that $u' \succcurlyeq_{\mathbf{g}(w)} v'$. Let z be an arbitrary world in $\psi \vee \chi$. Case 1: $z \in \psi$. Then there is a ϕ-world z' such that $z' \succcurlyeq_{\mathbf{g}(w)} z$, namely u. Case 2: $z \in \chi$. Then there is a ϕ-world z'' such that $z'' \succcurlyeq_{\mathbf{g}(w)} z$, namely u'. Since z was arbitrary, we conclude that for every $z \in (\psi \vee \chi)$, there is a $z' \in \phi$ such that $z' \succcurlyeq_{\mathbf{g}(w)} z$; thus, by the definition of $\succcurlyeq^{s}_{\mathbf{g}(w)}$, (3.15c) holds.

In fact, the validity of the inference in (3.15) was built into the axiomatization of comparative possibility by Kratzer's predecessor Lewis (1973: 52ff) and by Halpern (1997), whose Lewis-inspired discussion of comparative possibility is extremely close, perhaps equivalent, to Kratzer's theory. However, the damaging consequences of (3.15) for Kratzer's theory of modality were apparently not recognized in the literature until they were independently noticed by Yalcin (2010) and by Lassiter (2010a). (As Halpern (1997, 2003) notes, this property is shared by several other representations of uncertainty, e.g. possibility logic and fuzzy logic. The problem noted here is shared by these frameworks, as far their usefulness for modeling natural language is concerned.)

The reason that (3.15) is a problem is simply that this inference schema is clearly invalid as applied to epistemic comparatives and equatives in English. This is particularly clear when it is applied repeatedly, which can lead to extremely counter-intuitive results. For example, suppose someone tells you:

(3.16) "For any baseball team you like, the Blue Jays are at least as likely to win the World Series this year as that team is."

This is a reasonable claim that someone could make, if it happens that the Blue Jays are very good this year, and clearly better than any other team in Major League Baseball.

Now, (3.16) is obviously a weaker claim than (3.17):

(3.17) "The Blue Jays are at least as likely to win the World Series as they are not to win."

In a league with 30 teams, the Blue Jays would have to be really stellar for (3.17) to be true, simply because they will have so many opportunities to lose against the odds. Nevertheless, on Kratzer's theory, (3.16) — in combination with a few simple facts about baseball — actually **entails** (3.17). This is clearly unacceptable: first, the weaker statement should not entail the stronger; and second, (3.18) just isn't a self-contradictory set of claims.

(3.18) The Blue Jays stand out as the best team in the Major Leagues this year. For any team you like, the Blue Jays are at least as likely to win the World Series this year as that team is. But they are still more likely not to win than they are to win — they're not *that* good.

To see why (3.16) entails (3.17) in Kratzer's theory, let $\mathbf{T} = \{\mathbf{team}_1, ..., \mathbf{team}_{29}\}$ be the other 29 Major League Baseball teams. (3.19) is a reasonable rendition of the truth-conditions of (3.16):

(3.19) $\forall x \in \mathbf{T}$: It is at least as likely that the Blue Jays win as it is that x wins.

Let p be the proposition *The Blue Jays win*, and let q_n be the proposition *Team$_n$ wins*. (3.19) is equivalent to (3.20):

(3.20) $(p \succsim^s_{\mathbf{g}(w)} q_1) \wedge (p \succsim^s_{\mathbf{g}(w)} q_2) \wedge ... \wedge (p \succsim^s_{\mathbf{g}(w)} q_{29})$

Feeding (3.20) into schema (3.15) repeatedly, we see that (3.21) follows as well.

(3.21) $p \succsim^s_{\mathbf{g}(w)} (q_1 \vee q_2 \vee ... \vee q_{29})$

Since one of the thirty teams must win, the only way that the Blue Jays can fail to win is if someone else does — that is, *The Blue Jays do not win* is true iff $(q_1 \vee q_2 \vee ... \vee q_{29})$ is true. So we can rewrite (3.21) as (3.22):

(3.22) *The Blue Jays win* is at least as likely as *The Blue Jays do not win*.

Putting this all together: according to Kratzer, (3.16) and the rules of baseball entail (3.22), which is equivalent to (3.17). But (3.17) is clearly a much stronger claim than (3.16), not an entailment.

In general, in Kratzer's theory ϕ will be ranked as high as a disjunction ψ, no matter how large, if it is ranked as high as every disjunct in ψ. The only way that ψ can overtake ϕ is if one of the disjuncts itself outranks ϕ or, in some cases, is incomparable to ϕ. Is this a plausible prediction about valid inferences from sentences involving *likely*? The answer seems to be a clear "no". Somehow, a disjunction of lower-ranked possibilities must be able to "gang up" to overpower a higher-ranked possibility.

A further example may bring home just how damaging this feature of Kratzer's semantics is. Yalcin (2010) points out a special case of the Disjunctive Inference which is even more clearly absurd than the baseball example:

(3.23) a. $\phi \succcurlyeq^{s}_{\mathbf{g}(w)} \neg\phi$

b. $\therefore\ \phi \succcurlyeq^{s}_{\mathbf{g}(w)} (\phi \vee \neg\phi)$

This follows because $\succcurlyeq^{s}_{\mathbf{g}(w)}$ is reflexive, and so we apply (3.15) substituting $\neg\phi$ for ψ and ϕ for χ. But since $\phi \vee \neg\phi$ holds at every world, this means that $\phi \succcurlyeq^{s}_{\mathbf{g}(w)} W$.

If Kratzer's theory were right on this point, then the following pattern of inference would be intuitively correct:

(3.24) a. It is as likely that it will rain as it is that it will not rain.

b. So, it is as likely that it will rain as it is that $2+2=4$.

This is, as Yalcin puts it, an "egregious" result: the prediction is that, if ϕ is at least as likely as not-ϕ, then it is as likely as any tautology or necessary truth.

(The disjunctive inference is also predicted to be valid in a slightly restricted but similarly damaging form by the modified definition of comparative possibility proposed by Kratzer (2012); see §3.7 below for discussion.)

3.3.2 Problem 2: Degree Modification and Interpretability

A different but related objection to Kratzer's semantics becomes apparent when we examine the theory's predictions about degree modification in light of the discussion in the last chapter. The following expression-types are all sensible in English, and it is not hard to find natural examples of their use — we will see a number of such examples in §3.3.4 (cf. Portner 2009; Yalcin 2010).

(3.25) a. ϕ is twice as likely as ψ.

b. It is half certain that ϕ.

c. It is 95% certain that ϕ.

These sentences ought to come out as interpretable in the semantics. Thinking in terms of RTM, this means that they ought to have a truth-value which is stable across all admissible $\mu_{\mathcal{K}}$ relative to a Kratzer-structure $\mathcal{K}$. Furthermore, we know what truth-conditions these sentences ought to receive according to a standard semantics of degree: their interpretation should be the same as the corresponding sentences in (3.26), i.e. (3.27).

(3.26) a. Glass A is twice as full as glass B.
b. Glass C is half full.
c. Glass D is 95% full.

(3.27) a. $[\![(3.25a)]\!]^{\mathcal{M},w,g} = 1$ iff $\mu_{likely}(\phi) \geq 2 \times \mu_{likely}(\psi)$.
b. $[\![(3.25b)]\!]^{\mathcal{M},w,g} = 1$ iff $\mu_{certain}(\phi) \geq \frac{1}{2} \times \mu_{certain}(\top_{certain})$, where $\top_{certain}$ is the maximum element of the scale that *certain* is associated with.
c. $[\![(3.25c)]\!]^{\mathcal{M},w,g} = 1$ iff $\mu_{certain}(\phi) \geq 0.95 \times \mu_{certain}(\top_{certain})$.

In order to be interpretable these statements must, of course, be true not just in some particular μ_{likely} or $\mu_{certain}$ but in all measure functions that are admissible for the relevant scale.

Suppose that we take $\mathcal{K}$-structures as providing the scale for *likely* and *certain*. On this account, the qualitative order on propositions, for both *likely* and *certain*, is given by $\succsim^{s}_{\mathbf{g}(w)}$, and interpretable statements are those whose truth-value remains constant on all $\mathcal{K}$-admissible $\mu_{\mathcal{K}}$. It's not too hard to see that none of the statements in (3.25) are semantically interpretable on such a theory. As noted in (3.12) and (3.14), $\mathcal{K}$-structures are weaker than ordinal scales, and — as with ordinal scales — all monotone increasing transformations of admissible $\mu_{\mathcal{K}}$ are admissible. It follows that ratios and proportions are not held constant across admissible $\mu_{\mathcal{K}}$ for the same reason that they are not for ordinal scales (see ch.2, §2.1.2.1). The situation is of course even worse for $\mathcal{K}$-structures than it is for ordinal scales, since some non-increasing transformations are admissible as well (3.14). But even without this taking this feature of the theory into account, the prediction is clear: the sentences in (3.25) should be just as bizarre as (3.28) (which, recall from ch.2 §2.1.2.3, are associated with interval scales and so do not preserve ratios or proportions across admissible μ).

(3.28) a. # 9PM is twice as late as 4:30PM.
b. # Atlanta is 95% hot.

A bit more bluntly: *If Kratzer's comparative possibility relation determines the scale with respect to which epistemic adjectives are interpreted, none of the sentences in (3.25) should be intelligible.* Since this prediction is plainly false, we can conclude that building a degree semantics on top of Kratzer's theory in the way that we have been considering will not provide us with an empirically adequate account of epistemic adjectives and their interaction with degree modifiers.

3.3.3 Problem 3: Too Many Incomparabilities

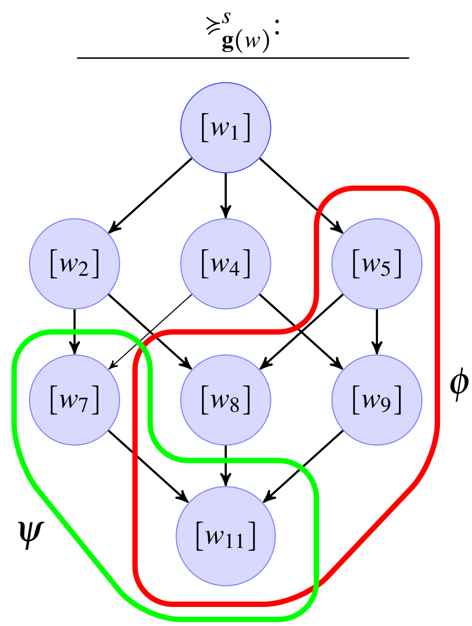
As we have seen, Kratzer's $\succsim_{\mathbf{g}(w)}$ and $\succsim^{s}_{\mathbf{g}(w)}$ relations are pre-orders, reflexive and transitive but not connected. Because of this, her theory predicts that a large number of epistemic comparatives and equatives will be undefined. Although, arguably, epistemic comparatives are sometimes undefined — and my own proposal will make room for this possibility — Kratzer's theory goes overboard, ruling many comparatives undefined which are intuitively quite reasonable. The problem is fundamentally that her theory makes no room for one expectation or norms being stronger than another; instead any conflict of expectation leads to incomparability.

For a simple example, consider the little model I gave above, with just three propositions in $\mathbf{g}(w)$. The definition of $\phi \succsim^s_{\mathbf{g}(w)} \psi$, "$\phi$ is at least as likely as ψ", requires that **every** ψ-world be weakly dominated by some ϕ-world: $\forall u \in \psi\ \exists v \in \phi : v \succsim_{\mathbf{g}(w)} u$. As a result, if there is even one ψ-world which is not comparable with any ϕ-worlds, then the comparative and equative will be undefined, even if all other ψ-worlds are strictly dominated by ϕ-worlds in every other case.

It is actually very easy for this condition to be fulfilled even in small models, and it gets easier as the models get richer and more realistic. Recall the small model above with three propositions in $\mathbf{g}(w)$:

(3.29) Set of expectations $\mathbf{g}(w)$:
E1. Bill is at home by 6PM.
E2. Bill drives his car.
E3. Bill has macaroni for dinner.

(3.30) w_1: E1-E3 are obeyed.
w_2, w_3: Only E1 violated.
w_4: Only E2 violated.
w_5, w_6: Only E3 violated.
w_7: Only E1 and E2 violated.
w_8: Only E1 and E3 violated.
w_9, w_{10}: Only E2 and E3 violated.
w_{11}, w_{12}: E1, E2, and E3 all violated.



Let $\phi = \{w_5, w_6, w_8, w_9, w_{10}, w_{11}\}$ and $\psi = \{w_7, w_{11}, w_{12}\}$. (In this model, ϕ picks out all of the worlds in which Bill does not have macaroni for dinner, and ψ picks out the worlds in which he does not arrive home by 6PM and does not drive his car — say, in these worlds he takes the bus instead, which slows him down.)

ϕ contains a number of worlds with equal or better plausibility to any of the worlds in ψ, at least in terms of how many expectations they fulfill; in particular, ϕ contains w_5 which only violates one expectation, while ψ contains only one world, w_7, which fulfills **any** of the expectations in $\mathbf{g}(w)$. Nevertheless, the logic of Kratzer's theory leaves us no choice but to declare ϕ and ψ incomparable in this case. This is because there is a ϕ-world — w_7 — for which there is no $u \in \psi$ such that $u \succsim w_7$. Also, there are several ψ-worlds v for which there is no ϕ-world v' such that $v' \succsim v$. As a result, neither $\phi \succsim^s_{\mathbf{g}(w)} \psi$ nor $\psi \succsim^s_{\mathbf{g}(w)} \phi$.

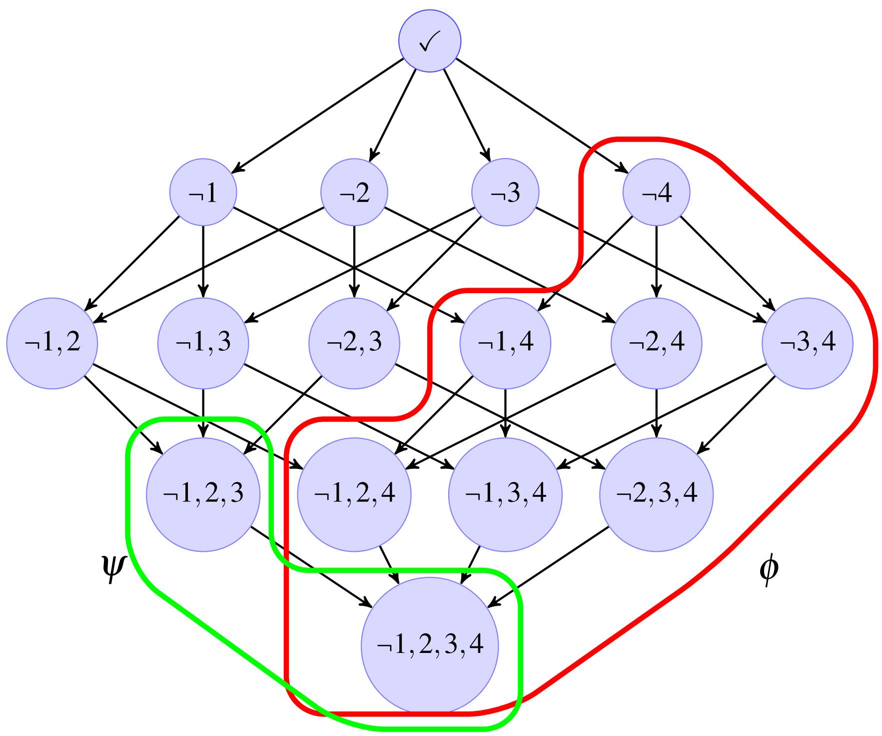
A comparison between these propositions should be incapable of taking a truth-value in a situation like this, then. This conflicts sharply with intuition:

(3.31) Bill is extremely predictable, and he almost always drives to and from work, arrives home by 6PM, and has macaroni for dinner. But it is more likely that Bill will have something other than macaroni for dinner than it is that he will both fail to be home by 6PM **and** fail to drive his car.

Surely it should not be impossible for the second sentence of (3.31) to be true (or false) in a situation like this. But Kratzer's semantics does not seem to leave us any choice.

This is the situation already in very small models. It turns out that, as $\mathbf{g}(w)$ gets larger, the size of $\succcurlyeq_{\mathbf{g}(w)}$ and $\succcurlyeq^{s}_{\mathbf{g}(w)}$ will increase exponentially, as will the number of incomparabilities. Here is what it looks like when we add just one more logically independent proposition to $\mathbf{g}(w)$:

Reduction $\succcurlyeq^{*}_{\mathbf{g}(w)}$ *to equivalence classes when* $\mathbf{g}(w) = \{1,2,3,4\}$:



Again, in this model ϕ and ψ are incomparable, as there are ϕ-worlds u for which no ψ-world violates a subset of the propositions in $\mathbf{g}(w)$ that u does. This is the case even though all ψ-worlds violate almost all expectations, and there may well be many ϕ-worlds fulfill almost all expectations. The situation rapidly grows more extreme as the models grow; with 5 or more propositions in $\mathbf{g}(w)$, approximately 50% of non-trivial comparisons between worlds are undefined.

These are very strong predictions made by Kratzer's theory, then: in any model of even moderate size, about half of comparisons between worlds will be undefined; and second, it doesn't matter *at all* for the comparison between ϕ and ψ whether a proposition satisfies almost all expectations or almost none of them, unless the set of expectations satisfied by ϕ happens to be a super- or subset of the set satisfied by ψ. Since even one systematic incomparability is enough to falsify a comparison between propositions, we expect that a huge number of epistemic comparatives will be without truth-value. The intuitive nature of the problems here is actually very simple: the theory doesn't leave any room for expectations being stronger or weaker than another. Presumably (3.31) is true to the extent that our expectations about Bill's travel habits are firmer than our expectations about his dinner plans. Kratzer's theory cannot capture this, however — expectations are all-or-nothing.

3.4 Adjectival Epistemic Modals and Boundedness

As we have seen, Kratzer's theory encounters some serious empirical problems which we want to avoid. In the rest of the chapter we will do so by developing a semantics for epistemic modals from the ground up, starting with basic questions about the structure of the scales that underlie epistemic modality. This section begins the project by taking a closer look at the adjectival epistemic modals *possible*, *probable*, *likely*, and *certain*. Tests for adjective type and boundedness taken from the literature allow us to use these expressions to diagnose structural properties of the scale(s) that epistemic adjectives are associated with, which will constrain other structural and logical properties that we can attribute to epistemic modals considerably.

3.4.1 Tests for Boundedness and Adjective Type

1. *Completely*-modification. *Completely* is a polysemous modifier, but one common function is as a maximizer. Kennedy & McNally (2005) argue that, if an adjective can be modified by *completely* with a "maximum" interpretation, it is a maximum-standard adjective, and thus denotes a function whose range is an upper closed scale. This explains the contrast in (3.32), where *bent* and *tall* are unacceptable with degree-modifying *completely* because their scales do not have a maximum element.

(3.32) a. The room is completely full.
b. # This basketball player is completely tall.
c. # The rod is completely bent.

2. *Slightly*-modification. According to Kennedy & McNally (2005), *x is slightly A* is true just in case *x* has property *A* to a degree which differs by a small amount from the minimum degree in $\mathcal{S}_A$. Thus if an adjective can be modified by *slightly*, it is associated with a scale with a bottom element, and, in general, is a minimum-standard adjective.

(3.33) a. The rod is slightly bent.
b. # This player is slightly tall.
c. # The room is slightly full.

3. "A but could be A-er": Minimum-standard and relative-standard adjectives are natural in the construction in (3.34), but maximum-standard adjectives do not (Kennedy 2007).

(3.34) a. The rod is bent, but it could be more bent.
b. This basketball player is tall, but he could be taller.
c. # The room is full, but it could be fuller.

4. Negation entails antonym? The negation of a minimum- or maximum-standard adjective entails that its antonym holds. Relative-standard adjectives have a "zone of indifference" (Sapir 1944; Kennedy 2007), so that it is possible to be neither *A* nor *not-A*.

(3.35) a. The rod is not bent. ⊨ The rod is straight.
b. The rod is not straight. ⊨ The rod is bent.
c. This player is not tall. ⊭ This player is short.

5. Type of antonym. The defining characteristic of antonymous pairs of adjectives is that *x is more* A_{pos} *than y* is true iff *y is more* A_{neg} *than x* is true. Since antonymy involves reversing the direction of the ordering relation, the maximum element of a scale, if there is one, is always the minimum element of its antonym, and vice versa (Rotstein & Winter 2004; Kennedy 2007). As a result, maximum-standard adjectives have minimum-standard antonyms and vice versa:

(3.36) a. This neighborhood is completely/#slightly safe.
b. This neighborhood is slightly/#completely dangerous.

(3.37) a. The rod is slightly/#completely bent.
b. The rod is completely/#slightly straight.

However, the antonym of a relative adjective is also a relative adjective.

(3.38) a. My car was #completely/#slightly expensive.
b. My car was #completely/#slightly cheap.

6. *Almost*. If an object is *almost A*, then it fails to be *A* by a small margin. Rotstein & Winter (2004) show that *almost A* is acceptable only with with maximum-standard adjectives *A*, and entails that the antonym of *A* with *slightly* holds of *x*:

(3.39) a. This area is almost safe. ⊨ This area is slightly dangerous.
b. #This player is almost tall.
c. #This rod is almost bent.

7. Proportional Modification. Proportional modifiers like *half*, *70%*, and *mostly* measure an object's location on the scale relative to both the maximum and minimum points. As a result, these modifiers occur only with adjectives such as *full* which are associated with a fully closed scale, since they are undefined if either of these points does not exist.

(3.40) a. The glass is half full/empty.
b. # My cousin is half tall/short.
c. # This road is half dangerous/safe.

In most cases proportional modifiers occur with maximum-standard adjectives. However, I will show below that they also occur in some cases with minimum- and relative-standard adjectives which are located on fully closed scales.

3.4.1.1 *Possible*

On all but one of our tests, *possible* behaves as a minimum-standard adjective. In order from (3.32) to (3.40):

(3.41) a. #It is completely possible that the Jets will win.
b. In a tense situation, it's slightly possible that an asteroid entering our atmosphere could trigger a nuclear war.[4]
c. It is possible that the Jets will win, but it could be more possible.[5]
d. It is not possible that the Jets will win. $\models$ It is impossible that they will.
e. It is completely/#slightly impossible that the Jets will win.
f. #It is almost possible that the Jets will win.
g. #It is half possible that the Jets will win.

However, *possible* differs from the minimum-standard adjectives *bent* and *dangerous* in accepting the proportional modifier *n%*:

(3.42) I felt that if it was 80-90 percent possible that [the cancer] hadn't spread, I didn't want the hysterectomy.[6]

These examples show that *possible* is a minimum-standard adjective. So we expect, based on the discussion of the minimum-standard adjectives *bent* and *dangerous* in chapters 1 and 2, the following truth-conditions for the positive form:

(3.43) $[\![\phi \textit{ is } \mathbf{pos} \textit{ possible}]\!]^{\mathcal{M},w,g} = 1$ iff $\phi \not\approx_{possible} \bot_{possible}$

(Recall that $\bot_P$ and $\top_P$ represent the inherent minimum and maximum, respectively, of the scale $\mathcal{S}_P$.) Equivalently,

(3.44) $[\![\phi \textit{ is } \mathbf{pos} \textit{ possible}]\!]^{\mathcal{M},w,g} = 1$ iff $\mu_{possible}(\phi) > \mu_{possible}(\bot_{possible})$.

And, of course, this means that the scale of possibility must have an inherent minimum.

Bracketing proportional modification (3.42) for the moment, the minimum amount of structure that we can attribute to $\mathcal{S}_{possible}$ is the following:

(3.45) $\mathcal{S}_{possible} = \langle \Phi, \succcurlyeq_{possible}, \ldots, \bot_{possible}, \ldots \rangle$, where
a. Φ is a set of propositions;
b. $\succcurlyeq_{possible}$ is a pre-order or weak order on Φ;
c. $\forall \phi \in \Phi : \phi \succcurlyeq_{possible} \bot_{possible}$.

(Ellipses leave room for further structure that we want to go back and fill in later.)

4 http://www.thefuturewatch.com/Catastrophe.html
5 Many speakers accept the comparative *more possible*, though some express discomfort, preferring *more likely* even for small values. I do not know the source of this preference, but it does not seem to be grammatical in nature: *more possible* is robustly attested in corpora. As we will see below with the proportional modification data, there seem to be soft preferences that are not grammatical in nature for using different adjectives which could express intermediate grades of possibility.
6 http://www.suggestadoctor.com/doctor_118_paulshuchiang_lin.htm

3.4.1.2 *Probable* and *Likely*

Probable and *likely* appear to be synonymous, although there are syntactic and register differences between them (cf. Horn 1989 and chapter 4 below). Almost without exception, they behave as relative-standard adjectives:

(3.46) a. #It is completely likely/probable that the Jets will win this year.[7]
b. #It is slightly likely/probable that the Jets will win.
c. It is likely that the Jets will win, but it could be more likely.
d. It is not likely that the Jets will win, but it is not unlikely either.
e. It is #completely/#slightly unlikely/improbable that the Jets will win.
f. #It is almost likely/probable that the Jets will win.
g. #It is half likely/probable that the Jets will win.

The exception is that, in contrast with the relative adjective *tall*, *likely* and *probable* occur with the proportional modifier *n%* quite often in corpora. Here are some examples:

(3.47) a. [I]t's 80 percent likely that the iPhone will be coming to T-Mobile ... [8]
b. [T]he IPCC ... said it was "very likely" or more than 90 percent probable that human activities ... had caused most of the warming in the past half century.[9]

I will account for this difference between *likely* and *tall* in §3.4.3.1 below.

Another way in which *likely* and *probable* resemble *tall* is in taking ratio modifiers:

(3.48) a. Sam is twice as tall as Harry.
b. It is twice as likely to rain as it is to snow.

As we discussed in chapter 2, ratio modifiers are only interpretable if the scale is a ratio scale; so, like $\mathcal{S}_{tall}$, $\mathcal{S}_{likely}$ is presumably a ratio scale. If this is right, then $\mathcal{S}_{likely}$ must have a concatenation operation which is positive and regular, as well as an inherent minimum. $\succcurlyeq_{likely}$ must also be a weak order in order for a ratio scale to be defined..

(3.49) $\mathcal{S}_{likely}$ is a ratio scale $\langle \Phi, \succcurlyeq_{likely}, \circ, \ldots \rangle$, where
a. Φ is a set of propositions;
b. $\succcurlyeq_{likely}$ is a weak order on Φ;
c. $\succcurlyeq_{likely}$ is positive, regular, and Archimedean.

Note that, as with *tall* in chapter 2, $\mathcal{S}_{likely}$'s inherent minimum need not be enumerated explicitly since its existence is entailed by the ratio scale axioms.

The behavior of *likely* and *probable* on these tests also suggests the following truth-conditions for the positive form, again following the treatment of relative adjectives like *tall* in the last chapter.

7 (3.46a) is acceptable when *completely* indicates a correction or is a marker of speaker confidence, but not as a degree modifier, at least on the usual assumptions.

8 Huffington Post, August 1, 2010 (http://www.huffingtonpost.com/2010/07/22/tmobile-iphone-in-2010-80_n_655459.html

9 Reuters, November 17, 2007 (http://www.reuters.com/article/idUSGOR68185720071117)

θ_{likely} is the vague threshold for counting as "likely". (In chapter 4 I will have more to say about how this threshold is determined.)

(3.50) $[\![\phi$ *is* **pos** *likely/probable*$]\!]^{\mathcal{M},w,g} = 1$ iff $\phi \succsim_{likely} \theta_{likely}$

3.4.2 *Certain*

On all of our tests, *certain* behaves like the maximum-standard adjectives *full* and *closed*, which fall on fully closed scales:

(3.51) a. It is completely certain that the Jets will win.
b. #It is slightly certain that the Jets will win.
c. #It is certain that the Jets will win, but it could be more certain.
d. It is not certain that the Jets will win. ⊨ It is (at least a little bit) uncertain.
e. It is slightly uncertain that the Jets will win.
f. It is almost certain that the Jets will win.
g. It is almost certain that the Jets will win. ⊨ It is slightly uncertain.
h. It is half/95% certain that the Jets will win.

These facts suggest the following truth-conditions for the positive form of *certain*:

(3.52) $[\![\phi$ *is* **pos** *certain*$]\!]^{\mathcal{M},w,g} = 1$ iff $\phi \approx_{certain} \top_{certain}$ iff $\mu(\phi) = \mu(\top_{certain})$.

Since all of these characteristics are shared with *full*, these adjectives should probably be associated with scales with the same structure (rather than, say, the weaker structure of $\mathcal{S}_{safe}$ which does not have a bottom element and does not allow proportional modifiers). In chapter 2 I argued that *full* is associated with a fully closed ratio scale, one with a minimum, maximum, and a concatenation operation which is positive and regular. This suggests the the scale of *certain* is also a fully closed ratio scale, as in (3.53).

(3.53) $\mathcal{S}_{certain} = \langle \Phi, \succsim_{certain}, \circ, \bot_{certain}, \top_{certain} \rangle$, where
a. Φ is a set of propositions;
b. $\succsim_{certain}$ is a weak order on Φ;
c. $\forall \phi \in \Phi : \top_{certain} \succsim_{certain} \phi \succsim_{certain} \bot_{certain}$;
d. $\succsim_{certain}$ is positive, regular, and Archimedean.

Note that $\mathcal{S}_{certain}$, like $\mathcal{S}_{likely}$, cannot be a pre-order because a ratio scale must contain at least a weak order.

3.4.3 Co-Scalarity of the Adjectival Epistemic Modals

3.4.3.1 Arguments for Co-Scalarity

A comparison of the proposed structures for $\mathcal{S}_{possible}$, $\mathcal{S}_{probable}$, $\mathcal{S}_{likely}$ and $\mathcal{S}_{certain}$ reveals that they are rather similar; the differences are in the degrees of freedom left available by the evidence we have seen so far. (I continue to bracket the availability of proportional modifiers with *possible, probable*, and *likely* for the time being.)

- The data we have seen underdetermine the position of $\mathcal{S}_{possible} = \langle \Phi, \succcurlyeq_{possible}, \ldots, \bot_{possible}, \ldots \rangle$ in our typology of scales in three ways:
  - $\succcurlyeq_{possible}$ may be a pre-order or a weak order;
  - There may or may not be a concatenation relation defined on $\mathcal{S}_{possible}$;
  - $\mathcal{S}_{possible}$ may or may not have a maximum element $\top_{possible}$.
- The data leave open only one parameter of variation for $\mathcal{S}_{probable}$ and $\mathcal{S}_{likely} = \langle \Phi, \succcurlyeq_{likely}, \circ, \ldots \rangle$ — they may or may not have an maximum element;
- $\mathcal{S}_{certain} = \langle \Phi, \succcurlyeq_{certain}, \circ, \bot_{certain}, \top_{certain} \rangle$ is fully determinate up to the expressive limits of the typology of scales developed in chapter 2.

I want to suggest that *possible*, *likely*, and *probable* in fact have more structure than the tests we have seen indicate, and that these scales are in fact the same: $\mathcal{S}_{possible} = \mathcal{S}_{probable} = \mathcal{S}_{likely} = \mathcal{S}_{certain}$. If this is right these items are all defined in terms of a single fully closed ratio scale $\mathcal{S}_{epistemic} = \langle \Phi, \succcurlyeq_{epistemic}, \circ, \bot_{epistemic}, \top_{epistemic} \rangle$. On this proposal, the positive form of *possible* refers to the minimum point of $\mathcal{S}_{epistemic}$; the positive form of the relative-standard adjectives *likely* and *probable* pick out a vague threshold somewhere in the midle of this scale; and the positive form of the maximum-standard adjective *certain* picks out the maximum point of $\mathcal{S}_{epistemic}$. In fact, this proposal seems quite natural given the clear semantic connections between these items, but it has been explicitly denied in recent work by Portner (2009), and so it is worth discussing in detail why it is probably the right account.

The first reason to think that these items fall on the same scale is that there are clear entailment relationships between these adjectives and their antonyms which are readily accounted for by this proposal.

(3.54) a. It is certain that we will win. $\vDash$ It is likely/probable/possible that we will win.
b. It is probable/likely that we will win. $\vDash$ It is possible that we will win.

(3.55) a. It is not possible that we will win. $\vDash$ It is not likely/probable/certain that we will win.
b. It is not likely/probable that we will win. $\vDash$ It is not certain that we will win.

These entailments are explained in a maximally simple fashion if these items fall on the same scale, and share an underlying order $\succcurlyeq_{epistemic}$.

Quantity implicatures point to the same conclusion:

(3.56) a. It is possible that we will win. $\rightsquigarrow$ It is not likely/probable/certain that we will win.
b. It is probable/likely that we will win. $\rightsquigarrow$ It is not certain that we will win.

Admittedly, neither of these arguments is an absolutely compelling reason to adopt the co-scalarity hypothesis. What they show is that the orderings $\succcurlyeq_{possible}$, $\succcurlyeq_{likely}$, and $\succcurlyeq_{certain}$ never disagree on the relative positions of two propositions. However, it could still be that one or more of these is a subset of the others. For example, we might consider the possibility that $\mathcal{S}_{certain}$ has a maximum element which $\mathcal{S}_{possible}$ and/or $\mathcal{S}_{likely/probable}$ lack (as suggested by Klecha 2011); in this case, $\succcurlyeq_{possible}$ and $\succcurlyeq_{likely/probable}$ would be proper subsets of $\succcurlyeq_{certain}$.

This question brings us back to proportional modification, which, recall, was the only data point on which *possible* and *likely* differed from *bent/dangerous* and *tall* respectively. We saw several examples of proportional modifiers with *possible*, *likely*, and *probable* above. Here are a few more:

(3.57) a. So overall a Slugger was about fifty percent likely to get better, a control was about thirty-three percent likely, and a Lagger was only about twenty-five percent likely to increase.[10]

b. Women who are 40 and older are only about five percent likely to have acne.[11]

c. **Q**: Do you think it would be difficult for a ross grad to match into a good radiology residency, and then go on to an interventional fellowship???
A: Yes. Longer answer: But not COMPLETELY 100 percent impossible. But probably REALLY HARD. I personally wouldn't set my heart on it. It's probably more like 99 percent impossible.[12]

d. It is very unlikely - 2% to 10% possible - but not impossible, that residential or occupational EMFs [electromagnetic fields] could be responsible for even a small fraction of birth defects, low birth weight, neonatal deaths, or cancer generally.[13]

If *possible, probable*, and/or *likely* were associated with scales that had no maximum element, these sentences would be just as bizarre as similar examples with *tall*:

(3.58) # Sam is 70 percent tall.

In support of this intuitive acceptability judgment, the difference in frequency between *percent likely/possible/probable* and *percent tall* is highly significant even by the very rough measure of Google hits. *Percent likely* occurs about 100 times as often as a proportion of total hits for *likely*, *percent probable* about 13 times as often, and *percent possible* about 240 times as often. (This does not take into account the use of the % sign, which Google does not register in searches, and which would very likely increase the ratio for *percent probable*).

On the other hand, the fact that these items differ from standard examples of minimum- and relative-standard adjectives *in precisely this one data point* is explained if their scales differ in having a maximum. Putting the point in RTM terms, the fact that sentences combining proportional modifiers with *possible* and *likely* can be interpretable entails that these adjectives are associated with scales which are upper- and lower-bounded and have a concatenation operation which is positive. But this means that the scales of all of these adjectives are isomorphic: they have the same underlying order on propositions, a positive concatenation operation, and are fully closed.

10 http://intl.feedfury.com/content/40488360-hittracker-data-part-ii.html
11 http://www.articlesbase.com/acne-articles/seven-myths-that-can-hinder-treatment-for-acne-767524.html
12 http://www.valuemd.com/ross-university-school-medicine/7387-ross-radiology-residency.html
13 California Health Dept. Report, April 2001, http://www.electric-fields.bris.ac.uk/Careport.pdf

3.4.3.2 Degree Modification Tests and an Argument Against Co-Scalarity

Portner (2009) points out two related problems for the idea that the adjectival epistemic modals occupy a single scale. Suppose that *completely* is a maximizer — *x is completely A* is true if and only if *x* has a maximum degree of property *A* — and that all of the adjectival modals that we are discussing fall onto a single fully closed scale. Then we do not seem to have any way to distinguish the meanings of the sentences in (3.59):

(3.59) a. It is completely possible that the Jets will win.
b. It is completely likely/probable that the Jets will win.
c. It is completely certain that the Jets will win.

Since these sentences clearly do not have the same meaning — in particular, only the example with *certain* seems to have a "maximum-degree" reading — Portner concludes that *certain* is on an upper-bounded scale, but *possible* and *probable* are not.

The second point is that *possible* and *probable* do not accept the same degree modifiers:[14]

(3.60) a. It is slightly possible that the Jets will win.
b. # It is slightly probable that the Jets will win.

Portner takes data like (3.60) as evidence that *possible* and *probable* do not occupy the same scale either. So we are back to the three-way split between $\mathcal{S}_{possible}$, $\mathcal{S}_{probable/likely}$, and $\mathcal{S}_{certain}$ that we started with.

The first thing to note here is that, if this argument is correct, then we are left with a serious problem in explaining the fact that all of these adjectives do take proportional modifiers (3.57). Even without bringing in this empirical problem, though, I don't think that the conclusion is inescapable. The conclusion follows only if we interpret the degree modification tests as "if and only if" statements about boundedness properties, which is probably too strong. I'll suggest that they are better seen, either as simple "if" statements about boundedness, or as "if and only if" statements about adjective type. On either interpretation, the argument against co-scalarity fails.

Suppose that an adjective allows modification by *slightly* with a "just above minimum" interpretation, as with our Kennedy (2007)-inspired examples repeated in (3.61).

(3.61) a. This antenna is slightly bent.
b. This neighborhood is slightly dangerous.

It seems clear that we can infer from this that $\mathcal{S}_{bent}$ and $\mathcal{S}_{dangerous}$ have minima. What do we say when the test comes up negative?

(3.62) a. # This neighborhood is slightly safe.
b. # This glass is slightly full.
c. # My brother is slightly tall.

14 Portner actually uses different examples to make this point. I use *slightly* because the examples that Portner marks as unacceptable (*completely/entirely probable*, *extremely/more possible*) are actually well-attested in corpora. *Slightly probable* is quite rare, and when attested does not have the "greater than minimum" reading that we are interested in here.

d. # This wine is slightly expensive.

In the case of (3.62a), we might guess that $\mathcal{S}_{safe}$ does not have a minimum — and, in this case, we would be right. However, making the same guess based on (3.62b) would be wrong — we know that $\mathcal{S}_{full}$ has a minimum because the scale associated with its antonym *empty* has a maximum (*completely empty*, i.e. zero fullness).

Similarly, we would be wrong to interpret (3.62c) and (3.62d) as indicating that *tall* and *expensive* do not have scales with minima. Since both adjectives allow ratio modifiers with the equative (*twice as tall/expensive*), we know that they are ratio scales, and this means that they cannot help but have an inherent minimum. (We might try to finess the problem with *tall* by pointing out that nothing in the adjective's domain can have zero height, but this would not work for *expensive*, since objects can have zero cost.)

However, all of the offending adjectives in (3.62) do share one important characteristic: none of them are minimum-standard adjectives. This, I suggest, is the reason that they do no accept *slightly*-modification. If I am right, then we have to interpret the *slightly*-modification test a bit more subtly, roughly as:

(3.63) a. If *x is slightly A* is felicitous with a "just above minimum" interpretation, then *A* is associated with a lower-bounded scale.

b. If *x* is in the domain of *A* and *x is slightly A* is infelicitous, then *A* is not a minimum-standard adjective.[15,16]

If this is the right way to interpret the *slightly*-modification test, then the argument against co-scalarity from data like (3.60) fails. The fact that *probable* does not accept *slightly*-modification doesn't show that its scale does not have a minimum element. Rather, it shows that *one or the other* of the conditions in (3.63) fails: either $\mathcal{S}_{probable}$ does not have a minimum, or *probable* is not a minimum-standard adjective. Since we have already seen that *probable* is a relative-standard adjective, and we have evidence from proportional modifiers that its scale has a maximum and a minimum, it seems clear which disjunct is operative in explaining the infelicity of (3.60a).

Similar reasoning explains the fact that *completely possible*, *completely probable*, and *completely certain* do not all mean the same thing, as in (3.59). The first thing to note is that *completely* is a

15 It might even be possible to recover an "if and only if" test by replacing (3.63a) with:

> If *x is slightly A* is felicitous with a "just above minimum" interpretation, then *A* is a minimum-standard adjective.

The interpretation of the test would now be, in effect: if *x* is in the domain of *A*, then *x is slightly A* is felicitous if and only if *A* is a minumum-standard adjective. Since being minimum-standard entails having a minimum, a positive result entails the condition (3.63a).

This interpretation is simpler, but it would seem to give the wrong result — as would the same move with *completely*, discussed below — because of the pair *open/closed*. As Kennedy (2007) points out, *slightly/completely open* and *slightly/completely closed* all seem to be acceptable in some contexts. I am not sure how to fit these adjectives into the typology of minimum-, relative-, and maximum-standard adjectives, but it would seem to be too strong to conclude — as the modified interpretation for *slightly* would have it — that both *open* and *closed* are minimum-standard adjectives; both of these also accept modification by *completely*.

16 Note, by the way, that there do seem to be tests that diagnose boundedness properties independent of scale type: the proportional modifer *n%* appears to function in this way, since it works for all our otherwise different epistemic adjectives.

polysemous modifier; in addition to its degree-maximizing meaning, it has the distributional reading exemplified in (3.64) and the emphatic reading exemplified in (3.65) (usually with a pitch accent on *completely*).

(3.64) This neighborhood is completely dangerous.
"It is dangerous everywhere"

(3.65) Mary: The president is not tall.
Sue: Nuh-uh! He is completely tall.

ϕ *is completely possible/probable/likely*, when they are acceptable, seem to have one of these meanings. For instance, Sue might respond to Mary's skepticism in (3.66) by using *completely possible* (with a pitch accent on *completely*).

(3.66) Mary: It's not possible that we will win this tournament.
Sue: Nuh-uh! It is completely possible.

Naturally-occurring examples generally have this character, for instance:

(3.67) It's completely possible to have a baby if you have BDD![17]

(3.68) It is completely possible to be a pro-life feminist. I am one.[18]

These are just two randomly chosen examples, but it's easy to find more; in contest, both examples are clearly framed as attempts to correct unwarranted skepticism, suggesting the emphatic/corrective interpretation as in (3.65).

The reason that *completely possible* and *completely probable* do not mean the same as *completely certain*, I suggest, is that the former are not examples of degree modification at all: when these are acceptable, they are examples of emphatic *completely*. This means that we are left with the question of why degree-modifying *completely* combines with *certain* but not with *possible, probable*, or *likely*. The situation is directly analogous to the case of *slightly*: degree-maximizing *completely* is not a diagnostic for the presence of a maximum element directly, but rather a diagnostic for whether an adjective is maximum-standard.

(3.69) a. If x *is completely* A is felicitous with a "maximum-degree" interpretation, then A is associated with an upper-bounded scale.
b. If x is in the domain of A and x *is completely* A is infelicitous, then A is not a maximum-standard adjective.

However, we can't conclude that $\mathcal{S}_A$ is not upper-bounded if A fails the test, because it may be that A simply isn't a maximum-standard adjective.

A further benefit of this interpretation is that it explains a related puzzle noted by Kennedy (2007). Since $\mathcal{S}_{expensive}$ is lower-bounded, the scale of its antonym *inexpensive* should be upper-bounded. So if the *completely*-modification test were an unambiguous diagnostic of upper-boundedness, then it would be a mystery why (3.70a) does not mean the same as (3.70b). Another example of the same puzzle is (3.71).

17 Body Dysmorphic Disorder, sufferers of which are preoccupied with their perceived physical defects. From http://bddcentral.com/forums/index.php?topic=9560.0.

18 http://stfusexists.tumblr.com/post/2143643768/it-is-completely-possible-to-be-a-pro-life-feminist-i

(3.70) a. This pizza is completely inexpensive.

b. This pizza is free.

(3.71) a. On that planet, you would be completely lightweight.

b. On that planet, you would be weightless.

The problem, and the solution, is exactly the same as with *completely likely/probable*: failure to maximize with *completely* is not necessarily indicative of the absence of an upper bound, but may simply indicate that the adjective is not maximum-standard. In these examples as well, other facts support this explanation, and the puzzle is resolved.

In sum, Portner's observations about differences in degree modification with adjectival epistemic modals are very much to the point, but they do not make a compelling case against co-scalarity. The fact that epistemic adjectives can behave differently on modification tests while also accepting proportional modifiers leads to a useful refinement of our understanding of the meaning of degree modification tests, though: as it turns out, in each case there are similar data involving non-modal adjectives which support the interpretation that I have given. Furthermore, if the co-scalarity hypothesis were false, it would be very difficult to explain the fact that all four adjectives accept proportional modifiers, exactly as we would expect if they inhabit a single, fully closed scale $\mathcal{S}_{epistemic}$.[19]

19 Without wishing to digress too much, I should note that this way of looking at the problem is rather different from Kennedy's (2007) approach. (Readers who are not invested in current debates in the theory of gradability can skip this lengthy footnote without loss of continuity.) One of Kennedy's goals in that paper is to derive adjective type from the boundedness properties of the scale that an adjective is on. Very roughly, the idea is that if a scale has a lower bound, then adjectives on it will be minimum-standard; if it has an upper bound, adjectives will be maximum-standard; if it has neither, they will be relative-standard; if it has both, they will be one or the other, or perhaps act like both. Kennedy attributes this pattern to a pragmatic principle "Interpretive Economy" which directs language users to minimize context-sensitivity in assigning truth-conditions. One of the advantages of Kennedy's approach is that tests for adjective type start to look like more informative "if and only if" statements again.

Reducing the number of independent dimensions of variation in adjective meaning is a very worthy goal, but the data that we have been looking at calls into question whether this reduction is possible. Since *n% likely* is acceptable, *likely* must have an upper and lower bound by Kennedy's own lights; but this means that, by Interpretive Economy, it should not be a relative adjective. Similarly, *expensive* and *tall* should not be relative adjectives since their scales — being ratio scales — have inherent minima; but they are. (Kennedy 2007 has a story about this in the case of *expensive*, but it is rather ad hoc — cf. Lassiter 2010a for discussion — and is not coherent with the way that ratio scales are defined, which requires the presence of a minimum. Similarly Klecha (2011) suggests that Interpretive Economy can be rescued if *likely* lacks a maximum element, but this claim is incompatible with his own argument that the scale of *likely* is a conditional probability measure. Ratio scales and probability scales have inherent endpoints, and it is not possible to remove them without drastically affecting other aspects of the scale.)

I do not know how to modify Kennedy's theory in order to capture this data. One possibility would be to weaken Interpretive Economy to a set of implicational statements about which adjective types can fall on which scales, e.g. "If an upper-bounded scale has only one adjective, then it will be maximum-standard; if a lower-bounded scale is ratio and has one adjective, it will be relative; if a lower-bounded scale is not ratio and has one adjective, it will be minimum", and so on. This approach is probably more in line with Potts's (2008) re-interpretation of Interpretive Economy as a historical-functional pressure, but preserves the core insight of Kennedy (2007) that there are non-arbitrary connections between scale type and adjective type. In any case, though, I suspect that the final story on this issue will be rather more complicated than any account currently available, and will rely on empirical generalizations that we do not currently possess; our current understanding of the semantics and pragmatics of degree modification really doesn't put us in a

3.4.4 Summary: The View from Degree Modification

The overall picture that emerges from the considerations of this section is this: the semantics of adjectival epistemic modals is built around a fully closed ratio scale $\mathcal{S}_{epistemic}$. *Possible* is a minimum-standard adjective, *probable* and *likely* are relative-standard, and *certain* is maximum-standard. These behave as expected from their adjective type with respect to degree modification, entailments, and implicatures, with one peculiarity: all of them, even the minimum- and relative-standard ones, accept certain proportional modifiers. This is explained by the fact that, unlike the other minimum- and relative-standard adjectives that we have seen, all of the epistemic adjectives fall on scales with an upper and lower bound.

3.5 The Scale of Epistemic Adjectives is a Probability Space

In this section I show that, if $\mathcal{S}_{epistemic}$ has the structure that I have argued for on the basis of degree modification data, it is provably equivalent to a finitely additive probability space. This is the core result of the chapter:

(3.72) Evidence from degree modification shows that the adjectival epistemic modals *possible, probable, likely*, and *certain* are all associated with a scale $\mathcal{S}_{epistemic}$ which is equivalent to a representation by finitely additive probability.

In §3.5.1 I will describe informally why this identification is reasonable; in §3.5.2 I will give some details of the proof and references to Narens (2007), with which it originates.

3.5.1 Qualitative and Quantitative Probability

By looking at the behavior of the adjectival epistemic modals in light of degree modification and the RTM semantics presented in chapter 2, we have arrived at the following picture of the epistemic scale:

(3.73) $\mathcal{S}_{epistemic} = \langle \Phi, \succsim_{epistemic}, \circ, \bot_{epistemic}, \top_{epistemic} \rangle$, where

a. Φ is a set of propositions;

b. $\succsim_{epistemic}$ is a weak order on Φ;

c. $\forall \phi \in \Phi : \top_{epistemic} \succsim_{epistemic} \phi$ and $\phi \succsim_{epistemic} \bot_{epistemic}$;

d. $\langle \Phi, \succsim_{epistemic}, \circ \rangle$ is a concatenation structure which is positive, regular, and Archimedean.[20]

The main empirical motivations for this particular choice of scale were the fact that proportional modifiers and ratio modifiers occur with the adjectival epistemic modals. In order to ensure that

position to constrain the theory of modality too strongly by assuming one particular theory such as Kennedy's (2007).

20 As usual in lower-bounded scales where some object may occupy the minimum, the positivity axiom has to be modified slightly to accommodate $\bot$, so that it reads:

Positivity: $\forall x \forall y$: if $(x \circ y)$ is defined and $x \not\approx \bot$, then $x \circ y \succ y$.

these come out as interpretable, we need a scale which is upper- and lower-bounded (3.73c) and for which all admissible μ are additive (equivalent to (3.73d), as shown in chapter 2).

It turns out that the structure $\mathcal{S}_{epistemic}$ is already very close to axiomatizations of qualitative probability discussed by Savage (1954); Krantz et al. (1971); Fine (1973); Fishburn (1986); Narens (2007) and many others. Authors who have contributed to this literature generally have one of two rather different motivations. On the one hand, qualitative probability is investigated with an eye to the logic of the English expressions "It is probable that ϕ", "It is more probable that ϕ than it is that ψ", and the like. This is essentially the goal of this chapter as well, although we are using the more general toolkit of modern formal semantics.

On the other hand, many authors have investigated qualitative probability in order to discover what assumptions are necessary in order for a qualitative probability structure to correspond to one of the various types of numerical probability, where propositions are mapped to real numbers in the range $[0,1]$. The approach here is generally to take a perspective essentially similar to RTM and ask what conditions need to be satisfied by a qualitative structure in order for it to uniquely characterize a given variety of numerical probability.

Our goals here are closer to those of the first set of theorists, but we can learn something interesting from the latter literature as well: although many different approaches to probability representations have been proposed, if $\mathcal{S}_{epistemic}$ has the structure in (3.73), then it shares most of the properties of qualitative axiomatizations corresponding to standard numerical probability. In fact, we can utilize a result due to Narens (2007) to prove that every admissible measure function on a fully closed ratio scale, including $\mathcal{S}_{epistemic}$, is isomorphic to a probability measure.

To see this, consider an axiomatization of numerical probability making use of FINITE ADDITIVITY. (N.B.: W is not required to be finite; rather, the name indicates that additivity is not required to hold for infinite sets of disjoint propositions, as in countably additive probability.)

(3.74) A **Finitely Additive Probability Space** is a triple $\langle W, \Phi, \mu \rangle$, where

a. W is a set of possible worlds;

b. $\Phi \subseteq \mathcal{P}(W)$ is an algebra of propositions (sets of worlds) containing W which is closed under union and complement;

c. $\mu : \Phi \rightarrow [0,1]$ is a function from propositions to real numbers in the range [0,1];

d. $\mu(W) = 1$;

e. **Additivity**: If A and B are in Φ and $A \cap B = \varnothing$, then $\mu(A \cup B) = \mu(A) + \mu(B)$.

These axioms entail that $\varnothing \in \Phi$ (since $W \in \Phi$ and Φ is closed under complement) and that $\mu(\varnothing) = 0$ (because of (3.74d) and (3.74e)).

Imagine that you have been given (3.74) and told that this μ is an admissible measure function for some qualitative structure $\mathcal{S}_P$ in the usual RTM fashion. We can learn a lot about $\mathcal{S}_P$ from (3.74): for instance, the domain of its underlying order $\succcurlyeq_P$ is a set of propositions; $\succcurlyeq_P$ is upper-bounded by W and lower-bounded by $\varnothing$. Furthermore, if concatenation is disjoint join as I argued in chapter 2, then — since join is union in the domain of propositions, and concatenation is defined only for disjoint objects — axiom (3.74e) corresponds to one of the familiar characteristics of ratio scales: $\mu(A \circ B) = \mu(A) + \mu(B)$ for all admissible μ. This suggests that $\mathcal{S}_P$ contains a concatenation operation $\circ$ which is additive (i.e., positive, regular, and Archimedean). A pretty good guess, then, is that $\mathcal{S}_P$ is a fully closed ratio scale.

3.5.2 Proof of the Equivalence

An important result due to Kraft, Pratt & Seidenberg (1959) shows that the most obvious ways to construct qualitative probability representations do not always have admissible measure functions which are consistent with (3.74). This result has led to a wide variety of proposals for further axioms to rule out the counter-examples — see Krantz et al. 1971; Fine 1973; Fishburn 1986 for surveys — and it might appear to pose problems for us as well.

However, it turns out that, due to the combination of RTM with algebraic formal semantics that we are making use of, we are able to bypass this issue. Narens (2007: 31-33) shows that all admissible measure functions on a fully closed ratio scale are isomorphic to finitely additive probability measures, as long as the set of propositions Φ is not extremely small, and a few structural axioms are satisfied. Kraft et al.'s (1959) problematic example did not make these additional assumptions; however, as it turns out, they all follow either from standard assumptions in formal semantics or from the ratio scale axioms discussed in chapter 2.

Specifically, Narens' proof relies on the following assumptions, all of which are satisfied by the fully closed ratio scale $\mathcal{S}_{epistemic}$ in combination with the RTM approach to scalar semantics laid out in chapter 2. Take $\langle X, \succcurlyeq, \circ, m \rangle$, where $\succcurlyeq$ is a weak order, m is a $\succcurlyeq$-maximal element, $\circ$ is a partial binary operation, and for all $x, y, z \in X$:

A. If $x \circ y$ is defined and $x \succcurlyeq u$ and $y \succcurlyeq v$, then $u \circ v$ is defined.

B. (Positivity) If $x \circ y$ is defined, then $(x \circ y) \succ x$.

C. (Associativity) If $x \circ (y \circ z)$ is defined, then $(x \circ y) \circ z$ is defined and $x \circ (y \circ z) \approx (x \circ y) \circ z$.

D. (Regularity) If $x \succ y$, then there is some $z \in X$ such that $x \succcurlyeq (y \circ z)$.

E. There are x and y in X such that $x \circ y = m$.

F. (Archimedean) Every strictly bounded standard sequence is finite.

(A) and (F) are part of the definition of a concatenation structure. (B) and (D) are definitional of ratio scales. (C) follows from the interpretation of $\circ$ as restricted join. (E) is fulfilled as long as the set of propositions forms a Boolean algebra, a standard assumption in formal semantics, and is not trivially small (i.e. there is more than one possible world).

Narens shows that all admissible measure functions relative to a structure satisfying these axioms are order-preserving and finitely additive.

Ex. There is a $\mu : X \to \mathbb{R}^+$ such that, for all $x, y \in X$, $x \succcurlyeq y \equiv \mu(x) \geq \mu(y)$, and if $x \circ y$ is defined, then $\mu(x \circ y) = \mu(x) + \mu(y)$.

Un1. For any two μ, μ' satisfying (Ex.), there is some $r \in \mathbb{R}^+$ such that $\mu(x) = r \times \mu'(x)$ for all $x \in X$.

Un2. For any μ satisfying (Ex.) and any $r \in \mathbb{R}^+$, there is a μ' satsifying (Ex.) such that $\mu'(x) = r \times \mu(x)$ for all $x \in X$.

These axioms characterize an infinite set of finitely additive probability measures, but there is only one with $\mu(m) = \mu(W) = 1$, which is the one which we conventionally focus on.

As a result, if we know that we are dealing with a scale $\mathcal{S}_{epistemic}$ which is fully closed and additive, we are necessarily dealing with a representation at least as rich as a finitely additive probability measure. Since the degree modification data seem to lead inexorably to the former, we have no choice but to embrace the latter as well: adjectival epistemic modals have a semantics built around probability.

For the remainder of this dissertation I will make use of this equivalence and employ the qualitative and quantitative characterizations of finitely additive probability interchangeably to characterize the scales of adjectival epistemic modals.[21]

In addition to drawing useful connections between RTM semantics for gradability and the literature on qualitative probability, the conclusion that adjectival epistemic modals have a scale built on probability will be useful in later chapters. There I will argue that the use of probabilistic representations, in combination with the RTM approach to scalar semantics sketched in chapter 2, explains a number of further puzzles which arise for epistemic, deontic, and bouletic modals. In chapter 4 I show that this approach explains the puzzling fact that judgments of probability in experimental settings are sensitive to the distribution of alternatives. In chapter 6 I will show that we can use probabilistic information, along with a slightly enriched notion of preference, to build scales for bouletic and deontic modals which avoid the deep problems which plague quantificational semantics for these expressions.

---

21 I would be remiss if I did not mention that we need further conditions to get a qualitative structure which picks out a countably additive probability measure, as axiomatized by Kolmogorov (1933):

(3.75) A **Countably Additive Probability Space** is a triple $\langle W, \Phi, \mu \rangle$ which is a finitely additive probability space and in addition:

a. Φ is a σ-algebra (closed under countable union);

b. **Countable Additivity**: If $\{A_1, A_2, ...\}$ is a (possibly infinite) set of mutually exclusive propositions each of which is in Φ, then

$$\mu\left(\bigcup_{i=1}^{\infty} A_i\right) = \sum_{i=1}^{\infty} \mu(A_i)$$

This is the standard form of probability employed in most of mathematics, statistics, etc. It is possible to restrict $\mathcal{S}_{epistemic}$ further so that the adjectival modals in English would be associated with countably additive probability measures, though I do not know whether it is desirable; Kolmogorov's axioms have also been questioned on a number of grounds. For example, Fine (1973: 65) describes them as "restrictive and arbitrary"; Narens (2007) argues that finite additivity is sufficient, although this claim is quite controversial.

As far as I can tell, the answer to this question is not too crucial for our purposes: natural language data will probably not enable us to distinguish finitely and countably additive probability, and some other kind of evidence is needed to determine whether the stricter requirement of countable additivity better characterizes the way that people reason using adjectival epistemic modals.

3.6 The Puzzles Resolved

Associating the adjectival epistemic modals with the fully closed ratio scale $\mathcal{S}_{epistemic}$ allows us to apply the RTM semantics for gradability and comparison from the last chapter directly to these items. In effect, this means that ϕ *is more likely than* ψ is true if and only if $\phi \succ_{epistemic} \psi$. This, in turn, is true if and only if, for some (equivalently, all) μ which is an admissible measure function for $\mathcal{S}_{epistemic}$, $\mu(\phi) > \mu(\psi)$.

Without loss of generality, we can restrict our attention to the unique admissible μ for which $\mu(W) = 1$. I will give this familiar μ a special name, *prob*, and use it throughout the rest of the dissertation. So, for example, instead of writing "For all $\mathcal{S}_{epistemic}$-admissible μ, $\mu(\phi) > \mu(\psi)$", I will just write "$prob(\phi) > prob(\psi)$". It should be clear that no information is lost in this practice, since all of the $\mathcal{S}_{epistemic}$-admissible μ are isomorphic.

With this convention in hand, this section will demonstrate that the three problems for Kratzer's semantics for epistemic modals that I pointed out in §3.2 do not arise for the scalar semantics proposed here.

3.6.1 Comparatives, Equatives, and Disjunction

The first problem noted in §3.2 involved the fact that Kratzer's semantics incorrectly predicts that the inference in (3.81) should be valid.

(3.76) **The Disjunctive Inference**

a. ϕ is at least as likely as ψ.

b. ϕ is at least as likely as χ.

c. ∴ ϕ is at least as likely as $(\psi \vee \chi)$.

As a counter-model to (3.81), suppose that $prob(\phi) = 0.3$, $prob(\psi) = 0.2$, $prob(\chi) = 0.2$, and $prob(\psi \wedge \chi) = 0$. Then (3.76a) and (3.76b) come out true, but the conclusion is false: by additivity $prob(\psi \vee \chi) = prob(\psi) + prob(\chi) = 0.4$, which is greater than $prob(\phi) = 0.3$. This is the result that we want: the inference is not valid in the scalar semantics that I have proposed simply because $\mathcal{S}_{epistemic}$ is **additive** with respect to concatenations. ($\mathcal{K}$-structures, in contrast, are **maximal** with respect to concatenations of comparable propositions, as noted in §3.3.2 above.)

In fact (3.81) is invalid for precisely the same reason that (3.77) is, or any similar inference pattern with ratio scale properties:

(3.77) a. x is longer than y.

b. x is longer than z.

c. ∴ x is longer than $(y \circ z)$.

The counter-model to (3.77) is identical: let x be 0.3 feet long, y 0.2 feet long, and z 0.2 feet long. Then $x \circ y$ is 0.4 feet long, contradicting the conclusion.

Similarly, in the baseball example, it is easy for the Blue Jays to have a better chance of winning the World Series than any other without being more likely to win than not to win. All we need is that, for example, the Blue Jays have a 20% chance of winning, and each of the other 29 teams has a chance less than 20%. It is still the case that the Blue Jays are overwhelmingly more likely not to win (80%) than they are to win (20%); so the intuitively invalid inference from *The Blue Jays are more likely to win than anyone else is* to *The Blue Jays are more likely to win than not to win* is not predicted.

3.6.2 Degree Modification and Interpretability

The second problem in §3.2 was that the sentence-types in (3.82) do not have stable truth-values across $\mathcal{K}$-admissible μ, and so are predicted to come out as uninterpretable in the RTM sense.

(3.78) a. ϕ is twice as likely as ψ.
b. It is half certain that ϕ.
c. It is 95% certain that ϕ.

These are all interpretable in the scalar theory proposed here. In particular, since all $\mathcal{S}_{epistemic}$-admissible μ are isomorphic to *prob*:

(3.79) a. $[\![(3.78a)]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) = 2 \times prob(\psi)$
b. $[\![(3.78b)]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) = 0.5 \times prob(W) = 0.5$
c. $[\![(3.78c)]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) = 0.95 \times prob(W) = 0.95$

Ratios, proportions, and the relative size of intervals are all preserved in all $\mathcal{S}_{epistemic}$-admissible $\mu_{epistemic}$, since the latter are all isomorphic to *prob*.

3.6.3 Incomparabilities

Kratzer's theory encountered a serious problem with incomparabilities: in many cases, the semantics makes it impossible to assign truth-values to epistemic comparatives and equatives which deserve to be evaluable. In contrast, the scalar theory that we arrived at by considering the adjectival epistemic modals contains no incomparabilities: $\succcurlyeq_{epistemic}$ is a weak order, and so for any ϕ, ψ in the domain Φ, either $\phi \succcurlyeq_{epistemic} \psi$ or $\psi \succcurlyeq_{epistemic} \phi$.

This feature of probability representations is a good result overall, at least compared to its prominent rival. However, the connectedness of probability has been criticized by some as too strong a property to ascribe to intuitive probability judgments (e.g. Keynes 1921). The point is not unreasonable, I think: many of us will have no intuitions at all about the relative probability of the two sentences in (3.80).

(3.80) a. It will be sunny in Zanzibar tomorrow.
b. I will get an A on my math test.

Now, our lack of intuitions about the proposition does not demonstrate conclusively that these propositions are epistemically incomparable. First, given the unrelated subject matter of these sentences, it is hard to imagine someone having access to the information needed to make an informed comparison here. Lack of access to the information needed to evaluate a sentence, even principled lack of access, is not the same thing as the sentence itself lacking a truth-value.

Second, given that these propositions appear to be totally unrelated, there may be a pragmatic explanation of why it seems strange to ask which is more likely. Indeed, if you try to imagine a situation in which the comparison between these propositions is relevant for some practical purpose, the perception of incomparability is much less strong.

However, if Keynes and other critics are right in saying that not all epistemic comparisons are defined, existing tools allow us to introduce incomparability as needed. As van Rooij (2009) points out with respect to adjectives like *clever*, it is possible to capture incomparability without weakening scales too much by treating a scale as a set of subscales. Each of these is built around a connected order, and the global truth-conditions of $x \succcurlyeq_P y$ rely on universal quantification over all the value of $x \succcurlyeq_p y$ in all subscales of $\mathcal{S}_P$. So, for example, the semantics would treat *John is cleverer than Mary* as true if John is cleverer than Mary with respect to every subscale in $\mathcal{S}_{clever}$, false if John is not cleverer than Mary with respect to any subscale in $\mathcal{S}_{clever}$, and undefined otherwise.

Similarly, we can avoid legislating epistemic comparability, if we wish to, without making drastic changes to $\mathcal{S}_{epistemic}$ by building the semantics around a set of scales each of which is a fully closed ratio scale. This is actually closely related to an approach to uncertainty about probability assignments which has been explored by a number of authors, using sets or ranges of probability measures (see Halpern 2003: ch. 2 for an overview).

The advantage of this semantics over Kratzer's, then, is pretty clear: it is not at all clear whether we want any epistemic incomparability; but if we do, the present proposal allows us to build models which contain exactly as much epistemic comparability as is warranted by the data, and no more. In contrast, Kratzer's semantics forces us to declare far too many epistemic comparatives undefined, including ones which are intuitively quite reasonable.

3.7 Kratzer (2012) on Orders and Probability

This section considers some additional issues relating to the critique of Kratzer's theory that I have offered that are raised by Kratzer 2012, a forthcoming book revising and updating Kratzer's classic papers on modality and conditionals.[22] In a section which has has already gotten a fair bit of attention before its publication, Kratzer offers some modifications to the influential theory that I have critiqued and discusses relevant issues involving graded modality and its relation to the concept of probability. There Kratzer acknowledges that there are problems with the originally definition of comparative possibility that we have been considering and suggests an alternative. This is intriguing since most of the objections to Kratzer's theory in Yalcin 2010; Lassiter 2010a and this chapter are specific to the original comparative possibility relation $\succcurlyeq^s_{\mathbf{g}(w)}$, and a different relation could in

22 This section differs from §3.7 in the version of this dissertation originally distributed in September 2011. In that version, I inadvertently made use of the definition of comparative possibility from Kratzer 1981 rather than the new definition proposed in Kratzer 2012, misrepresenting her position regarding probability in the process. This was pointed out to me by Paul Portner, Aynat Rubinstein, and Angelika Kratzer soon after the original version was released. I owe them thanks for the correction. Since the main points of the dissertation do not rely directly on the affected portion of this chapter, I have taken the liberty of correcting this mistake in the publicly available version for the benefit of subsequent readers and to combat any misinformation that may have been caused by my error. While the details of the argumentation in this section differ as a result, the conclusion is similar. I have also made some slight modifications to §3.8 to improve continuity given the changes in this section. (DL, 11/17/11)

principle avoid some or all of them. Even if not, the connection with probability that she offers might render viable a principled mixture of the two approaches, where epistemic adjectives are probabilistic while epistemic auxiliaries are (suitably restricted) quantifiers over possible worlds.

However, as I will show, both of these ideas have problems. While the new comparative possibility relation addresses a few of the counter-examples, it validates a slightly restricted version of the disjunctive inference which remains empirically problematic. This fact places strong constraints on the class of probability measures that can be defined consistently with the new comparative possibility relation, which are as a result too restricted to form the basis either of a semantics for epistemic modals or of a principled connection between modal semantics and probability as it is used in scientific practice.

3.7.1 Orders

Citing Yalcin's discussion, Kratzer (2012: 41) notes that the original definition of comparative possibility from Kratzer (1981, 1991) has "consequences that might be unwelcome for certain applications":

> Suppose, for example, that there is a world w that is better than any other world. We would now predict that all propositions containing w are equally good possibilities. W and $\{w\}$ are equally good possibilities, then.

This is a special case of the disjunctive inference discussed in some detail above: since W is the union of all the singletons in it, it is effectively a big disjunction of very small propositions. It follows from the proof of (3.15) in §3.3.1 that, whenever $\{w\}$ is (according to the original definition) as good a possibility as each of these singletons individually, it will also be as good a possibility as the entire set of worlds W.

Kratzer (2012) suggests adopting a different definition of comparative possibility, motivated in part by the need to respond to this concern. (Again, I have modified the notation to match with what I am using in this dissertation, so that Kratzer's $<_{\mathbf{g}(w)}$ order on worlds corresponds to $>_{\mathbf{g}(w)}$ here, etc.)

(3.81) **Comparative possibility** (new definition):
p is at least as good a possibility as q in a world w with respect to a modal base $\mathbf{f}$ and an ordering source $\mathbf{g}$ if and only if

$$\neg\exists u \in (q-p)\ \forall v \in (p-q)[u >_{\mathbf{g}(w)} v] \quad (\text{where } u, v \in \bigcap \mathbf{f}(w))$$

For notational convenience and to distinguish it from the original relation $\succcurlyeq^{s}_{\mathbf{g}(w)}$ I will give this new relation a name, $\succcurlyeq^{N}_{\mathbf{g}(w)}$ (with ^{N} for "new"). The definition in (3.81) is, of course, equivalent to (3.82)—

(3.82) $p \succcurlyeq^{N}_{\mathbf{g}(w)} q$ iff $\forall u \in (q \wedge \neg p) \exists v \in (p \wedge \neg q)\ [\neg(u >_{\mathbf{g}(w)} v)]$

—assuming again that u and v are in $\bigcap \mathbf{f}(w)$.

In words, the new rule ranks two propositions p and q by ignoring worlds that are in both, comparing only worlds that are in one but not the other. In addition to weakening the conditions on ordering among propositions by quantifying over a smaller set of worlds from each proposition, this definition places a weaker condition on the relationship between worlds in the portions of the propositions that are deemed relevant. The original order $p \succcurlyeq^{s}_{\mathbf{g}(w)} q$ required that all of the q-worlds be comparable to and weakly dominated by at least one p-world; but the new order $\succcurlyeq^{N}_{\mathbf{g}(w)}$ requires only that all of the relevant q worlds be either (a) weakly dominated by or (b) incomparable with at least one relevant p-world. This is a much easier condition to fulfill: for example, if a single $(p \wedge \neg q)$-world is $\succcurlyeq_{\mathbf{g}(w)}$-incomparable with all $(q \wedge \neg p)$-worlds, then p is automatically at least as good a possibility as q (Kratzer 2012: 42). In the old order $\succcurlyeq^{s}_{\mathbf{g}(w)}$, this would have led to incomparability between p and q. This difference may alleviate the problems with excessive incomparabilities that were discussed earlier in this chapter, though some incomparabilities will remain (and it will be necessary to consider in detail whether the predicted ones are empirically justified, a task I will not undertake here).

The new order on propositions is also an improvement with respect to the most troubling problem discussed in this chapter, the disjunctive inference. This inference is not valid with $\succcurlyeq^{N}_{\mathbf{g}(w)}$, a fact which resolves two disjunction-related problems that we have discussed. First, it solves the problem from the quote at the beginning of this section: if there is a best world w', it does not follow that $\{w'\} \approx^{N}_{\mathbf{g}(w)} W$. Second, the inference from $\phi \succcurlyeq^{N}_{\mathbf{g}(w)} \neg\phi$ to $\phi \succcurlyeq^{N}_{\mathbf{g}(w)} (\phi \vee \neg\phi)$ is not valid (as it was with $\succcurlyeq^{s}_{\mathbf{g}(w)}$, cf (3.23) and Yalcin 2010).

While these results are welcome, the new comparative possibility relation does not get to the heart of the disjunction issue. Recall the form of the disjunctive inference, given in (3.83) using the new relation.

(3.83) **Disjunctive inference**:

a. $\phi \succcurlyeq^{N}_{\mathbf{g}(w)} \psi$

b. $\phi \succcurlyeq^{N}_{\mathbf{g}(w)} \chi$

c. $\phi \succcurlyeq^{N}_{\mathbf{g}(w)} (\psi \vee \chi)$ (**not valid**)

Counter-models to (3.83) have a very specific form: the only way that the inference can fail is if one of the conditions in (3.84) holds.[23]

(3.84) a. $\exists u \in (\psi \wedge \neg\phi)\ \exists v \in (\phi \wedge \chi)[\neg(u \succ^{N}_{\mathbf{g}(w)} v)]$

b. $\exists u' \in (\chi \wedge \neg\phi)\ \exists v' \in (\phi \wedge \psi)[\neg(u' \succ^{N}_{\mathbf{g}(w)} v')]$

23 **Proof**: Suppose (3.83a) and (3.83b) are true and (3.83c) is false. Since (3.83c) is false, there is a world u in $((\psi \vee \chi) \wedge \neg\phi)$ such that, for all v in $(\phi \wedge \neg(\psi \vee \chi))$, $u \succ^{N}_{\mathbf{g}(w)} v$. World u is of course either in $(\psi \wedge \neg\phi)$ or in $(\chi \wedge \neg\phi)$. (All world variables in this and the next proof are implicitly restricted to $\bigcap \mathbf{f}(w)$, except for the distinguished actual world variable w.)

Suppose first that $u \in (\psi \wedge \neg\phi)$. Since (3.83a) holds, there is no world in $(\psi \wedge \neg\phi)$ that strictly dominates all worlds in $(\phi \wedge \neg\psi)$; and so in particular u does not strictly dominate all of these worlds, i.e. there is some world $v' \in (\phi \wedge \neg\psi)$ such that $\neg(u \succ^{N}_{\mathbf{g}(w)} v')$. But since u strictly dominates all worlds in $(\phi \wedge \neg(\psi \vee \chi))$, v' must be in $(\phi \wedge \neg\psi \wedge \chi)$. So, of course, v' is in $(\phi \wedge \chi)$. A parallel argument shows that, if u is in $(\chi \wedge \neg\phi)$, there must be some world v'' in $(\phi \wedge \psi)$ such that $\neg(u \succ^{N}_{\mathbf{g}(w)} v'')$. But since u is either in $(\psi \wedge \neg\phi)$ or in $(\chi \wedge \neg\phi)$, one of the two conditions in (3.84) must hold, and in particular there must be worlds either in $(\phi \wedge \psi)$ or in $(\phi \wedge \chi)$.

As a result, a counter-model to the disjunctive inference must contain worlds either in $(\phi \wedge \chi)$ or in $(\phi \wedge \psi)$. This explains why Yalcin's problem (3.23) and the problem that Kratzer notes in the quote above are avoided by adopting $\succcurlyeq^N_{\mathbf{g}(w)}$: substituting ϕ for ψ in (3.83) guarantees that there will be worlds in $(\phi \wedge \psi)$. The inference will only be valid if there are more than two disjoint alternatives under consideration.

If the alternatives are disjoint, though, there are no worlds in $(\phi \wedge \chi)$ or in $(\phi \wedge \psi)$ and the inference goes through. This was the case in the first empirical problem that was raised in §3.3.1, the baseball example. Given the structure of baseball, if the Blue Jays win the World Series (ϕ), the Red Sox do not ($\neg\psi$) and the Braves do not ($\neg\chi$). This presumably means that there are no worlds in $\bigcap \mathbf{f}(w)$ in which two or more teams win the World Series. But if this is the case neither of the conditions in (3.84) can be satisfied, and it is not possible to construct a counter-model; so the disjunctive inference in (3.83) goes through, and the problem remains.

More generally, if we modify the disjunctive inference by explicitly encoding the fact that we are comparing more than two mutually exclusive events, it is valid (with a proof very similar to the proof that the original disjunctive inference was valid with $\succcurlyeq^s_{\mathbf{g}(w)}$).

(3.85) **Modified disjunctive inference**:

a. $(\phi \wedge \neg\psi \wedge \neg\chi) \succcurlyeq^N_{\mathbf{g}(w)} (\neg\phi \wedge \psi \wedge \neg\chi)$

b. $(\phi \wedge \neg\psi \wedge \neg\chi) \succcurlyeq^N_{\mathbf{g}(w)} (\neg\phi \wedge \neg\psi \wedge \chi)$

c. $\therefore\ (\phi \wedge \neg\psi \wedge \neg\chi) \succcurlyeq^N_{\mathbf{g}(w)} ((\neg\phi \wedge \psi \wedge \neg\chi) \vee (\neg\phi \wedge \neg\psi \wedge \chi))$

Proof: By the definition of $\succcurlyeq^N_{\mathbf{g}(w)}$, (3.85a) means that, for any world $u \in (\neg\phi \wedge \psi \wedge \neg\chi)$ there is a world $v \in (\phi \wedge \neg\psi \wedge \neg\chi)$ such that $\neg(u \succ_{\mathbf{g}(w)} v)$. (The restrictions simplify because the propositions compared are disjoint, so the use of set subtraction in the new definition has no effect.) Similarly, (3.85b) means that for any world $u' \in (\neg\phi \wedge \neg\psi \wedge \chi)$ there is a world $v' \in (\phi \wedge \neg\psi \wedge \neg\chi)$ such that $\neg(u \succ_{\mathbf{g}(w)} v)$.

Let z be an arbitrary world in $((\neg\phi \wedge \psi \wedge \neg\chi) \vee (\neg\phi \wedge \neg\psi \wedge \chi))$. Case 1: $z \in (\neg\phi \wedge \psi \wedge \neg\chi)$. Then there is a world $z' \in (\phi \wedge \neg\psi \wedge \neg\chi)$ such that $\neg(z \succ_{\mathbf{g}(w)} z')$, namely v. Case 2: $z \in (\neg\phi \wedge \neg\psi \wedge \chi)$. Then there is a world $z'' \in (\phi \wedge \neg\psi \wedge \neg\chi)$ such that $\neg(z \succ_{\mathbf{g}(w)} z'')$, namely v'. Since z was arbitrary, we conclude that, for any $z \in ((\neg\phi \wedge \psi \wedge \neg\chi) \vee (\neg\phi \wedge \neg\psi \wedge \chi))$ there is some $z''' \in (\phi \wedge \neg\psi \wedge \neg\chi)$ such that $\neg(z \succ_{\mathbf{g}(w)} z''')$. This is equivalent to $(\phi \wedge \neg\psi \wedge \neg\chi) \succcurlyeq^N_{\mathbf{g}(w)} ((\neg\phi \wedge \psi \wedge \neg\chi) \vee (\neg\phi \wedge \neg\psi \wedge \chi))$ by the definition in (3.81), and so (3.85) is valid.

For the primary empirical objection to Kratzer's theory involving disjunction discussed in §3.3.1 the alternatives are disjoint anyway, and it does not make any difference if we mention this explicitly in the proof. Since the proof goes through in this case, modifying the ordering on propositions as in (3.81) does not resolve the problem: both the old and the new versions of comparative possibility yield incorrect predictions about the interaction between epistemic modals and disjunction.

The new version of Comparative Possibility is, admittedly, given with the caveat that it is "one option among many that should be considered" (Kratzer 2012: 41). One way to interpret the result here is simply that this particular idea is wrong, but that we will eventually be able to find a way of ranking propositions that gets the interaction with disjunction right. In a way, I think that this is exactly right (indeed, my own proposal has this form). However, I doubt seriously that this result will ever be achieved if the only tool that we have is quantification over possible worlds, however complicated we make the restrictions of the quantifiers. The main lesson of this chapter is that epistemic modality is **positive**, as defined in ch. 2, §2.2.1 — the likelihood of a disjunction of two disjoint propositions is strictly greater than the likelihood of either disjunct, unless one of them is impossible (has zero likelihood). More than that, epistemic modality is **additive** (ch.2, §2.1.2.2) — the likelihood of a disjunction of disjoint propositions is the sum of the likelihoods of the disjuncts. While I don't have an impossibility proof to offer, it strikes me as very doubtful that any straightforward modification of Kratzer's semantics, or other theories which trade primarily in quantification over possible worlds, will be able to achieve this empirically necessary result.[24]

There is, of course, a well-understood way to create a logic for epistemic modals and other additive properties using quantification; but that way is Measurement Theory, and it takes orders to be basic rather than derived, makes use of quantification over sets of worlds (equivalence classes) rather than worlds, and looks very different in many other details from theories that are typically used in modal semantics. An adequate semantics for epistemic modals will, I believe, need to be built on these or related tools, such as those that we have at hand in degree semantics.

3.7.2 Probability

The problem discussed in the last section involved the relation between the newly proposed ordering on propositions $\succsim^N_{\mathbf{g}(w)}$ and the English construction *as likely as*. Here is a possible way out, then: we might suppose that the meaning of *likely* is not given directly in terms of the ordering $\succsim^N_{\mathbf{g}(w)}$, but in terms of some scale of propositions which is defined indirectly using this order. Given the previous results in this chapter, this scale should have both top and bottom elements, and should be additive (which should allow us to avoid the disjunctive inference). Of course, this makes the scale equivalent to a probability measure (§3.5). Can we define a relationship of consistency between a particular choice of $\succsim^N_{\mathbf{g}(w)}$ and a probability measure that will draw a principled connection between comparative possibility and probability?

Here is one possibility, inspired by discussion in Kratzer (2012: 42-3).

(3.86) **Compatibility**:
$\succsim^N_{\mathbf{g}(w)}$ and $prob(\cdot)$ are *compatible* iff$_{df}$ $prob(p) \geq prob(q)$ whenever $p \succsim^N_{\mathbf{g}(w)} q$.

I should add that Kratzer does not propose this rule explicitly; I am inferring it from an example that she gives. (The text actually suggests a stronger interpretation, with "when and only when" replacing "whenever", but the formulation in (3.86) is preferable because it will also work in models in which $\succsim^N_{\mathbf{g}(w)}$ is not connected.)

Now, Kratzer clearly does not intend for probability measures so constructed to be used in the semantics: as she says, "[o]ur semantic knowledge alone does not give us the precise quantitative notions of probability and desirability that mathematicians and scientists work with" (p.25). Instead, a construction like Compatibility allows comparative possibility to provide the "conceptual launch

24 This is a challenge to those who wish to defend the analysis of modals as restricted quantifiers over possible worlds: your first order of business should be to devise a semantics for epistemic modals which does not validate the disjunctive inference in any form, and — in particular — is additive.

pads for mathematical explorations to take off from" (ibid.). This reasoning seems to assume that there is no semantic or psychological motivation for treating probability as a deep part of our cognitive lives, as opposed to just math; of course, the results of this chapter and the next, as well as the considerable use of probabilistic models in modern psychology, call this assumption into question. Indeed, as the discussion in earlier sections of this chapter has hopefully made clear, the distinction between qualitative (comparative) and quantitative (numerical, metrical) probability is much less sharp than the heavy mathematics employed in some uses of probability theory might suggest. It is possible — and very natural from the standpoint of an intensional semantics outfitted with tools for scalar reasoning — to define probability without mentioning numbers at all (§3.5). All of this points to the fact that the point of using probabilistic models is not numerical manipulation but deriving qualitatively correct patterns of inference, e.g. the empirically necessary property of positivity. This should be the focus of the discussion about the relevance of probability to semantics, rather than whether or not quantitative techniques are used. As Glenn Shafer put it: "Probability is not really about numbers; it is about the structure of reasoning" (cf. Pearl 1988:15).

In any case, there is good reason to consider both the intended interpretation of Compatibility and a more ambitious one. Since the new version of Comparative Possibility still has serious problems with disjunction and in particular does not capture positivity or additivity, a hybrid semantics which connects the new version of Comparative Possibility with a probabilistic semantics for *likely* appears to be the best hope for Kratzer's theory given the results of the last section. For this reason, I will consider here both the use that Kratzer proposes and a semantic use that goes beyond the intended one. As it turns out, though, both are problematic: the class of probability measures that can satisfy Compatibility in any reasonably large model is a tiny subset of the probability measures that can be defined in accordance with the usual axiomatization, and the restrictions added by the need to cohere with a $\succcurlyeq^N_{\mathbf{g}(w)}$ relation limit the empirical coverage of probability excessively for either purpose.

The first thing to note is that Compatibility is not — and is not meant to be — a way to *derive* probability from Comparative Possibility.[25] As Kratzer points out, there will always be many probability measures that preserve the structure of a particular instantiation of $\succcurlyeq^N_{\mathbf{g}(w)}$. In fact, all Compatible probability measures are admissible measure functions on $\succcurlyeq^N_{\mathbf{g}(w)}$ in the sense that this term was defined in chapter 2, §2.1.2.1. But there will also be many admissible measure functions on $\succcurlyeq^N_{\mathbf{g}(w)}$ that are not probability measures, and the ones that are have no special theoretical status *qua* measure functions. For example, most admissible μ will not be additive, so that additivity is not a property of a $\mathcal{K}$-structure built around a $\succcurlyeq^N_{\mathbf{g}(w)}$ relation. In addition, if $\succcurlyeq^N_{\mathbf{g}(w)}$ and *prob* were related in the way familiar from RTM, we would find that ratio and proportional modifiers are uninterpretable with epistemic adjectives since their truth-values will not generally be consistent across $\succcurlyeq^N_{\mathbf{g}(w)}$-admissible μ. Compatibility is meant instead as a way to check consistency between a pre-existing probability measure and a qualitative order; it does not tell us where the probability measure comes from.

25 The earlier version of this section misconstrued Kratzer on this point. Additional discussion in a recent version of the book manuscript, as well as subsequent discussions with Paul Portner, have helped me see that the proposed connection between comparative possibility and probability is weaker than I had thought.

Very well, then, what kind of probability measures satisfy this constraint?[26] Since the disjunctive inference holds in $\succcurlyeq^N_{\mathbf{g}(w)}$ for disjoint propositions ϕ, ψ, χ (as I proved in the previous subsection), it will not be possible to construct a Compatible probability measure which assigns a greater probability to the disjunction of two lower-ranked propositions ψ and χ than it does to a higher-ranked proposition ϕ. This seems odd, since probability is additive by definition. However, the disjunctive inference remains valid in probability measures constructed in the way described because Compatibility rules out in advance all measures that fail to assign the higher-ranked proposition ϕ a probability at least as great as the sum of the probabilities of the lower-ranked propositions ψ and χ.

A related restriction leads to further problems with the connection to probabilities. (This is the counterpoint to the problem Kratzer notes for the old order $\succcurlyeq^s_{\mathbf{g}(w)}$ in the quote at the beginning of §3.7.1.) If there is a world w' that is better than all other worlds, Compatibility requires us to assign a probability greater than 0.5 to the unit set $\{w'\}$. If we did not, then we would have an illicit violation of the disjunctive inference: $\{w'\} \succ^s_{\mathbf{g}(w)} \{w''\}$ for all $w'' \neq w' \in \bigcap \mathbf{g}(w)$, and so

$$\{w'\} \succ^N_{\mathbf{g}(w)} \bigcup_{w'' \neq w' \in W} \{w''\}$$

and so the probability of $\{w'\}$ must be greater than the probability of $W - \{w'\}$. In general, if there is a set A of undominated and equally ranked worlds, then for *each* world w'' in A the unit set $\{w''\}$ must have probability greater than $W - A$. If A is large then the amount of probability available to be distributed to non-maximal worlds is tiny (specifically, $\frac{1}{|A|+1}$, a vanishingly small quantity even for moderately sized A).

More generally, if there is a set $\mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w))$ of $\succcurlyeq^N_{\mathbf{g}(w)}$-undominated worlds — that is, if the ordering source is either finite or consistent — Compatibility requires that this set receive probability greater than 0.5.[27] These features of Compatibility place increasingly stronger restrictions on allowable probability measures as the number of propositions in $\mathbf{g}(w)$ increases. Roughly, the amount of probability available to be distributed among worlds decreases exponentially as we move down the ordering $\succcurlyeq^N_{\mathbf{g}(w)}$, subject also to considerations of the size of each equivalence class. When the ordering source is large, a small number of worlds must receive nearly all of the probability mass while an overwhelming majority receive very little.

These constraint that every proposition must have a probability at least as great as the disjunction of all propositions ranked below it is an excessive restriction to place on probability measures. It severely limits the empirical coverage of probability models, including (but certainly not limited to) the fact that we cannot construct empirically adequate models of situations in which our epistemic alternatives are disjoint, as in the baseball example. As far as the semantic application that we have considered is concerned, we will not be able to get the meaning of "as likely as" right even in a probabilistic semantics if we restrict attention to Compatible probability measures, for the same

26 I'll assume that $\mathbf{g}(w)$ is not empty. In models in which it is, Compatibility imposes no constraints at all, and any well-defined probability measure is vacuously consistent. The problems noted here don't arise in such models, but obviously it wouldn't do as a general solution to restrict attention to such models.

27 I am not certain what follows if $\mathbf{g}(w)$ extends upward infinitely. Apparently every $\succcurlyeq^N_{\mathbf{g}(w)}$-equivalence class will receive probability 0, since each equivalence class must be less probable than an infinite number of higher-ranked equivalence classes; but I don't know what follows from this for arbitrary $\phi \subseteq W$.

reason that we could not do so when we attempted to define the meaning of *likely* directly in terms of $\succcurlyeq^N_{\mathbf{g}(w)}$.

Furthermore, probability measures coherent with the newly defined comparative possibility relation will have quite limited applicability for the purposes for which Kratzer introduces this construction, to relate Comparative Possibility to the concept of probability as it is used in scientific practice (and many other domains outside formal semantics). Nearly all probabilistic models used in scientific practice will fail to be consistent with virtually any non-trivial ordering $\succcurlyeq^N_{\mathbf{g}(w)}$. The reason is that, regardless of your philosophy of probability, probabilities are meant to track the relative frequencies of events in one way or another.[28] Relative frequencies are not, however, distributed in the way that Compatibility would require. In particular, the counterpart of the (modified) disjunctive inference clearly does not hold of relative frequencies: an event that is more frequent than each of a set of disjoint alternatives will often turn out to be less frequent than the disjunction of these alternatives. For example, from the fact drunk driving accidents are pairwise more frequent in Texas than in each other U.S. state we cannot conclude that drunk driving accidents are more frequent in Texas than in the whole United States excepting Texas. This mismatch does not mean that there is anything wrong with adopting the construal of probability under discussion, on a technical level; it just means that the probability measures that we encounter elsewhere will almost never be Compatible with an order $\succcurlyeq^N_{\mathbf{g}(w)}$. The scope of this concept of probability is rather limited, then.

The modified comparative possibility relation does not get to the heart of the disjunction problem, and is not compatible with an empirically adequate approach to the measurement of uncertainty in semantics or beyond, at least not by a straightforward rule such as (3.86). The best way to avoid the problems with disjunction and degree modification — and the only way that is theoretically well-motivated, as far as I can see — is the straightforward approach: we have to treat epistemic modals as scalar expressions built around structures at least as rich as fully closed ratio scales. Implementing this generally will require a scalar semantics for the epistemic auxiliary modals as well, to which I now turn.

3.8 Epistemic Auxiliaries Must Be Probabilistic Too

In this section I will consider the epistemic auxiliaries *must*, *might*, *should*, and *ought* in light of the conclusions that we have reached. Even though it is difficult to find direct evidence of gradability in this domain, I will argue that these items should be analyzed as non-gradable scalar expressions which also place conditions on probabilities. The key observation is that, supposing that the epistemic adjectives have a probabilistic semantics as I have argued, attempts to treat *must* as having a quantificational semantics along the lines of the best available quantificational theory (Kratzer's) fail to validate clearly valid inferences generated by the interaction of the epistemic adjectives and auxiliaries. A probabilistic account of *must*, however, gets the logical relations between these items right and is overall much simpler than a hybrid theory with complex machinery designed just for a handful of modal auxiliaries.

On usual assumptions about the relationship between *must* and *might*, this indicates that *might*

---

28 If you're a frequentist, this is a truism; if you're a Bayesian, it holds once you've observed enough data; etc. See e.g. Hacking 2001; Mellor 2005 for discussion of this point.

also expresses conditions on probabilities. The auxiliaries *should* and *ought* are also discussed briefly, and I suggest that the same conclusion is reasonable in their case. I conclude by considering a semantics due to Yalcin (2007) which makes use of quantification over probabilistically defined sets of worlds and does not have the logical problems of the Kratzerian hybrids considered. I argue that Yalcin's proposal can rewritten in thoroughly scalar terms with slightly better empirical coverage, a fact which suggests that the straightforward scalar account proposed here is preferable.

3.8.1 Troubles for the Hybrid Approach

Despite the conclusion of the last several sections, quantificational semantics for the epistemic auxiliaries may still seem attractive because these items do not participate in comparatives or allow degree modification. Although — as I pointed out in ch.1, §1.2 — lack of gradability is not conclusive evidence against scalarity, these facts certainly leave room for doubt that the auxiliaries have a scalar semantics.

(3.87) COMPARISON:
Sam must/should/might (*more than Bill) be at home (*more than Bill).

(3.88) DEGREE MODIFICATION:
Sam must/should/might (*very much) be at home (*very much).

Despite the lack of gradability, though, I'll argue that we still need a probabilistic semantics for these items if we want to explain inferences generated by the interaction of the epistemic auxiliaries and adjectives.

Suppose that we were to adopt a theory on which the epistemic adjectives have a scalar semantics as argued in this chapter, but continue to treat auxiliaries are still quantifiers over $\succcurlyeq_{\mathbf{g}(w)}$-undominated worlds as in Kratzer's account. Naturally, we need to rule out models in which, for example, $must(\phi)$ is true but $prob(\phi) = 0$. In order to avoid rendering possible such monstrous scenarios, a hybrid theory must contain some kind of bridging principles to ensure that the ordering of propositions is at least minimally consistent between the probability measure and the ordering on worlds.

Which bridging rules are worth considering will depend on whether we are using the definition of Comparative Possibility in Kratzer 1981, 1991 or the definition in Kratzer 2012. Consider first the original definition of this relation, which I have been abbreviating as $\succcurlyeq^{s}_{\mathbf{g}(w)}$:

(3.89) $\succcurlyeq^{s}_{\mathbf{g}(w)} =_{df} \{(\phi, \psi) \mid \forall u \in \psi\ \exists v \in \phi : v \succcurlyeq_{\mathbf{g}(w)} u\}$, where $u, v \in \bigcap \mathbf{f}(w)$.

The obvious bridging rule to use would be (3.90) (modeled after the proposal for $\succcurlyeq^{N}_{\mathbf{g}(w)}$ discussed a moment ago).

(3.90) If $\phi \succcurlyeq^{s}_{\mathbf{g}(w)} \psi$ then $prob(\phi) \geq prob(\psi)$.

This rule is a non-starter, though: it would entail, absurdly, that each $\succcurlyeq^{s}_{\mathbf{g}(w)}$-branch B either has zero probability or has a singleton set $\{w_B\}$ which is maximal in that branch, and that $\bigcup_B \{w_B\}$ has probability 1.[29] A slightly better option would be (3.91).

29 Each branch with non-zero probability must have a maximum since each $\succcurlyeq^{s}_{\mathbf{g}(w)}$-equivalence class A is exactly as likely as the infinite union of A with all equivalence classes that A strictly dominates. Since $prob(A) = prob(\bigcup\{A' \mid A \succcurlyeq^{s}_{\mathbf{g}(w)} A'\})$,

(3.91) If $\phi \succ^s_{\mathbf{g}(w)} \psi$ then $prob(\phi) > prob(\psi)$.

(3.91) does not tell us how to assign probabilities to propositions whose maximal worlds fall into the same equivalence class, and so does a bit better. It does, however, restrict *prob* to measures which validate the (disjoint) disjunctive inference. This is fatal (see §§3.3.1 and 3.7).

Furthermore, the constraint in (3.91) does not validate the obviously correct inference in (3.92):

(3.92) If $must(\phi)$ is true and ψ is at least as likely as ϕ, then $must(\psi)$ is true.

Whenever there are worlds in the modal base which satisfy all propositions in the ordering source (i.e. $\bigcap\mathbf{f}(w)\cap\bigcap\mathbf{g}(w) \neq \varnothing$) there will be models consistent with (3.91) in which — for various choices of ϕ and ψ — $must(\phi)$ is true and ψ is as likely as ϕ, but $must(\psi)$ is false. For example, take any ϕ which holds of all the worlds in $\bigcap\mathbf{f}(w)\cap\bigcap\mathbf{g}(w)$, and any ψ which holds of some but not all of them. In this case $must(\phi)$ is true and $must(\psi)$ is false. But since the best worlds in each of these are equally ranked by $\succcurlyeq_{\mathbf{g}(w)}$, we have $\phi \approx^s_{\mathbf{g}(w)} \psi$, and so (3.91) of course admits models in which ϕ and ψ have equal probability, and (3.92) is not valid. ((3.91) even allows that $prob(\psi)$ is **greater** than $prob(\phi)$ here.)

Neither of the bridging rules based on the original version of Comparative Possibility $\succcurlyeq^s_{\mathbf{g}(w)}$ is very promising, then. There are other options that could be considered: for instance, we might think to constrain probability measures directly using the ordering on worlds $\succcurlyeq_{\mathbf{g}(w)}$ rather than with the $\succcurlyeq^s_{\mathbf{g}(w)}$ relation that is defined in terms of it. As far as I can tell, though, this would work only in finite models. In general, the prospects for a hybrid semantics built around $\succcurlyeq^s_{\mathbf{g}(w)}$ do not look good.

Things are not much better if we turn to the new version of Comparative Possibility, $\succcurlyeq^N_{\mathbf{g}(w)}$. If we adopt a bridging rule like the one that Kratzer (2012) suggests (cf. (3.86)), we run into the fatal problems with disjunction again, as discussed at length in §3.7. A semantics for the auxiliary modals built around this idea also fails to validate plainly correct inferences, such as (3.93).

(3.93) If $must(\phi)$ is true then ϕ is much more likely than $\neg\phi$.

For a counter-model, let $\phi = \bigcap\mathbf{f}(w)\cap\bigcap\mathbf{g}(w)$ where this set is not empty, and let $\psi = W - \phi$. Then $must(\phi)$ is true, but the only constraint imposed by Compatibility is that $prob(\phi) > prob(\psi)$. This bridging rule allows any probability measure ones which meets this constraint, including ones in which $prob(\phi) = 0.5000001$ and $prob(\psi) = 0.4999999$. So this bridging rule predicts that there will

it follows from the probability axioms that $prob(A') = 0$ — i.e, the probability of any equivalence class which is strictly dominated by any other equivalence class is 0. Since we can repeat this demonstration for each equivalence class in an infinite branch, the total probability in any infinite branch is zero.

When there is a maximum and it has non-zero probability, the maximum must be a singleton. To see why, imagine that some $\succcurlyeq^s_{\mathbf{g}(w)}$-branch B has a non-singleton maximum $\{w_1, ...w_n\}$. (I'm assuming B is finite, though this does not affect the proof.) By the definition of $\succcurlyeq^s_{\mathbf{g}(w)}$ we have $\{w_1,...w_n\} \approx^s_{\mathbf{g}(w)} \{w_1\} \approx^s_{\mathbf{g}(w)} ... \approx^s_{\mathbf{g}(w)} \{w_n\}$. So by (3.90), $prob(\{w_1,...w_n\}) = prob(\{w_1\}) = ... = prob(\{w_n\})$. From the probability axioms we can infer that all but one of $prob(\{w_1\})...prob(\{w_n\})$ is 0. This contradicts the fact that $prob(\{w_1\})...prob(\{w_n\})$ are all equal to $prob(\{w_1,...w_n\})$. So all branches with non-zero probability have a singleton maximum.

The union of all such singletons has probability 1 because $\{w'\} \approx^s_{\mathbf{g}(w)} \bigcup\{A \mid \{w'\} \succcurlyeq^s_{\mathbf{g}(w)} A\}$, and so $prob(\{w'\}) = prob(\bigcup\{A \mid \{w'\} \succcurlyeq^s_{\mathbf{g}(w)} A\})$. From this it follows that $prob(\bigcup\{A \mid \{w'\} \succcurlyeq^s_{\mathbf{g}(w)} A\} - \{w'\}) = 0$, and so all of the probability mass in every branch is concentrated in the singleton maximum.

be many models in which $must(\phi)$ is true even though ϕ is just barely more likely than its negation, which is absurd.

I am not, of course, in a position to rule out the possibility that a clever defender of quantificational semantics will eventually be able to devise a bridging rule that does not make demonstrably incorrect predictions about the logical relationship between the epistemic adjectives and the auxiliaries. I cannot come up with one, though, and a hybrid theory of this type will not be viable unless someone does. Even if this can be done, we will still want to consider whether there is any advantage of adopting such a theory beyond theoretical inertia. In particular, the Kratzerian approach brings in a good deal of extra logical and pragmatic machinery, and we pay a considerable price in terms of the complexity of the theory if we retain this apparatus exclusively to provide denotations for a handful of modal auxiliaries. If the auxiliaries can be given a simple and straightforward account in the same terms as the adjectives, on the other hand, we have a maximally simple account of the logical relations between the varieties of epistemic modals with no need for stipulated bridging principles to connect the two.

3.8.2 Probabilistic Semantics for Epistemic Auxiliaries

Since the hybrid approach is not very promising, we should consider the possibility that *must*, *should*, and *ought* are also scalar expressions which place conditions on probabilities. If this is right, we do not need complicated extra machinery for the auxiliaries plus bridging rules to connect the two types of epistemic modals. Instead, simple and intuitively correct logical relations between the epistemic adjectives and auxiliaries follow immediately from the definitions.

One such proposal for *must* was made by Swanson (2006), and is reproduced as (3.94).

(3.94) $[\![must\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) \geq 1 - \alpha$ (where $\alpha \geq 0$ is a contextual parameter).

As the reader can verify, none of the problems noted in the previous section involving the relationship between *must* and *likely* arise if we adopt the definition in (3.94) (as long as α is not too large).[30]

If the proposal in (3.94) is right, then — making use of the standard assumption that $might(\phi)$ is equivalent to $\neg(must(\neg\phi))$ — we have a plausible denotation for *might* in (3.95) (also from Swanson 2006).

(3.95) $[\![might\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) > \alpha$

Note that this is very close to the meaning of *possible* in the positive form which was proposed earlier, and equivalent if $\alpha = 0$.

Should and *ought* seem to be of intermediate strength, as discussed by von Fintel & Iatridou (2008) and in chapters 5 and 6 below: for example, (3.96a) seems OK, but (3.96b) is quite strange.

30 Two quick notes about possible variants of this proposal. First, von Fintel & Gillies (2010) argue against Kratzer (1991) that *must* is not weak. If this is right, we can incorporate it easily: just set α to zero, so that $must(\phi)$ is equivalent to ϕ *is certain*. Another possible modification suggested by von Fintel & Gillies (2010) is to treat epistemic *must* as an evidential indicating an inferential or otherwise indirect source of information. It is not obvious precisely how to incorporate this insight into the present framework, in part because there is currently no widely accepted semantics for evidentiality. However, one promising approach which could readily be integrated with the present approach is the probabilistic theory of evidentiality due to Davis, Potts & Speas (2007). A more detailed consideration of this issue will have to wait for another occasion, though.

(3.96) a. Mary ought to be at home, but I guess I can't say she must be.

b. # Mary must be at home, but I guess I can't say she ought to be.

Indeed these have a similar flavor to (3.97), which the semantics we have from §3.4 accounts for:

(3.97) a. It is likely that Mary is at home, but I guess I can't say it is certain.

b. # It is certain that Mary is at home, but I guess I can't say it is likely.

One way to capture these facts is to suppose that *should* and *ought* are just the verbal equivalents of the relative adjectives *likely* and *probable*: in effect, they are relative-standard modal verbs.

(3.98) $[\![should/ought\ \phi]\!]^{\mathcal{M},w,g} = 1$ iff $prob(\phi) > \theta_{epistemic}$, where $\theta_{epistemic}$ is the same threshold that controls the interpretation of *likely* and *probable*.

(3.98) seems plausible, but may well be incorrect in the details; what motivates it is primarily the intuition that these expressions are used appropriately in the same situations in which *likely* and *probable* are appropriate, and that all of them are neg-raisers, which indicates that they occupy a similar mid-range position in their scale (cf. Horn 1989 and discussion in ch.5-6 below).

3.8.3 Continuous Sample Spaces and Approximation

Yalcin (2007: 1015-7) points out a problem which affects both the proposed semantics for *possible* and *might* and the issue of whether we can dispense with quantification altogether in our semantics for epistemic modals. Suppose that we are dealing with some continuous space, e.g. the length R in feet of Sam's next attempt at the standing long jump. This is a continuous variable, and the value of R can in principle be any positive real number — an uncountably infinite range. Unless our ability to predict Sam's jumping ability is extraordinarily precise, then, the probability that $R = r$ will be zero for any particular real number r. Nevertheless, it is clear that there is a fair range of values between, say, 2-7 feet — uncountably many, in fact — for which it true to say that it is possible that Sam will jump that far, or that Sam might jump that far. So there must be events which are possible and yet have probability zero; as a result, "it would be a mistake to collapse epistemic possibility with nonzero probability" (Yalcin 2007: 1016).

This is a serious worry, of course, and seems to suggest that the truth-conditions proposed for *possible* earlier in this chapter were too strict. Yalcin (2007) proposes a solution which is relevant in addition since it provides an interesting way to combine scalar and quantificational semantics for epistemic modals while avoiding the problems just noted for Kratzer's approach. On Yalcin's account, we start by partitioning logical space into a set of alternatives Π (the "modal resolution" of Yalcin 2011). We then single out a subset π of Π, which is distinguished by the fact that all of the probability mass is concentrated in this subset. Finally, we define *possible*, *might*, and *must* as quantifiers over cells ι of π (*might* and *possible* have the same meaning):

(3.99) a. $[\![must(\phi)]\!]^{\mathcal{M},w,g} = 1$ iff $\forall \iota \in \pi : \forall w \in \iota : [\![\phi]\!]^{\mathcal{M},w,g} = 1$

b. $[\![might(\phi)]\!]^{\mathcal{M},w,g} = 1$ iff $\exists \iota \in \pi : \forall w \in \iota : [\![\phi]\!]^{\mathcal{M},w,g} = 1$

c. $[\![\phi\ is\ possible]\!]^{\mathcal{M},w,g} = 1$ iff $\exists \iota \in \pi : \forall w \in \iota : [\![\phi]\!]^{\mathcal{M},w,g} = 1$

On this proposal, we can make sense of the judgment that, for instance, *It is possible that Sam will jump exactly 3.4454789458743598 feet* is true: all that is required is that there be some cell $\iota \in \pi$ such that, for every world in ι, Sam jumps exactly 3.4454789458743598 feet. This can be true even if $prob(\iota) = 0$.

This is a quantificational semantics of sorts, although I don't expect that this will be much comfort to defenders of traditional approaches to modality: the quantifiers range over sets of worlds rather than worlds, and the sets in question are crucially defined using probability, a scalar concept. Even if this is the right approach, then, we will end up affirming the importance of scalar concepts and eliminating reference to direct quantification over possible worlds in our semantics for epistemic modals. The question that I am interested in, though, is whether it is possible to achieve the effect of this analysis — or even improve on it — without bringing in quantification even over cells of a partition. The reason for wishing to do so is that, if Yalcin is correct in saying that *possible* is an existential quantifier over possible worlds, it is not clear how to explain the data from degree modification, entailments, and implicatures which motivated us to classify *possible* as a minimum-standard gradable adjective in §3.4.

The response to Yalcin's objection, I think, is that we must take into account the fact that natural language expressions are frequently given APPROXIMATE interpretations. To take a simple case, a sentence like *Sam is 5 feet 10 inches tall* will rarely be taken to pick out a precise value — if it did, it would almost certainly be false, or at best true by some massively improbable accident. Even in the long jump example, where for competitive reasons we are uncommonly concerned with precise measurement, the measurements reported are never meant to pick out a particular real number of inches, but rather some number of inches or centimeters plus or minus some granularity or margin for error.

The semantics and pragmatics of approximate interpretation has been considered among others by Lewis (1979); Krifka (2007a); Sauerland & Stateva (2007); Bastiaanse (2011). Without wanting to go into the issues surrounding approximate interpretation in detail, I will note a few fairly obvious facts here. The granularity of interpretation can vary for any number of reasons, linguistic or non-linguistic, and seems to depend on the purposes of the conversation and various other factors. Most notably for us, the granularity will generally be responsive to the actual value reported: for example, *5 feet tall* is interpreted more generously than *5.3 feet tall* (or for that matter *5.0 feet tall*).

The ubiquity of approximate interpretation suggests that natural language expressions of degree are rarely, perhaps never, interpreted with anything like the precision of the real number system; instead, they are generally interpreted as picking out a value d with some contextually variable granularity g. If this is right — as others have argued before me — the objection that Yalcin brings up can be addressed without bringing in quantification over cells of partitions: as long as the granularity g is greater than zero, the probability that Sam will jump $r \pm g$ feet will be

$$\int_a^b prob(\textit{Sam jumps r meters})\,\mathrm{d}r$$

where $a = r - g$ and $b = r + g$, a non-zero value for many reasonable probability measures and choices of r and g.

Although Yalcin's objection is correct as it goes, then, it probably does not apply to natural language as it is actually used. This means that we can retain the simple semantics for *possible* that was proposed on analogy with Kennedy & McNally's (2005) semantics for *dangerous*, *bent*, etc.: all modals are scalar expressions, and possibility is non-zero probability. In addition to explaining the evidence that *possible* is a minimum-standard gradable adjective, this account receives independent motivation from general features of approximate interpretation.[31]

3.9 Confidence, Probability, and Question-Embedding *Certain*

In a late-breaking development, Klecha (2011) brings up the following pair of examples in reply to Lassiter (2010a) (an early version of the arguments given in §§3.4-3.6 of this chapter).

(3.100) a. Obama's reelection couldn't be less certain.

b. Obama's reelection couldn't be less likely.

Klecha notes that, if *couldn't be less A* has the obvious interpretation as a degree minimizer, these two sentences are expected to mean the same thing on the theory argued for here. It seems clear that they do not, though: for example, if there is a 50% chance that Obama will be reelected, (3.100b) is clearly false but (3.100a) would seem to be true. These data problematize the argument given above and in Lassiter (2010a) that *certain* is associated with a probability scale.

This example is very suggestive, and I want to discuss two possible accounts briefly here. The less interesting (but perhaps correct) response is to note that *couldn't be less A* does not always signal a minimum degree of property *A*. Naturally occurring examples of this phrase frequently involve properties that do not appear to have minima, e.g. "they couldn't be less friendly/interested/excited about work" (and many more that are easily found on Google). Similarly a *USA Today* reviewer writes that the 2007 film version of *Beowulf* "couldn't be less faithful to the original epic poem" (11-15-07). (I presume that the reviewer didn't mean to suggest, for example, that the *Beowulf* film had zero plot elements in common with the original.) These are adjectives which intuitively have no minima, an intuition corroborated by the standard tests — and necessitated by the fact that they are relative adjectives, for those who endorse Interpretive Economy. If *couldn't be less A* could only be interpreted as a degree minimizer, then, these uses would not be felicitous. In light of this data, it seems that *couldn't be less A* is in at least some uses an idiom meaning "is far from being *A*" or "is clearly not *A*". If so, the arguments for interpreting *certain* as a maximum-standard adjective on a scale of probability are in the clear: it would be completely unreasonable to describe an event with a 50% probability of occurring as "certain" on my account, but it's much less of a stretch to call the same event "likely".

A more interesting — but quite speculative — way to handle this issue is to develop Klecha's suggestion that, in the example at hand, *certain* is not associated with the probability scale but with a scale of confidence. It turns out that there is a well-motivated way to measure confidence probabilistically using the information-theoretic notion of **surprisal** (also known as **self-information**).[32]

31 See however Yalcin (2011) for arguments that the concept of modal resolution has independent motivation as well. My proposal is not really in conflict with this idea, by the way: there are other reasons to think that this is a useful concept, but I don't think that the issue about continuous spaces is decisive either in favor of Yalcin's (2007) account of epistemic modals or against the theory that I have given.

32 Self-information and entropy were originally defined by Shannon (1948), and information-theoretic ideas have been extremely influential in many fields; for introductory presentations and discussion of many applications see Cover &

As you might surmise, the surprisal of a proposition ϕ is a measure of how surprised we would be to learn that the actual world is in ϕ: the less probable it is that ϕ holds, the more surprised we will be if we subsequently learn that it is, and the more information we will gain if this occurs. Formally surprisal I is defined as negative log probability, i.e. the log of the reciprocal of the probability of a proposition:

$$I(\phi) = \log_2 \frac{1}{prob(\phi)} = -\log_2 prob(\phi).$$

(The choice of base 2 for the logarithm is arbitrary, but standard.)

Surprisal is a measure of uncertainty, rather than certainty: greater probability leads to lower I. To get a measure of certainty we have to invert the scale, so that we end up with the positive log probability of a proposition. On this account *certain* would be defined as:

(3.101) $[\![certain]\!]^{\mathcal{M},w,g} = \lambda p_{\langle s,t\rangle}[-I(p)] = \lambda p_{\langle s,t\rangle}[\log_2 prob(p)]$

One useful property of this measure is that it has the same upper bound as the probability scale: when a proposition has probability 1, it has the maximum possible amount of certainty under this definition, i.e. the minimum possible amount of surprisal: $\log_2(1) = 0$. If ϕ is certain it has probability 1, then, just as we argued above. This would even have the benefit of explaining the entailments and implicatures between *possible, likely*, and *certain* noted in §3.4.3.1 of this chapter.

This way of thinking about certainty has problems in light of the broader theory of degree modification, however. As we noted, the availability of ratio and proportional modifiers (e.g. *half/95%*) with *certain* requires that the scale have both and maximum and a minimum, and that the measure be additive. If certainty is measured as log probability, though, it is not additive and does not have a minimum: $\log_2(0) = -\infty$. As a result, (3.101) would lead us to expect that many of the degree modifiers that we saw earlier in this chapter should not be possible.

Perhaps we have not been looking far enough afield, though. English uses *certain* and *uncertain* to embed not only propositions (*It is certain that ϕ*) but also overt and concealed questions:

(3.102) a. It is certain which horse will win the race.

b. The winner is certain.

This is a point of difference with the other epistemic adjectives discussed here, none of which can embed questions.

(3.103) * It is possible/probable/likely which horse will win the race.

Ideally we would like to have a semantics for *certain* which makes sense of both question- and proposition-embedding *certain* in the same terms.

Questions are often treated as denoting partitions of W (Groenendijk & Stokhof 1984). As it happens, there is also a standard way to measure confidence over partitions, known in information theory as **entropy**. Let $\mathcal{A} = \{A_1, ...A_n\}$ be a partition of W. We can define the entropy H of partition $\mathcal{A}$ as the expected surprisal of the cells of $\mathcal{A}$:

$$H(\mathcal{A}) = \mathbb{E}_{A\in\mathcal{A}}[I(A)] = \sum_{i=1}^{n} prob(A_i) \times I(A_i) = -\sum_{i=1}^{n} prob(A_i) \times \log_2 prob(A_i),$$

Thomas (1991); MacKay (2003). See also van Rooij (2003, 2004) for some interesting applications of information theory to the semantics and pragmatics of questions.

where $I(A_i)$ is the surprisal of A_i as defined above and the expectation $\mathbb{E}(\cdot)$ of a function is a probability-weighted average (cf. ch.6, §6.3).

The entropy H of a probability distribution is a measure of how dispersed the distribution is; translated into the terminology of intensional semantics, this means that the larger $H(\mathcal{A})$ is, the less confident we are in our guess about where in $\mathcal{A}$ the actual world lies. As with surprisal, we will need to consider negative entropy, since greater entropy means a more dispersed distribution, and so lower overall confidence with respect to that partition. We can treat question-embedding *certain* as a maximum-standard adjective with the following definition:

(3.104) $[\![certain]\!]^{\mathcal{M},w,g} = \lambda \mathcal{Q}_{\langle st,t\rangle}[-H(\mathcal{Q})] = \lambda \mathcal{Q}_{\langle st,t\rangle}\left[\sum_{p\in\mathcal{Q}} prob(p) \times \log_2 prob(p)\right]$

On this interpretation, certainty is upper-bounded because entropy is lower-bounded. The upper bound corresponds to the minimum possible value of H, i.e. the minimum possible uncertainty. If there is some cell p of the partition $\mathcal{Q}$ such that $prob(p) = 1$, then $prob(p') = 0$ for all other cells $p' \in \mathcal{Q}$, and so

$$H(\mathcal{Q}) = -1 \times \log_2(1) - \left[\sum_{p' \neq p \in \mathcal{Q}} 0 \times \log_2(0)\right] = -1 \times 0 - 0 = 0$$

(where $\log_2(0)$ is defined to be 0, as usual in information theory). 0 is the minimum entropy that a distribution can have, and is only seen in case there is some cell p in the partition for which $prob(p) = 1$.

The interpretation of confidence as negative entropy gives us reasonable truth-conditions for the examples in (3.102), as well as Klecha's example (3.100a). First consider (3.102a). Intuitively, *It is certain which horse will win* tells us that there is some horse x for which it is certain that x will win, but doesn't indicate which horse this is. Let $\mathcal{Q} = [\![which\ horse\ will\ win]\!]^{\mathcal{M},w,g}$. This comes down to a partition $\{\{w \mid x\ wins\ in\ w\} \mid x \in X\}$, where X is the set of horses in the race. The truth-conditions predicted by definition (3.104) are:

(3.105) $[\![It\ is\ certain\ which\ horse\ will\ win\ the\ race]\!]^{\mathcal{M},w,g} = 1$ iff $\mathbf{pos}([\![certain]\!]^{\mathcal{M},w,g})(\mathcal{Q})$

$= 1$ iff $[\![certain]\!]^{\mathcal{M},w,g}(\mathcal{Q}) = \mathbf{max}(\mathbf{D}_{certain})$

$= 1$ iff $\sum_{p\in\mathcal{Q}} prob(p) \times \log_2 prob(p) = 0$

$= 1$ iff $\exists p' \in \mathcal{Q} : prob(p') = 1$

This derives the intuitively correct meaning for this example: (3.102a) is true just in case there is some partition p of $\mathcal{Q}$ such that $prob(p) = 1$; this is the same as saying that there is a horse x such that the probability that x will win is 1.

This treatment extends immediately to a slight variant of Klecha's key example.

(3.106) It couldn't be less certain whether Obama will be reelected.

Note first that confidence as defined here is lower-bounded once a partition is fixed. The minimum possible amount of confidence we could have regarding a particular question is what we have when we have absolutely no information about which cell of the partition the actual world is in. This is the point of **maximum entropy**, where probability is spread equally throughout the partition: $prob(p) = \frac{1}{|\mathcal{Q}|}$ for each $p \in \mathcal{Q}$. In this case, each possible answer to the question at hand is equally likely.

The embedded question *whether Obama will be reelected* denotes a partition $\mathcal{Q} = \{\{w \mid \textit{Obama is reelected in } w\}, \{w \mid \textit{Obama is not reelected in } w\}\}$. Let's allow that *couldn't be less A* is a degree minimizer in this instance. The interpretation of (3.106) is now

(3.107) $[\![(3.106)]\!]^{\mathcal{M},w,g} = 1 \text{ iff } [\![certain]\!]^{\mathcal{M},w,g}(\mathcal{Q}) = \mathbf{min}(\mathbf{D}_{certain})$

$$= 1 \text{ iff } \forall p \in \mathcal{Q} : prob(p) = \frac{1}{|\mathcal{Q}|} = \frac{1}{2}$$

The derivation relies on an implicit assumption that the domain $\mathbf{D}_{certain}$ is relativized to the question $\mathcal{Q}$ at hand: we don't look for the minimum certainty (maximum entropy) that any distribution on *any* partition could have (no matter how large), but the minimum amount of certainty that any distribution on this partition could have, $\log_2(\frac{1}{|\mathcal{Q}|})$. (This probably would not be acceptable as a general assumption about adjective semantics, but it makes sense here given the nature of the example: *this* partition *couldn't* be less certain, given the question that it reflects.) The analysis derives intuitively correct truth-conditions for (3.106) — it is true just in case there are even odds on Obama's reelection. I won't try to deal with Klecha's original example where *certain* modifies the noun *reelection* here, since this would force us to consider the semantics of event nominals in more detail than seems necessary. However, it seems likely that this case will yield to the same analysis as (3.106): *Obama's reelection couldn't be less certain* is equivalent to *It couldn't be less certain whether Obama will be reelected*, and negative entropy gives us a reasonable account of the latter.

None of this tells us how to interpret *certain* when it embeds a proposition rather than a question, though. There may be a way to use information-theoretic notions to get a definition of confidence for propositions that is more closely related to the entropy-based formulation in (3.104), but it is not obvious what this would be: negative suprisal won't work, for instance, because it is not lower-bounded or additive, and so would render ratio and proportional modifiers meaningless. My best guess for the moment is that *certain* simply has different meanings depending on whether it embeds a question or a proposition. ϕ *is certain* maps ϕ to a probability scale and checks whether it occupies the maximum point of this scale; $\mathcal{Q}$ *is certain*, on the other hand, maps $\mathcal{Q}$ to a confidence scale and checks whether it occupies the maximum of this scale. As it happens, the confidence scale is defined in such a way that a positive answer to the latter question entails that there is some proposition in $\mathcal{Q}$ for which a positive answer to the former question is in order. They are, nevertheless, simply different scales.

Returning to the main point of this chapter, it should be clear that however the detailed issues surrounding *certain* ultimately pan out, the arguments for associating *likely, probable*, and *possible* with a probability scale are still compelling.[33] In fact, the results of this section — and in particular

33 Klecha (2011) tries to cleave off *possible* as well, arguing that this adjective is not gradable at all. The judgments reported are questionable, though, and do not reflect the actual usage of English speakers. He doesn't address several other crucial points, in particular the fact that *impossible* is clearly a maximum-standard gradable adjective, which on standard assumptions entails that *possible* is a minimum-standard gradable adjective.

In any case, we could concede all of this and it would do nothing to save Kennedy's (2007) Interpretive Economy constraint (one of Klecha's goals): the probability scale has inherent maximum and minimum elements but *likely* and the ease with which we were able to give a plausible scalar semantics for question-embedding *certain* in probabilistic terms — serve to underline the main point of Lassiter (2010a) and this chapter: epistemic modals are fundamentally scalar items, not quantifiers over possible worlds; and probability plays a crucial role in their semantics.

3.10 Epistemic Conditionals and Conditional Probability

To close the chapter, I want to say a few quick words about the way that the scalar semantics interacts with Kratzer's (1986) restrictor account of conditionals, as reviewed and recast slightly in ch.1, §1.6. The short version is that, just as $M(\psi)$ on its own (M = any epistemic modal) is true iff the probability of ψ is at least as high as the appropriate threshold θ_M, the conditional *If* ϕ *then* $M(\psi)$ will be true just in case the **conditional** probability $prob(\psi|\phi)$ is at least θ_M. This is, by all indications, the right result. Some details follow. (Since conditionals are mostly a side concern here I will simply describe the results in general terms, referring readers to the literature on qualitative probability, e.g. Krantz et al. 1971; Fishburn 1986; Narens 2007, for more detail.)

The scale $\mathcal{S}_{epistemic}$ discussed in earlier sections is built around an order on propositions, rather than worlds. To some extent this feature was a necessary result of our method: the structure of this scale was inferred from the acceptability of various linguistic expressions of possibility, certainty, likelihood, and so on, and the things to which we apply these expressions are typically propositions. Arguably events, states, and actions can also have these properties as well, but it does not look as if English has the resources to talk directly about the degree of possibility of an individual world. Nevertheless, it is possible to build the epistemic scale in an equivalent way by taking worlds to be basic, and doing so is useful for several reasons: it makes for a simple and explanatory approach to conditionals, and it also sets up some features of the semantics that I will propose in chapter 6 for deontic modals and desire verbs.

Let $\succcurlyeq^W_{epistemic}$ be a weak order on W, where the superscript ^{W} reminds us that this is an order over worlds rather than propositions. (We simplify by assuming that W is finite.) A scale with only these elements would be an ordinal scale which carries no quantitative information; in particular it would not determine a unique assignment of probabilities to (unit sets of) worlds. To extend this into something capable of determining a probability measure, we need to consider a fully closed ratio scale $\mathcal{S}^W_{epistemic} = \langle W, \succcurlyeq^W_{epistemic}, \circ, \bot, \top \rangle$ which satisfies the usual axioms (cf. ch.2, (2.41)). This looks a bit strange initially — what is the concatenation of two worlds? — but it is isomorphic to a scale of propositions which looks much more familiar: the fully closed ratio scale $\mathcal{S}^P_{epistemic} = \langle \Phi, \succcurlyeq^P_{epistemic}, \circ, \bot, \top \rangle$ discussed in detail earlier in this chapter. Here $\Phi = \mathcal{P}(W)$, $\circ = \cup$, $\bot = \varnothing$, $\top = W$, and the superscript ^{P} reminds us that we are dealing with an order on propositions now. The latter scale is, as we noted in §3.5.2, isomorphic to a finitely additive probability measure.

The translation between $\mathcal{S}^W_{epistemic}$ and $\mathcal{S}^P_{epistemic}$ is straightforward, and relies on ensuring that unit sets of worlds are ordered in the same way that their constituent worlds are and that concatenation is associated with set union. Note that the world variables range over both singular and plural worlds here.

probable are neither maximum- nor minimum-standard adjectives, and so we still have two clear counter-examples to Interpretive Economy.

(3.108) **Consistency**: let $ATOMS(x)$ be the set of atomic subparts of x, i.e. $\{y \mid y \sqsubseteq x \wedge \forall z : z \sqsubseteq y \rightarrow z = y\}$, where $\sqsubseteq$ is the part-of relation on worlds corresponding to Link's (1983) part-of relation on individuals. Then an order on (plural) worlds $\succcurlyeq^W$ and an order on sets of worlds $\succcurlyeq^P$ are *consistent* just in case

$$\forall w, w' \in W : w \succcurlyeq^W w' \text{ iff } ATOMS(w) \succcurlyeq^P ATOMS(w').$$

If $\succcurlyeq^W_{epistemic}$ and $\succcurlyeq^P_{epistemic}$ are consistent, the scale of worlds $\mathcal{S}^W_{epistemic}$ and the scale of propositions $\mathcal{S}^P_{epistemic}$ are equivalent, with plural worlds interpreted as propositions in disguise.

The advantage of going this route is that we can apply Kratzer's (1986) restrictor analysis of conditionals, slightly modified along lines proposed in ch.1, §1.6, directly to $\mathcal{S}^W_{epistemic}$. Recall that the restriction $\succcurlyeq \restriction A$ of a binary order to a set A is defined as the order $\succcurlyeq' = \{(x,y) \mid (x,y) \in \succcurlyeq \wedge x \in A \wedge y \in A\}$. As I showed there, Kratzer's account is equivalent to taking the antecedent of a conditional to restrict a binary order which determines the interpretation of modals in the consequent. Letting $\mathbf{h}$ be a function from worlds to a binary order appropriate for the theory at hand,

(3.109) $[\![\textit{If } \phi \textit{ then } \psi]\!]^{\mathcal{M},w,g}_{\mathbf{h}} = [\![\psi]\!]^{\mathcal{M},w,g}_{\mathbf{h}'}$
where, for all w, $\mathbf{h}'(w) =_{df} \mathbf{h}(w) \restriction \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$.

According to the theory that I have proposed $\mathbf{h}(w)$ gives us a binary order satisfying the conditions that we have placed on $\succcurlyeq^W_{epistemic}$, a fully closed ratio scale which is equivalent to a finitely additive probability measure. Call this measure $prob_W(\cdot)$. The order $\mathbf{h}(w) \restriction \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$ is the restriction of $\succcurlyeq^W_{epistemic}$ to the set of worlds in which the antecedent holds. Once all of the atoms which fail to satisfy ϕ are removed, all plural worlds which have them as subparts are removed as well (as a consequence of the way that concatenation is defined). In the equivalent order on propositions, this corresponds to removing all propositions which do not satisfy ϕ throughout. Note that the operation of restriction cannot have any effect on the order of propositions; all it can do is to ensure that particular comparisons in the unrestricted order are not found in the restricted order.[34]

The restricted order is, then, equivalent to a restricted scale $\mathcal{S}^V_{epistemic} = \langle V, \succcurlyeq^V_{epistemic}, \circ, \bot, \top \rangle$, where the domain W has been replaced with the set $V = \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$ and the order $\succcurlyeq^V_{epistemic}$ is consistent with the original order $\succcurlyeq^W_{epistemic}$. Since all of the worlds in V are ϕ-worlds, the latter is also equivalent to a fully closed ratio scale on propositions, this time $\mathcal{S}^Q_{epistemic} = \langle \Psi, \succcurlyeq^Q_{epistemic}, \circ, \bot, \top \rangle$, where $\Psi = \mathcal{P}(V)$, $\top = V$, etc. Again, the only propositions in the domain of the scale are ones for which ϕ holds throughout, since V is just the set of ϕ-worlds.

On this construction, for all propositions χ, χ' which remain after application of the restriction operator $\restriction$, the ordering of χ and χ' in $\succcurlyeq^Q_{epistemic}$ is the same that it was in $\succcurlyeq^P_{epistemic}$. Since this is a fully closed ratio scale, all unique admissible measure functions on $\mathcal{S}^Q_{epistemic}$ will of course be isomorphic to a unique finitely additive probability measure $prob_V(\cdot)$, where $prob_V(\phi) = 1$ since

34 There is one small adjustment that we also have to make to the semantics from ch.1: the restriction operator needs to apply to the entire scale, rather than just the binary order, so that it can also reset the top element $\top$ to be the concatenation of all worlds that remain after the order is restricted, i.e. $\{w' \mid [\![\phi]\!]^{\mathcal{M},w',g} = 1\}$. Making this adjustment quite generally would not affect anything else, including in Kratzer's semantics.

ϕ holds throughout V. This is enough to guarantee that $prob_V(\cdot)$ is the same as the conditional probability measure $prob_W(\cdot|\phi)$.

The prediction of the restrictor analysis as applied to our semantics, then, is that epistemic modals in the consequent of a conditional should be interpreted as placing restrictions on the conditional probability of the consequent given the antecedent. As noted by Kratzer (1986); Egré & Cozic (2011) and others, this seems to be correct. In Grice's (1989) example (discussed in some detail by Kratzer), (3.110) is clearly interpreted in this way rather than — as other analyses of conditionals might lead us to expect — as an unconditional probability statement which is claimed to be true if the antecedent is.

(3.110) If Zog had white, there is an $\frac{8}{9}$ probability that he won.

a. ✓ $prob(Zog\ won|Zog\ had\ white) \geq \frac{8}{9}$

b. # $Zog\ had\ white \rightarrow [prob(Zog\ won) \geq \frac{8}{9}]$

This account makes similar predictions for the other epistemic modals discussed here as long as these expressions are scalar and are associated with a probability scale.

3.11 Conclusion

The fact that many expressions of epistemic modality are gradable does not seem at first glance to be a serious problem. The standard theory of modality due to Kratzer resembles the RTM semantics for gradability and comparison developed in the last chapter being built around a binary order, and indeed it is not difficult to extract a semantics for gradable modals from Kratzer's theory using the usual RTM methods. However, §3.2 of this chapter showed that this theory does not give us what we need. A theory faithful to Kratzer's semantics predicts that many or most epistemic comparatives should be undefined; it validates clearly incorrect inferences involving disjunction; it fails to validate correct inferences between epistemic comparatives, *must*, and *might*; and it does not give stable truth-conditions to sentences with ratio and proportional modifiers used with epistemic adjectives.

By starting from the other direction — examining the data involving degree modification in light of the general semantics for gradability developed in earlier chapters — we arrived at the conclusion that the epistemic adjectives *possible*, *probable*, *likely*, and *certain* are all associated with a scale $\mathcal{S}_{epistemic}$ which is fully closed and additive. A mathematical result due to Narens (2007) shows that, if $\mathcal{S}_{epistemic}$ has this structure, it is isomorphic to a well-known type of numerical probability. This indicates — as Swanson (2006); Yalcin (2007, 2010) and others have claimed before on somewhat different grounds — that the semantics of epistemic modality is based on numerical probability or something closely related to it. Happily, this approach also avoids the three logical problems which affect Kratzer's theory noted in §3.2, and makes reasonable and intuitively correct predictions regarding disjunction, degree modification, and comparability.

In §3.8 we considered whether the epistemic auxiliaries also have a semantics built around the scale $\mathcal{S}_{epistemic}$. The lack of degree modification and comparison with these items seem to suggest a negative answer; however, I showed that an attempt to build a hybrid theory where auxiliaries are quantificational but adjectives are probabilistic leads to unacceptable predictions. I also considered apparent counter-examples involving question-embedding *certain* and *possible* with continuous sample spaces, showing that these issues also receive a straightforward account in the proposed theory. The unified probabilistic semantics for epistemic modals that I propose treats where the adjectives and auxiliaries alike as scalar expression, even though the auxiliaries are not gradable — the maximally parsimonious theory as well as the most empirically adequate, it seems. To be sure, the conclusions of this chapter leave many questions about epistemic modals undecided; however, they shed considerable light on the structural aspects of this domain, a fundamental issue for many purposes.

In the next chapter we will take some more small steps in this direction by considering a problem from the psychology of reasoning literature involving *likely*, *probable*, and some related expressions of epistemic uncertainty. Experimental subjects' judgments of the likelihood of events often show sensitivity to the distribution of alternatives, a fact which has been taken as evidence that people are not capable of reasoning probabilistically. Building on Yalcin (2010), I will offer an alternative semantic account which fits together with the probabilistic semantics developed in this chapter and with the theory of gradability more generally. This account has two useful features: it clarifies how the context-sensitive standard for the relative adjectives *likely* and *probable* is set; and it allows us to explain the results of reasoning experiments without ascribing massive cognitive error to experimental subjects.

CHAPTER 4

Setting the Standard: *Probable*, Alternatives, and Rationality

4.1 Probability in the Philosophy and Psychology of Reasoning

Linguists and philosophers have built theories of modality largely around data from the auxiliary modals — *might, must, can, should*, for example — neglecting adjectival modals to a large extent. At the same time, modal logic and its extensions (notably the dominant theory due to Kratzer) treat all modals as restricted quantifiers over possible worlds. In general, authors in this tradition have not considered the possibility that modals are scalar expressions rather than quantifiers, and in particular have not treated probability as a serious contender to undergird the semantics of epistemic modality. As we will see in Chapters 5-6, the same holds for expressions of desire and obligation, which have also been assumed to be quantifiers over possible worlds — wrongly, I will argue.

The story in psychology is quite different, although the net result has been the same: a neglect of probability. Two differences are particularly notable for us. First, psychological work on uncertainty has not ignored the modal adjectives, but has made them the main object of investigation. The two words *likely* and *probable* have been the focus of an inordinate amount of work in the psychology of reasoning, some of which we will consider in the present chapter.

Second, psychologists' attempts to grapple with uncertainty have not ignored numerical probability, but have taken it as the starting point of the entire enterprise. In particular, a great deal of psychological work on reasoning has assumed that probability provides the unique normatively correct framework for reasoning about uncertainty. Support for this assumption comes from a large body of work in philosophy and logic showing that, on the assumption that beliefs come in degrees, only assignments of degrees of belief that conform to the probability calculus are guaranteed to be consistent. Against this background, apparent deviations from the probability calculus in experimental settings appear not only as evidence that subjects do not reason probabilistically, but also as evidence that subjects' reasoning is mistaken and defective.

Such evidence has been provided, in spectacular form, by the "Heuristics and Biases" tradition inaugurated in the early 1970's by the psychologists Daniel Kahneman and Amos Tversky (see the collection of papers in Kahneman, Slovic & Tversky 1982). Prior to this work, it was widely assumed that humans do reason probabilistically, more or less as Laplace famously claimed as far back as 1814:

> We see in this essay that the theory of probability is basically nothing but good sense reduced to calculation; it allows us to assess with precision that which clear minds feel by a sort of instinct, without often being able to recognize it. (Laplace 1829)

Before the 1970's there were slight caveats regarding these assumptions, for example, results suggesting that subjects are sometimes too conservative in Bayesian updating (e.g., Edwards 1968; see Gigerenzer 2000 for discussion). Soon after the publication of Kahneman & Tversky's early work, the current of opinion had shifted dramatically. Many psychologists working on reasoning since would agree with the sentiment of Slovic, Fischhoff & Lichtenstein (1976):

> It may be argued that we have not had the opportunity to evolve an intellect capable of dealing conceptually with uncertainty.

This kind of skepticism about ordinary subjects' reasoning abilities has characterized dominant trends in reasoning from the mid-1970's until the present day. The conclusion that humans cannot and do not reason probabilistically has been highly influential, not only in psychology but also in economics — Daniel Kahneman won the Nobel Prize in Economics in 2002 — and in other related fields as well as popular science (Gould 1992; Piatelli-Palmarini 1994). More recently, probabilistic models have enjoyed greater prominence in a number of areas of cognitive psychology (see Chater, Tenenbaum & Yuille 2006; Griffiths, Kemp & Tenenbaum 2008 for overviews). However, the results of many experiments which have been interpreted as showing that people are incapable of reasoning coherently about probability have not been met head-on in many cases. This chapter addresses one such case, distinguished by the fact that it is informative about both reasoning and about the semantics of modality: the alternative-sensitivity of expressions of epistemic modality such as *likely* and *probable*.

4.2 Overview of Chapter 4

If it is true that humans are incapable of probabilistic reasoning, this is a problem for the conclusions of Chapter 3. There I argued, on the basis of a variety of linguistic tests, that the semantics of epistemic modals is built on a scale which is isomorphic to (at least finitely additive) numerical probability. However, on the assumption that speakers of a language must have the cognitive resources to reason about the meanings of expressions in their language, this is in conflict with one of the leading conclusions of the Heuristics & Biases literature.

In this chapter we will consider experimental results showing that subjects' judgments about probability statements are sensitive to contextual alternatives. These results have been taken to support the conclusion that humans do not represent and reason about uncertainty probabilistically. I will argue that this conclusions are motivated by mistaken semantic assumptions: the results of the experiments are compatible with a probabilistic semantics for the crucial test items, which I will provide and motivate on independent linguistic grounds, building on discussion in Kennedy (2007); Beaver & Clark (2008); Yalcin (2009, 2010) and the results of the last chapter. The key observation is that *likely* and *probable* are relative-standard adjectives, a class of expressions which are grammatically sensitive to COMPARISON CLASSES. In the case of proposition-embedding adjectives, the comparison class is a set of propositional alternatives which can (but need not) be determined by the placement of FOCUS.

The conclusion will not be that we have direct psychological evidence that humans **do** represent uncertainty probabilistically, although the linguistic evidence provided in chapter 3 and results that I will establish here may be taken, less directly, as such. Nor will the conclusion be that we do **not** have evidence against probabilistic representations and reasoning — the Heuristics & Biases literature is simply too large for us to respond to every such claim. Rather, I will show that one particularly direct argument against probabilistic reasoning has been based on incorrect semantic assumptions; when these assumptions are corrected, the experimental results simply do not support the conclusions that have been drawn.

The particular topic addressed here is important not only for the psychological reasons just sketched, and because it must be addressed in order to support the conclusions of the previous chapter. In addition, dealing with alternative-sensitivity will provide a window into how the context-dependent threshold is determined for these adjectives.

4.3 What do *Likely* and *Probable* Mean? Semantic Assumptions and Experimental Results

4.3.1 Semantic Assumptions

Reasoning experiments are typically conducted using verbal or written materials, and as such their interpretation depends on implicit or explicit assumptions about the meanings of the test items and the pragmatics of the experimental situation. The experiments that we will examine in this chapter and the next rely on a number of semantic assumptions. Most prominent among these are a set of assumptions about the meaning of the items *likely* and *probable*. As it happens, these assumptions are by-and-large the same that are usually made by those who study epistemic modality in linguistics; and, I will show, one of them is crucially incorrect.

First, *likely* and *probable* are usually treated as synonyms in both linguistics and psychology. For example, Horn (1989) notes that, despite the different register and syntactic behavior of the two items, there does not seem to be any deep semantic difference between them. As far as I have been able to ascertain, this assumption is correct.

The second assumption is given in the following equivalence:

(4.1) $[\![\phi \textit{ is likely/probable}]\!] = 1$ if and only if ϕ is more likely/probable than $\neg\phi$.

This is the definition of *probable* given in Kratzer (1991), who does not make use of numerical probability. If we think that probability provides the domain of denotation for these items — as I do, and as nearly everyone in the psychology of reasoning assumes implicitly — then (4.1) is equivalent to (4.2):

(4.2) $[\![\phi \textit{ is likely/probable}]\!] = 1$ if and only if $prob(\phi) > 0.5$.

The assumption that the basic meaning of *likely/probable* is "more likely/probable than not" is widely held and intuitively plausible; but I will argue that it is not correct as a general characterization of the meaning of these items. To see why, we will begin by looking at some psychological experiments which make trouble for (4.1) and (4.2) and the far-reaching conclusions about human reasoning that have been drawn from these experiments.

4.3.2 Key Results

To get a feeling for the methods and subject matter, let's take an impromptu survey similar to those commonly employed in reasoning experiments.

(4.3) Imagine you've applied for a job where there are four other applicants, and you know that you and the other applicants are all equally qualified. How would you rate the following as descriptions of your chances, on a scale from 1 (completely inappropriate) to 10 (completely appropriate)?

a. It is certain that you will get the job. ______
b. It is likely that you will get the job. ______
c. It is somewhat likely that you will get the job. ______
d. It is unlikely that you will get the job. ______

After writing down your answers, consider a different case.

(4.4) You've applied for a job, but you just found out that someone else has been offered the position. You've been told confidentially that you'll get it if the other candidate withdraws, which happens (in the company's long experience) about one time in five. How would you rate the following as descriptions of your chances, on a scale from 1 (completely inappropriate) to 10 (completely appropriate)?

a. It is certain that you will get the job. ______
b. It is likely that you will get the job. ______
c. It is somewhat likely that you will get the job. ______
d. It is unlikely that you will get the job. ______

When Teigen (1988) conducted the experiments on which (4.3) and (4.4) are based — in better-controlled conditions, of course — he found a striking result: subjects responded very differently to (4.3) and (4.4). Even though the probability of getting the job (20%) was held constant, subjects rated descriptions like "likely" and "somewhat likely" as more appropriate in scenarios like (4.4), when the event of getting the job was being implicitly compared to a number of different events of similar probability, than in scenarios like (4.3), when it was compared to a single event of much higher probability.

An experiment reported by Yalcin (2009) demonstrates this effect clearly. Yalcin told one group of subjects that Team X had a 42 percent chance of winning a soccer championship, and a 58 percent chance of not winning. He then asked them whether the sentence "Team X will probably win the championship" is true or false in this scenario. 76% of subjects responded that this sentence is false, 15% judged it true, and 9% were unsure. To a second group of subjects, Yalcin described a situation in which the following distribution of chances was given:

(4.5) a. Team A: 12 percent change of winning the championship.
b. Team B: 11 percent change of winning the championship.
c. Team C: 13 percent change of winning the championship.
d. Team D: 42 percent change of winning the championship.
e. Team E: 12 percent change of winning the championship.
f. Team F: 10 percent change of winning the championship.

Even though the probability of winning was held constant (42%), the results were completely reversed: 76% of Yalcin's subjects endorses the description "Team D will probably win the championship", and only 22% judged it false. Although the use of numerical estimates may complicate these results somewhat, the pattern of responses is in line with Teigen's results. Apparently, subjects do not evaluate *likely* and *probable* (and their adverbial counterparts) by comparing a proposition to its negation, as in (4.1), but by comparing it to whatever other alternatives are provided by the context.

Across a variety of manipulations of the test conditions and the precise language employed, subjects tend to rate chances as better in situations like scenario (4.3), and as worse in scenarios like (4.4). This was true even if the question is phrased in terms of "good" versus "poor chances", or "not improbable" versus "not probable". (You may have avoided doing this yourself, but if you did not give different responses in (4.3) and (4.4), consider whether it was self-correction due to the fact that you saw the questions back-to-back.) This effect is summarized in (4.6):

(4.6) **Alternative-Sensitivity (first effect)**: An event may be rated as more probable when it is presented in contrast to a number of outcomes with similar or lower probability than when it is presented in contrast to a single focal outcome with much higher probability.

Teigen's results have been replicated and extended to other expressions, for example, by Windschitl & Wells (1998) and Yalcin (2009). Alternative-sensitivity appears to be an extremely robust phenomenon in the language of uncertainty.

A second important effect that Teigen found was that, when there was a set A of mutually exclusive and roughly equiprobable outcomes each of which was more probable than all alternatives not in A, subjects were often willing to judge all members of A as "probable". Here is Teigen's description of his experiment:[1]

> Ten days before the finals in the European Song Contest were to take place in Bergen (May 1986), 99 students in an introductory psychology course were given lists of the 20 nations participating in the contest and were asked to estimate of guess the chances for each participant to be elected winner. At that time, the Song Contest was the central current event in Bergen and the chances of individual participants were publicly and privately discussed. ...
>
> Subjects in *Group 2* ($n = 35$) were asked for each participant whether they thought it was a *probable* or *not probable* winner. There was also a third response alternative, *neither probable nor not probable*, for those cases where neither expression was felt to be appropriate.
>
> For *Group 3* ($n = 33$) the response alternatives were *improbable, not improbable*, and *neither*.

Teigen's results are given in Table 1. Group 2's results are particularly striking: on average, subjects rated 7.8 participants as "probable" winners. This behavior is clearly inconsistent with the definition of *probable* given above, where *probable* means "more probable than not" or "probability greater than 50%". Since *x wins* and *y wins* are mutually exclusive for $x \neq y$, they cannot both be probable

1 I omit discussion of Group 1, who were asked to produce numerical estimates of their subjective probabilities. Group 1 failed to do this in a normatively correct way, instead producing "subadditive" probability judgments ranging as high as 1050% total. I do not find these results particularly troubling, simply because the task of introspecting point estimates of subjective probabilities is exceedingly unnatural. Verbal expressions of uncertainty, on the other hand, provide a much more "ecological" window into subjects' representations of uncertain information.

according to the standard definitions; and yet Teigen's participants routinely judged ten or more such mutually exclusive events "probable".

Group 2			Group 3		
Expression	Mean	*SD*	Expression	Mean	*SD*
probable	7.8	3.0	not improbable	6.7	2.3
not probable	8.4	3.8	improbable	9.7	3.8
neither	3.8	3.3	neither	3.6	3.7

Table 1 Mean number of countries (of 20) judged to be probable and improbable winners of the 1986 European Song Contest in Teigen (1988).

These results indicate a second feature of alternative-sensitivity observed in this and a number of related experiments:

(4.7) **Alternative-Sensitivity (second effect)**: Multiple mutually exclusive events may be judged "probable" or "likely" when (i) they are all roughly equiprobable, and (ii) no other event is substantially more likely.

4.3.3 Psychological Interpretations

If "ϕ is likely/probable" means "ϕ is more likely/probable than $\neg\phi$", then, if speakers are reasoning correctly about probability, we expect two effects. First, no two mutually exclusive events can both be "likely" or "probable". Second, whether or not ϕ is "likely" should depend exclusively on whether or not ϕ has probability greater than 50%, and the distribution of alternative events with which to compare ϕ should not matter at all. Since subjects in the experiments we have considered routinely violate both of these predictions, we are forced to draw one of the two possible conclusions: either subjects are reasoning incorrectly, or "likely" does not mean what we thought.

Teigen (1988) takes the first route, characterizing his overall pattern of results as showing "overestimation of chances", and the behavior of subjects as "violations of the distributive law of probability theory". This interpretation is generally in line with the usual assumption that unexpected results in reasoning experiments indicate that subjects are making mistakes. Windschitl & Wells (1998) go well beyond Teigen in their analysis of alternative-sensitivity, though, providing a detailed model of alternative-sensitivity in the framework of dual-system models of cognition (see Evans 2008; Frankish 2010 for overviews).[2] The general idea is that human cognition is divided into two types of systems, a *rule-based* system and an *associative system*.

> The rule-based system represents information in relatively abstract terms and operates according to formal rules of logic and evidence ... Associative processing is relatively quick and spontaneous, but less flexible ... it is often an automatic product

2 Windschitl & Wells (1998) do not appear to be aware of Teigen's work, but three of their six experiments — and all of the ones directly relevant to us — replicate experiments reported in Teigen (1988).

that can be accompanied by an intuitive or gut-level sense. Associative processing represents information in more concrete terms and operates according to principles of similarity and contiguity. (Windschitl & Wells 1998: 1412)

Windschitl & Wells' "associative system" is often referred to as "System 1" in the dual-systems literature, and their "rule-based system" is often called "System 2".

Windschitl & Wells argue that their hypothesis of two parallel systems for reasoning about uncertainty explains alternative-sensitivity in the following way. Numerical expressions of uncertainty like *70 percent* are interpreted with respect to a learned, culturally shared system of numerical probability. This is rule-based, abstract reasoning, and does not show alternative-sensitivity, they claim.[3] On the other hand, non-numerical probability expressions like *likely* and *probable* are interpreted by the associative system, which relies on "pairwise comparisons between the focal and alternative outcomes", where "the comparison between the focal outcome and the most likely alternative has critical importance" (Windschitl & Wells 1998: 1413):

> The more this comparison favors the focal outcome (or the less it favors the most likely alternative), the greater the perceived likelihood for the focal outcome.

This model derives alternative-sensitivity as described in (4.6) straightforwardly: pairwise comparison of an outcome ϕ with one or a small number of higher-ranked alternatives $\psi_1 ... \psi_n$ will result in its being considered relatively unlikely, while pairwise comparison with a large number of lower-ranked alternatives $\chi_1, ... \chi_m$ will result in its being considered relatively likely. This holds even if the total probability of the two sets of alternatives is the same, and so we can derive the first effect of alternative-sensitivity summarized in (4.6). It is not entirely clear whether Windschitl & Wells' model derives the fact that multiple equiprobable outcomes may all be judged "probable" (4.7), since they do not give any semantic details for their experimental stimuli. However, if we suppose that the associative system returns 'true' for ϕ *is likely* as long as ϕ does not *lose* any pairwise comparison with alternatives, Windschitl & Wells can account for the second effect (4.7) as well.

There are two particularly striking features of Windschitl & Wells' account. First, although it does predict alternative-sensitivity as described in (4.6) and (4.7), it does so essentially by fiat: the process of comparing a focal outcome ϕ to alternatives is stipulated, as is the fact that ϕ's likelihood is determined by comparing it to the most likely alternative (rather than, say, the least likely, or the two hypotheses closest in likelihood to ϕ). This type of model could reasonably be criticized as a post hoc explanation which could be modified to explain any data set at all (cf. Gigerenzer 1991, who claims that most models in the Heuristics & Biases literature have this character).

Second, no faithful interpretation of Windschitl & Wells' model is coherent with currently existing theories of natural language semantics. Suppose that we take the proposed model at face value as giving the meanings of non-numerical probability expressions. The problem is simply that formal semantics is generally thought to be rule-based rather than associative, and this

3 Yalcin's experiment described above (4.5) calls this assumption into question, but we will leave this issue to the side for now. The alternative I will propose predicts alternative-senstivity with numerical expressions as well, though this may be masked by conscious hyper-correction in some cases.

applies to modal adjectives such as *probable* and *likely* just as well as any other type of linguistic expression. Furthermore, semantic theories do not generally contain mechanisms which shuttle different expressions off to different cognitive systems for evaluation. This is what would appear to be needed for Windschitl & Wells' account of the purported differences between numerical and non-numerical probability expressions to hold up. While this is not, of course, an insuperable problem, it points to the need for a great deal more detailed semantic work before a model of this type can compete seriously with the well-established results of modern semantics.

A second possibility is to suppose that the MEANING of *probable* and *likely* is "more likely than not", as in (4.1). Then it becomes a mystery why subjects do not get the right results without any difficulty: evaluating ϕ *is more likely than* $\neg\phi$ involves making a pairwise comparison, and subjects are assumed to be capable of this task. Furthermore, on this interpretation, these expressions should not be sensitive to alternatives.

A third option is that the meaning of *probable* and *likely* is "probability > 0.5", as in (4.2), but that subjects do not have the cognitive resources to evaluate probability expressions, and fall back on associative shortcuts that have little to do with the true meaning of the expressions. This interpretation, while more plausible than the previous two, is in conflict with widely held assumptions in linguistics and philosophy. In particular, it is not clear that it even makes sense to attribute a meaning m to an expression u in a language L if speakers of L do not USE u intending to communicate m, or, worse, if speakers of L are not psychologically capable of representing m. In general, it seems that Windschitl & Wells' approach to alternative-sensitivity, if it can be made to work at all, would require some pretty drastic theoretical and conceptual revisions to current conceptions of natural language semantics.

4.4 A Semantic Interpretation

I will argue for an alternative account of alternative-sensitivity that does not require any special semantic or psychological assumptions. Instead, it follows almost immediately from the semantics for *likely* and *probable* given in Chapter 3, and is based directly on numerical probability. The hypothesis is this:

(4.8) *Likely* and *probable* are **semantically** sensitive to alternatives: like other relative adjectives, they are evaluated by comparing their argument to a set of contextually salient alternatives. In the case of *likely* and *probable* the alternatives are often, but not always, provided by the denotation of the current Question Under Discussion (QUD, Roberts 1996).

This hypothesis derives the main characteristics of alternative-sensitivity (1,2) directly, and makes a number of new predictions. Many of these predictions will be shown to be correct here, although a few will require experimental verification in the future.

For the linguist, this project should be interesting because it illuminates the semantics of relative-standard epistemic modals and their interaction with context and focus. For the psychologist, this conclusions of this chapter have several useful characteristics. First, they eliminate the need to ascribe massive cognitive error to experimental subjects in this case, or to give radically different meanings to numerical and non-numerical expressions of probability. They also demonstrate the importance of careful attention to semantics and pragmatics in designing and interpreting verbally conducted reasoning experiments. Finally, this chapter illustrates the advantages of the methodology, standard in linguistics but apparently less widespread in work on reasoning, of assuming that naïve subjects and informants know what they are talking about. When we find unexpected results, rather than assuming that subjects are making mistakes, the default assumption should be that **our** models are not sophisticated enough. Of course this may turn out to be wrong, but it may lead beyond superficial assumptions to a deeper understanding of the phenomena, as I will argue it is in the present case.

4.4.1 Relative Adjectives, *Likely*, and *Probable*

4.4.1.1 Tests for Adjective Type

As we saw in Chapter 3, *likely* and *probable* pattern on numerous tests with RELATIVE ADJECTIVES like *tall*. Two of the most straightforward tests for adjective type are repeated here. Adjective type largely determines which degree modifiers are accepted, as illustrated by the close correspondence in (4.9).

(4.9) Degree Modification

a. Jeffrey is ✓very/??completely/#mostly/??slightly/#half tall.

b. It is ✓very/??completely/#mostly/??slightly/#half likely/probable that it will rain.

Furthermore, relative adjectives, unlike other adjective types, have a "zone of indifference" (Sapir 1944) in which neither they nor their antonym holds (4.10).

(4.10) Zone of Indifference

a. ✓Sam is not tall, but he is not short either.

b. ✓It is not likely that we will win, but it is not unlikely either.

This is in contrast to minimum-standard adjectives like *bent* and maximum-standard adjectives like *full*, which take different degree modifiers:

(4.11) a. This gold is ✓very/✓completely/✓mostly/#slightly/#half pure.

b. This neighborhood is ✓very/??completely/??mostly/✓slightly/??half dangerous.

The acceptable degree modifiers in (4.11) are different, in both cases, from those of *tall* and *likely* in (4.9). (Note that the expressions marked "??" are sometimes acceptable in certain contexts, but they generally do not function as degree-modifiers in these contexts.) Minimum and maximum adjectives also do not have a robust zone of indifference.[4]

(4.12) a. #? This gold is not pure, but it is not impure either.

b. #? This neighborhood is not dangerous, but it is not safe either.

These and numerous other correspondences suggest that *likely* and *probable* are, in most respects, ordinary relative adjectives.

---

4 See however Rotstein & Winter (2004), who show that examples like those in (4.12) can be acceptable in certain marked contexts. The contrast between absolute and relative adjectives is robust nonetheless: no special context is needed for the sentences in (4.10) to be true.

4.4.1.2 Comparison Classes and Significant Deviation

I will argue that the proposition-embedding relative adjectives *likely* and *probable* are similar to ordinary relative adjectives in yet another way: the effect of alternatives with these adjectives is essentially the same as the effect of comparison classes with adjectives like *tall* whose argument is an individual. To see why this is the case, I will quickly review some relevant characteristics of comparison classes here.

As we saw in chapter 1, comparison classes appear to play an important crucial role in constraining the vague, context-sensitive standard for comparison that characterizes the positive form of relative adjectives (Klein 1980; Fara 2000; Kennedy 2007). For example, the sentences in (4.13) mean very different things, and (4.14) is not self-contradictory.

(4.13) a. Sam is tall for a three-year-old.

b. Sam is tall for a professional basketball player.

(4.14) Michael Jordan is tall for an adult male, but he is not tall for a professional basketball player.

This property is not generally shared by maximum- and minimum standard adjectives, which do not have a shifting standard which needs to be fixed.

(4.15) a. #? This rod is bent for an antenna.

b. #? This room is full for a living room.

These examples are at best hard to interpret, and ungrammatical according to some authors.

The correct semantic account of comparison classes is debated: see e.g. Kennedy (2007); Bale (2011); Solt (2011). The details do not matter a great deal for this chapter; what is important is essentially that, when *A* is a relative adjective and *C* is a comparison class, *x is A for C* indicates that *x* is tall relative to some mean/median/expected/normative value which is calculated on the basis of the distribution of property *A* in class *C*.[5]

Various authors have proposed a "greater than average for *C*" semantics along these lines for the positive form of adjectives with comparison classes. This is not quite right, though: as Fara (2000) points out, if the average for comparison class *C* is $5'6''$, someone who is $5'6.2''$ will be "taller than average for *C*", but not "tall". (4.16) makes the same point (cf. Kennedy 2007):

(4.16) Steve is slightly taller than average/normal/expected, but he isn't tall.

If "tall" meant "taller than average", (4.16) would be self-contradictory. Fara suggests that SIGNIFICANT deviation from the mean/norm is what is crucial:

(4.17) *x is **pos** tall for C* is true iff **height**$(x) \geq \theta_{tall}$, where θ_{tall} is a value **significantly** greater than the average/normal/expected height for comparison class *C*.

5 Using a mean value is perhaps the most commonly assumed approach, although it's not generally recognized that this will only work if all relevant scales are at least as strong as interval scales. The use of a median value is suggested by Solt (2011). Fara (2000) suggests that typicality/normality is the correct characterization. See Solt (2011) for arguments that details of the statistical distribution of objects in a comparison class are relevant, and Barner & Snedeker (2008); Schmidt et al. (2009) for experimental investigations of various statistical models designed to predict subjects' judgments about the positive form of the relative adjective *tall* relative to various comparison classes.

The value of "significantly" is assumed to be given by context.

When *tall* is used without an overt comparison class, I will assume that C is provided by the domain of the adjective, subject to general contextual domain restriction processes. This gives us:

(4.18) x *is* ***pos*** *tall* = 1 iff **height**$(x) \geq \theta_{tall}$, where θ_{tall} is a value significantly greater than the average/normal/expected likelihood of the subset of *tall*'s semantic domain that is relevant in context.

Effectively, we are treating comparison classes as explicit domain restrictions, as in Kennedy (2007). (Other implementations are possible and would not affect the main point of this chapter.)

In addition, recall from chapter 1 that the use of a comparison class is only appropriate if the individual argument is a member of the comparison class (Kennedy 2007).

(4.19) a. # Sam is heavy for a jockey, but he isn't a jockey.

b. # This Honda is cheap for a BMW.

In general, an expression of the form *x is A for C* presupposes $x \in C$ (though there are certain exceptions, cf. Solt 2011).

4.4.1.3 *Likely/Probable* and *Tall*: Similarities and Differences

In §4.4.1.1 I claimed that *likely* and *probable* are, on most tests, ordinary relative adjectives. If this is right, then we should expect them (ceteris paribus) to have denotations roughly as in (4.20), the minimal modification of (4.17) for adjectives that take propositions rather than individuals as arguments.

(4.20) ϕ *is* ***pos*** *likely* = 1 iff $prob(\phi) \geq \theta_{likely}$, where θ_{likely} is a value significantly greater than the average/normal/expected likelihood of the subset of *likely*'s semantic domain that is relevant in context.

(4.20) contrasts sharply with the usual definition of *likely*, which compares a proposition with its negation, does not make reference to a comparison class, and does not require a significant difference. Here it is again:

(4.21) ϕ *is* ***pos*** *likely* = 1 iff ϕ is more likely than $\neg p$.

A significant difference seems to be needed here as well: as Yalcin (2010) points out, ϕ does not seem to be "likely" if it is just barely more likely than its negation. This problem can be patched up easily by adding a significance parameter to (4.21), however:

(4.22) ϕ *is* ***pos*** *likely* = 1 iff ϕ is significantly more likely than $\neg p$.

We might be tempted to stop here and not go as far as (4.20). Indeed, there appears to be some empirical support for this move: *likely* and *probable* do not appear to allow overt comparison classes (Yalcin 2009).

(4.23) #? It is likely that it will rain for a summer's day.

I find it difficult to understand (4.23) as meaning “Rain is more likely than is typical in the summer”. (Some speakers find (4.23) acceptable, though awkward.)

However, it would be too hasty to conclude on the basis of (4.23) that *likely* differs from other relative adjectives in important ways. Even if *likely* cannot combine with comparison classes, we expect (4.23) to be awkward because of the fact that *x is A for C* presupposes $x \in C$. This presupposition cannot be fulfilled here, simply because the proposition that it will rain is not an instance of a summer’s day. So at this point we just don’t know whether *likely* is semantically sensitive to alternatives, though by parity with other relative adjectives we expect it.

4.4.2 Focus and Alternatives

4.4.2.1 Focus-Sensitivity: Data

What we need in order to show that *likely* and *probable* are grammatically sensitive to alternatives is a linguistic means to evoke a set of propositions and show that affects the truth-conditions of sentences with *likely*. This set must contain the proposition denoted by *likely*’s complement in order to satisfy the presupposition of comparison classes. As it happens, FOCUS provides exactly what we need: it is standardly treated as triggering sets of propositional alternatives, one of which is the ordinary denotation of the sentence (Rooth 1992). If *likely* and *probable* are focus-sensitive, we should be able to make a direct argument that (4.20) is a better proposal than (4.22), and these expressions are grammatically sensitive to alternatives as I am claiming.

In fact *likely* and *probable* do appear to be sensitive to focus.[6] Imagine a lottery with a million tickets, in which one individual, Mr. Burns, is determined to win and buys 300,000. The rest are evenly distributed among the inhabitants of Springfield. Many speakers find a contrast between (4.24a) and (4.24b) in this scenario (reading capitalization as prosodic focus on the relevant phrases):

(4.24) a. It is likely that [MR. BURNS will win the lottery].
b. It is likely that [Mr. Burns will WIN THE LOTTERY].

In my informal surveys, everyone judges (4.24b) to be false in this scenario. However, many — perhaps most — speakers judge (4.24a) to be true. This is interesting because it indicates again that probability greater than 0.5 is not always necessary for a proposition to count as “likely”. What is new here, however, is that the contrast in (4.24) suggests that the mechanism by which the difference between (4.24a) and (4.24b) is derived must be closely related to the semantics and pragmatics of focus. This turns out to be the key to understanding alternative-sensitivity more generally.

4.4.2.2 Focus-Sensitivity and Discourse Structure

If we understand how and why *likely* and *probable* are sensitive to focus, I claim, we will also have an account of their sensitivity to contextual alternatives, the main problem which we are trying to

6 This example emerged from conversations with Salvador Mascarenhas and Seth Yalcin, one of whom probably came up with it originally.

explain. The key is the theory of information structure in discourse given by Roberts (1996), which connects focus and discourse context with a notion of alternativehood. Recent developments due to Beaver & Clark (2008) give an explicit account of *pragmatic* association with focus — that is, indirect association with focus due to the fact that focus makes salient sets of propositions. Beaver and Clark use this mechanism to explain the pragmatic focus-sensitivity of quantificational adverbs like *always*.

I will show that we can explain the focus-sensitivity of *likely* in essentially the same way: focus makes salient a set of propositional alternatives which then acts as a domain restriction on *likely/probable* in the same way that comparison classes act on *tall*. If this is correct, then (4.22) is wrong and (4.20) is correct: *likely* and *probable* really are **grammatically** sensitive to alternatives, rather than having an absolute 0.5 threshold as is generally assumed. Furthermore, the details of this theory will allow us in later sections to answer some prima facie objections to the theory that I am arguing for, and will also play an important role in the semantics for deontic and bouletic modals proposed in chapter 6.

In the theory of Alternative Semantics developed by Rooth (1985, 1992), the FOCUS SEMANTIC VALUE $[\![\cdot]\!]_F^{\mathcal{M},w,g}$ of an expression is calculated from the ordinary semantic value $[\![\cdot]\!]_O^{\mathcal{M},w,g}$ by replacing focused expressions with objects of the appropriate type from the domain of discourse. For example, the focus semantic value of (4.25a) is (4.25b).

(4.25) a. Mary is married to SAM.
b. $\{\text{Mary is married to } x \mid x \in D_e\}$

If the contextually relevant domain of discourse is $\{Mary, Sam, Harry, Lou\}$, then (4.25b) is equivalent to (4.26).

(4.26) {Mary is married to Mary, Mary is married to Sam, Mary is married to Harry, Mary is married to Lou}

(Some of these alternatives are excluded on plausibility grounds, of course.) If focus is on MARY instead of SAM, the ordinary semantic value is the same, but the focus semantic value changes:

(4.27) a. MARY is married to Sam.
b. $\{x \text{ is married to Sam} \mid x \in D_e\}$
c. {Mary is married to Sam, Sam is married to Sam, Harry is married to Sam, Lou is married to Sam}

A straightforward application of this theory is to explain question-answer congruence, illustrated by (4.28) and (4.29).

(4.28) Who is Mary married to?
a. ✓ Mary is married to SAM.
b. # MARY is married to Sam.

(4.29) Who is married to Sam?
a. # Mary is married to SAM.
b. ✓ MARY is married to Sam.

In the semantics for questions proposed by Groenendijk & Stokhof (1984), the denotation of the question *Who is Mary married to?* in (4.28) is a partition of W in which every pair of worlds in each cell agrees on the value of $\{x : \text{Mary is married to } x \text{ in } w\}$.[7] In the case at hand, this is equivalent to the focus semantic value of *Mary is married to SAM*, which is also the acceptable answer to the question in (4.28). The same holds in example (4.29).

To a first approximation, then, we can formalize the requirement of question-answer congruence as in (4.30) (Roberts 1996; Beaver & Clark 2008).

(4.30) a. S is CONGRUENT to Q if and only if the focus semantic value of some subpart of S is equal to the ordinary semantic value of Q.

b. S is an appropriate answer to Q if S is congruent to Q.[8]

The influential theory of Roberts (1996) uses the relation of question-answer congruence as the foundation for a theory of discourse structure. The basic idea is that discourses are structured around questions. It would of course be maximally useful to have an answer to the Big Question — What is the Way Things Are? — but this question is rather unmanageable. Instead, speakers adopt strategies for inquiry that involve dividing the Big Question into smaller questions which are more tractable and more directly relevant to current purposes.

The question that is implicitly or explicitly under discussion at a given moment in a discourse is the QUESTION UNDER DISCUSSION (QUD). Various conditions are needed to ensure a cooperative discourse strategy, but the most important for our purposes is the requirement of congruence to the QUD:

(4.31) A declarative sentence S is an appropriate assertion at time t if and only if S is congruent to the Question Under Discussion at t.

This is a generalization of the Question-Answer congruence requirement in (4.30). Since, when a question is explicitly asked, it becomes the QUD by default, one effect of this requirement is to enforce (4.30b). However, (4.31) is more general than (4.30b) because it also requires that assertions be congruent to **implicit** QUDs: to a first approximation, an utterance like (4.32a) is not appropriate unless speaker and listener take themselves to be discussing the question in (4.32b).

(4.32) a. ASSERTION: Mary is married to SAM.

b. IMPLICIT QUD: Who is Mary married to?

---

7 Beaver & Clark (2008) argue that Hamblin's (1973) semantics for questions, which does not assume that the denotation of a question is a partition, is more useful for a theory of focus. I am not sure which semantics makes better predictions in the case at hand; the examples that we will consider are all partitions of W anyway.

8 Two side remarks may be useful.

First, note "if" rather than "iff" in (4.30b). There are other ways to answer a question appropriately, e.g. by saying something that implicates a congruent answer.

Second, other versions of (4.30a) are stricter, requiring that S be congruent to Q simpliciter (e.g., Roberts 1996). Beaver & Clark (2008) add the "subpart" qualification in order to explain the acceptability of (4.30a) as an answer to (4.30b).

(4.30) a. Who is married to Sam?

b. I think MARY is.

Even though discourse participants' understanding of the structure of a conversation is not always perfectly aligned, it is often possible to infer from the focus structure of a speaker's utterance what she takes the QUD to be, and when appropriate to accommodate this. So, for example, someone who utters (4.32a) can be assumed to be treating (4.32b) as the QUD, whether or not the other discourse participants know this in advance.

4.4.2.3 Pragmatic Focus-Sensitivity

Beaver & Clark (2008) use the theory just described to explain the phenomenon of pragmatic focus sensitivity, using *always* as the flagship example. They use a variety of linguistic tests to distinguish the pragmatic focus-sensitivity of *always* from the grammatical focus-sensitivity of operators such as *only* (Beaver & Clark 2003); (Beaver & Clark 2008: 160-181). For example, grammatically focus-sensitive expressions cannot associate with elements that have been extracted from their complement into a higher clause, but pragmatically focus-sensitive operators can.

(4.33) a. We should thank the man who$_i$ Mary always took t_i to the movies.
b. We should thank the man who$_i$ Mary only took t_i to the movies.

Beaver & Clark (2008: 163) note that (4.33a) can be read as meaning "We should thank the man such that, if Mary took someone to the movies, it was him", while (4.33b) cannot mean "We should thank the man such that Mary took only HIM to the movies".

Likely and *probable* pattern with *always* rather than *only* here: for example, (4.34a) and (4.34b) can both be read as presupposing (4.24a), with focus on *Mr. Burns*.

(4.34) a. Mr. Burns is the one who$_i$ is likely t_i to win the lottery.
b. We should kiss up to Mr. Burns, who$_i$ is likely t_i to win the lottery.

On this and variety of other tests which Beaver & Clark (2008) use to distinguish these two operators, the focus-sensitivity of *likely* and *probable* pattern with the pragmatically focus-sensitive operator *always*.

The basic idea of Beaver & Clark's explanation of pragmatic focus-sensitivity is that quantificational adverbs like *always* have an implicit domain argument, and that this argument can — but does not have to — be filled by the set of alternatives evoked by focus. On this proposal, the **semantic** meaning of (4.35) is just (4.36).

(4.35) Mary always goes to LEEDS.
a. $\approx \lambda\sigma$: Every event in σ is an event of Mary going to Leeds.

That is, *always* has a free variable σ which restricts the class of events that it quantifies over. In principle σ can be filled by any salient class of events (and there are indeed cases in which focus is ignored in favor of, e.g., presupposition). However, because of the congruence requirement (4.31), (4.35) is only an appropriate response if the QUD is *Where does Mary go?*.[9] The partition made salient by this question is {*Mary goes to Birmingham, Mary goes to Leeds, Mary goes to Edinburgh,*

9 Or if the QUD is *Where does Mary always go?* The question discussed in the main text is more interesting for our purposes, though.

...}. Unless some other set of events is highly salient, Beaver & Clark argue, the denotation of the QUD preferentially fills in the implicit argument of *always*. The net effect is that (4.35) is roughly equivalent to (4.36):

(4.36) a. $\forall w$: Mary goes somewhere in $w \rightarrow$ Mary goes to Leeds in w.
b. "When Mary goes somewhere, it is invariably Leeds."

This is the correct interpretation of (4.35). (We are skipping over a lot of details of the derivation here: see Beaver & Clark (2003, 2008) for the full story.)

4.4.2.4 Explaining Focus-Sensitivity of *Likely* and *Probable*

A very similar story can be applied to derive the focus-sensitivity of *likely* and *probable* while also maintaining the analysis of these items as ordinary relative adjectives. First, suppose that these items have the denotation in (4.20), repeated here.

(4.37) ϕ *is **pos** likely* = 1 iff ϕ is significantly more likely than θ_{likely}, where θ_{likely} is the average/normal/expected likelihood of the subset of *likely*'s semantic domain that is relevant in context.

The proposal is simply that focus makes salient a set of propositions which then restricts the domain of *likely*. The restricted domain is then used to calculate the value of the standard value θ_{likely}, with the result that the same string can be assigned different truth-values if the focus is shifted. As initial support for the hypothesis, note that various other relative-standard modal expressions are sensitive to focus, e.g. *good*:

(4.38) a. It is good that you spilled WHITE wine on the carpet.
b. It is good that you spilled wine on the carpet.[10]

(4.38a) does not entail (4.38b). Now, *good* passes all of the standard tests for being a relative-standard adjective; in this light, the explanation for the missing entailment is presumably that the standard θ_{good} is being calculated on the basis of a different set of propositions in the two sentences. In (4.38a), spilling white wine on the carpet is being compared for goodness to this sentence's focus-triggered alternative set — i.e., other kinds of wine that could be spilled. Relative to that set of alternatives, spilling white wine is presumably above par. Nevertheless, the fact that *good* ranks high relative to this restricted domain tells us nothing about whether it ranks high relative to the domain of (4.38b), whatever it may be (spilling wine vs. spilling nothing, say). *Want*, which — as I will argue in chapter 6 — is also a relative-standard modal, displays similar focus-sensitivity (Villalta 2008).

Similarly, I suggest, when an item in the complement of *likely* is focused the value of θ_{likely} is calculated relative to the average/expected likelihood of the items in the denotation of the QUD. This provides an immediate account of the crucial data:

10 This is a modification of an example due to Krifka (2007b), who uses it to illustrate the focus-sensitivity of the sentential adverb *fortunately*. (It is perhaps not a coincidence that the latter is the adverbial form of the relative adjective *fortunate*.) I owe the observation that *good* is focus-sensitive to Dean Pettit.

(4.39) a. It is likely that [MR. BURNS will win the lottery].

b. It is likely that [Mr. Burns will WIN the lottery].

(4.39a) is a congruent answer to the question *Who will win the lottery?*. The denotation of this question is (4.40).

(4.40) a. $\{x \text{ will win the lottery} \mid x \in D_e\}$

b. {Bart will win, Lisa will win, Mr. Burns will win, ...}

If — by analogy to *always* and *good* — the domain of proposition-embedding operators is by default the denotation of the QUD, then (4.39a) should be interpreted by default as conveying (4.42).

(4.41) The likelihood that Mr. Burns will win is significantly greater than the average/expected likelihood of the propositions in (4.40b), i.e., the likelihood of winning among all the residents of Springfield.

In the scenario that we described — Mr. Burns has 300,000 tickets, while the other 700,000 are held by one person each — this condition is fulfilled. The prediction is that (4.39a) should be judged true in this scenario by someone who takes (4.40) as giving the domain of *likely*. By my judgment, as well as other speakers I have talked to, this is indeed the comparison that is being made when speakers judge (4.39a) true in the case at hand: (4.39a) is true because Mr. Burns is much more likely to win than anybody else is.

On the other hand, (4.39b) is a congruent answer to the question *What will Mr. Burns do?*. This triggers the alternative set in (4.42a). This is presumably contextually equivalent to (4.42b).

(4.42) a. $\{\text{Mr. Burns will } x \mid x \in D_{VP}\}$

b. {Mr. Burns will win, Mr. Burns will lose}

If (4.42b) provides the value of P, we predict that (4.39b) should be understood as meaning:

(4.43) The likelihood that Mr. Burns will win is significantly greater than the average/expected likelihood of the propositions in (4.42b).

Since the average probability of the propositions in (4.42b) is necessarily 0.5, (4.39b) should be judged true only if the probability of Mr. Burns' winning is significantly greater than 0.5. But this condition is not fulfilled in the scenario at hand. Thus we predict, correctly, that speakers should judge (4.39b) false.

This account derives the focus-sensitivity of *likely* and *probable* directly from the standard denotation of relative adjectives along with plausible assumptions about domain restriction connected with Beaver & Clark's (2008) account of pragmatic association with focus. Two questions remain, however. First, *most* speakers judge (4.39a) true in the scenario described, but why do some persist in judging it false? The existence of some contextual and inter-speaker variability is actually a direct prediction of the pragmatic account. Beaver & Clark (2008) note the same phenomenon with respect to *always*. There is no grammatical requirement for the QUD to supply the value of the implicit argument of *always*, but only a pragmatic default; other factors can override this preference, such as contextual salience. Speakers who insist that (4.39a) is false in this scenario are presumably interpreting it with respect to a different domain (most likely the one in (4.42)).

Second, why is the traditional account of *likely* and *probable* as having a standard fixed at 0.5 so intuitively plausible? This fact is, I suggest, due to a tendency to interpret a decontextualized declarative sentence *It is likely that* ψ as evoking a default QUD $?\psi$. For example, if I give you the example in (4.44a), you are likely to interpret it as responding to the question in (4.44b) unless you have a specific reason to do otherwise.

(4.44) a. DECONTEXTUALIZED SENTENCE:
It is likely that it will rain tomorrow.

b. DEFAULT QUD:
Will it rain tomorrow?

If (4.44b) supplies the domain on the basis of which θ_{likely} is calculated, then θ_{likely} will indeed be (roughly) 0.5, because the denotation of $?\psi$ is $\{\psi, \neg\psi\}$, and the average probability of the propositions in this set is necessarily 0.5. This effect, I suggest, is why the usual interpretation of ψ *is likely* as "ψ is more likely than $\neg\psi$" is so initially compelling. In many contexts, these sentences will indeed have the same truth-conditions. In contexts where sentences are presented without a clear discourse context, as in the usual style of presentation in academic papers in formal semantics, they will nearly always be equivalent.

Summing up the results of this section, the pragmatic account of focus-sensitivity that Beaver & Clark (2008) propose for *always* extends readily to explain the focus-sensitivity of *likely* and *probable*. This provides strong support for the hypothesis that *likely* and *probable* have a context-sensitive standard, just as other relative adjectives do. Furthermore, it explains why there is contextual and inter-speaker variability in judgments, and gives an account of why so many scholars have mistakenly thought that *likely* means "more likely than not". The default interpretation of ψ *is likely* when context or focus does not provide a value for the comparison class turns out to be equivalent to this paraphrase.

4.5 Explaining the Experimental Results

4.5.1 Yalcin's (2009) Experiment

The hypothesis that *likely* and *probable* are semantically sensitive to contextual alternatives goes a long way toward explaining the experimental results showing alternative-sensitivity in probability judgments. Consider first one of Yalcin's (2009) experiment, whose stimuli were given in (4.5). When told that Team X has a 42 percent chance of winning and a 58 percent chance of losing, 76% of subjects judged the sentence "Team X will probably win" to be false. On the other hand, when this team was situated among five other teams with lower probabilities of winning, this sentence was judged true by 76% of subjects, even though the probability of winning was held constant at 42%.

Assuming that *probably* ϕ is equivalent to ϕ *is probable*, the account of alternative-sensitivity just given predicts these results immediately. In the first case, the set of salient alternatives {Team X wins, Team X does not win} supplies the comparison class which is used to calculate $\theta_{probable}$. The prediction is that "Team X will probably win" should be true if and only if Team X's chances of winning (.42) are significantly greater than $\frac{.42+.58}{2} = .5$, which is false here.

On the other hand, "Team X will probably win" should be true in the second context just in case their chances of winning (.42) are significantly greater than the average of the alternatives $\frac{.12+.11+.13+.42+.12+.10}{6} \approx .17$, which is likely to come out as true (depending on how "significantly" is interpreted in context, as usual).

This result shows that this approach predicts the first effect of Alternative-Sensitivity, repeated here:

> **Alternative-Sensitivity (first effect)**: An event may be rated as more probable when it is presented in contrast to a number of outcomes with similar or lower probability than when it (or another event with the same probability) is presented in contrast to a single focal outcome with much higher probability.

4.5.2 Teigen's European Song Contest Experiment

The theory also accounts for the results of Teigen's (1988) experiment involving the European Song Contest. Recall that many subjects rated a large number of contestants as "probable" winners, some as many as 11 out of 20. The domain in this context is presumably the set of propositions $\{x \text{ wins} \mid x \in D\}$, where D is restricted to the participants in the competition. Teigen's results indicate that many of his subjects judged that eight or more contestants had a probability of winning significantly greater than average for D. Since there are 20 teams and one of them will win, the average probability of winning is necessarily $\frac{1}{20} = .05$. On this interpretation, then, Teigen's subjects simply judged that a considerable number of contestants had a probability significantly greater than .05 of winning. Again, depending on how "significantly" is interpreted, this may not be not particularly shocking.[11]

In this way, we can explain for the second effect of alternative-sensitivity:

> **Alternative-Sensitivity (second effect)**: Multiple mutually exclusive events may be judged "probable" or "likely" when (i) they are all roughly equiprobable, and (ii) no other event is substantially more likely.

This account applies, mutatis mutandis, to Windschitl & Wells's (1998) Studies 1-3 as well.

4.5.3 New Distributional Predictions

The explanation of Teigen's results just given suggests a new prediction of the current theory which is in need of experimental investigation. If it holds up under testing, this will be a strong point in favor of the current theory against Windschitl & Wells (1998), who do not make this prediction.

Consider a subject in Teigen's European Song Contest experiment who judged that eight contestants have a probability of winning significantly greater than the average of .05. Suppose (arbitrarily) that the level of significance is .04, so that a contestant must have probability at least .09 in order to count as a "probable" winner in this context. Then eight contestants have probability

11 This interpretation does not predict the results of Teigen's Group 1. However, remember that the task of producing point estimates of subjective probabilities is unnatural and these results are not very reliable.

at least .09 of winning, and so the probability that one of these contestants will win is at least $.09 \times 8 = .72$. It follows that the probability that one of the other 12 contestants will win cannot be greater than .28, and so the average probability of winning among the 12 contestants who are not "probable" winners is at most $\frac{.28}{12} \approx .023$, and very likely even lower.

The prediction, then, is this: if a large number of contestants are "probable" winners, then there should be a number of other contestants that have a very **low** probabilty of winning. It cannot be the case, for example, that the probability of winning is distributed equally among the contestants, and all of them are "probable" winners; instead we predict that the best answer will be "somewhat probable" or the like. In contrast, Windschitl & Wells's (1998) account does not make any predictions about the distribution of probabilities among lower-ranked alternatives: their interpretation heuristic for likelihood focuses on the comparison between an outcome ϕ and the most likely alternative ψ such that $\phi \neq \psi$, without making any detailed predictions about the relationship between these and other alternatives.

These predictions can be tested experimentally in the following way. Recall the job search experiment described above.[12] Suppose that the options are "certain", "likely", "somewhat likely", and "unlikely" in (4.3) and (4.4). When competing with someone who is far more likely to get the job, as in (4.3), the optimal response is predicted to be "unlikely". If there are five equally qualified applicants as in (4.4), however, the prediction is that the optimal response should not be "likely" but "somewhat likely". The reason is that, with five applicants with probability .2 each of getting the job, no one is significantly more likely than average, so "likely" is not the most appropriate response. But no one is "unlikely" either, since no one is significantly *less* likely than average to get the job. The best response, then, should be "somewhat likely", the intermediate response among those offered here. This differs from the predictions of Windschitl & Wells (1998), who — at least on the charitable re-interpretation we suggested above to deal with equiprobability — would predict that all five applicants should be "likely" to get the job.

A set of modified experiments should help to disentangle these predictions and to confirm or falsify the predictions of the present theory. Rather than having five equally qualified candidates, there should be a large number, say 20, where two candidates A and B are better than the rest. The experimental manipulation involves how much better A and B are than the rest of the pack. If A and B are described as "slightly more qualified than any of the others", more subjects should choose the response "likely" than did when there are five equally qualified, but most will still choose "somewhat likely". However, if A and B are described as "much better than any of the others", more subjects should endorse the description of them as "likely" to get the job.

If fine-grained manipulations in the distribution of probabilities among lower-ranked alternatives affects the results of the experiments described here, we have evidence that the current theory is on the right track, and that alternative-sensitivity really is grammatical in the way described here. Furthermore, since the logic behind the prediction relied crucially on details of the probability distribution, this result would count as evidence that subjects are reasoning using standard

12 We do not have ready results to report because the stimuli which Teigen (1988) used in conducting the experiment on which (4.3) and (4.4) are based were different. Teigen used "reasonable hope", "not improbable", "doubtful" and "improbable" as alternatives, a rather heterogeneous set for which we don't yet have a detailed account. However, the general approach outlined here should apply to these results as well, cf. §4.6.3.

probability.

4.6 Further Considerations

There are, I suspect, a number of further issues relating to this proposal that need to be considered. Here I briefly consider three that seem particularly pressing: first, that proposal seems to allow the threshold for being "probable" to be too low; second, that alternative-sensivitity is more limited with proposition-embedding than with adnominal *likely* and *probable*; and third, how the account generalizes beyond *likely* and *probable* to other expressions of uncertainty.

4.6.1 Thresholds and Context Change

One worry about the account proposed here is simply that the standard for counting as *likely* or *probable* seems too low. Yalcin (2009), discussing the general idea of lowering the threshold for *likely* below 50%, makes this point clearly:

> [T]his would mean that in some contexts, both ϕ and $\neg\phi$ are probable, meaning that the inference pattern *Probably to not probably not* is not really valid. This is hard to believe.

This is a serious worry, but it does not apply to the present account. The reason is that a context in which we are comparing the likelihood of ϕ and $\neg\phi$ is ipso facto a context in which the QUD is $?\phi$. In contrast, cases in which the "threshold" falls below 50% are always cases in which the QUD is not $?\phi$ but a partition for which the value of θ_{likely} may be significantly lower than 50%.

In other words, ϕ and $\neg\phi$ cannot both be probable in the same context because, if we are in a context where the threshold is low, ϕ and $\neg\phi$ cannot both be among the alternatives. If we are in a context where the threshold is low and ϕ is one of the alternatives, considering the question whether $\neg\phi$ is probable *changes* the context so that the QUD is $?\phi$. This conversational move will raise the threshold back to 50% (plus the value of the significance parameter). The inference from *probably* ϕ to *not probably not* ϕ is valid, but only if you hold the context fixed. Asking the question whether this inference holds may change the context in a way that makes it appear invalid, but this is an illusion (cf. Yalcin 2010:932 for a related diagnosis developed independently).

4.6.2 Proposition-Embedding vs. Adnominal *Likely* and *Probable*

A complication for the account given here is that the second effect — the possibility of multiple mutually exclusive events all counting as "likely" or "probable" — seems to be sensitive to the grammatical position of the epistemic adjective.

(4.45) [*A* and *B* are two equally matched contestants in a race, and both are far better than all other entrants.]

a. ✓ *A* and *B* are both likely winners.

b. # It is likely that *A* will win, and it is likely than *B* will win.

There are relevant differences between adnominal *likely* in (4.45a) and proposition-embedding *likely* in (4.45b), then. A similar pattern obtains for *probable*. However, it is still the case that proposition-embedding *likely* and *probable* display the first effect of alternative-sensitivity, the possibility of a threshold lower than 0.5, as we saw already in (4.24). The question is why equiprobable events cannot both be described using *It is likely that ...*, even though the possibility of a threshold lower than 0.5 would seem to make this available in principle.

I suspect that this effect is due to the information structure of sentences such as *It is A that* ϕ. Note, first of all, that (4.46) is just as bad as (4.45b) in the case at hand.

(4.46) # *A* is the likely winner, and *B* is the likely winner.

This is unsurprising, of course: the use of *the* here brings with it a presupposition that there is a unique likely winner for the contest at hand, which is incompatible with the meaning of the sentence. (4.45b) can be explained in the same terms if we can find evidence that sentences of the form *It is A that* ϕ quite generally lead to the inference that it is not *A* that ψ, where ψ is one of the propositional alternatives of ϕ.

This is a question worthy of more space than we can dedicate to it here, but the initial results are encouraging. For example, the (a) sentences in the following two examples strongly imply the negation of the same sentence for each of the contextual alternatives, and the (b) sentences are rather odd ways to make the intended point.

(4.47) [I am the night shift manager at a diner. One of the day shift workers tells me on the way out the door:]

a. It's doubtful that SUE will make it to work tonight.
$\rightsquigarrow$ The speaker can't say that it's doubtful that x will come, for any $x \neq$ Sue among our workers.

b. #? It's doubtful that SUE will come tonight, and it's doubtful that BILL will come tonight.
(contrast: "It's doubtful that Sue or Bill will come.")

(4.48) [Mary has three ex-boyfriends, Sam, Stan, and Larry, any of whom might appear at the birthday party being thrown by her new boyfriend.]

a. It's good that SAM came to the party.
$\rightsquigarrow$ It would not have been good if Stan or Larry had shown up.

b. #? It's good that SAM came to the party, and it's good that LARRY came to the party.
(contrast: "It's good that Sam and Larry came.")

If it is correct that sentences of this form tend to imply that the adjective holds uniquely of its complement among the contextual alternatives, then we have an account of the oddity of (4.45b): the example leads strongly to an inference which is contradicted by its content, and one which could have been avoided by choosing a minimally different structure to communicate the intended meaning.

4.6.3 Beyond *Likely* and *Probable*

Throughout this chapter we have focused exclusively on the adjectival epistemic modals *likely* and *probable*, giving an account which derives the basic facts of alternative-sensitivity without attributing massive error to experimental subjects, and without abandoning our main claim that epistemic modals have a probabilistic semantics. Now, even though *likely* and *probable* are indeed among the most popular items of investigation in the psychology of reasoning, the experimental results that we have discussed deal with a somewhat broader range of linguistic items than the two that we have focused on.

For example, the relevant experiments in Windschitl & Wells (1998) ask subjects to rate a range of modified expressions: *certain, extremely likely, quite likely, fairly likely, ... , quite unlikely, extremely unlikely, impossible*. The account given here extends directly to these, with one caveat. Relative adjectives are generally sensitive to alternatives even when modified: for example, they still take comparison classes.

(4.49) a. Jeffrey is extremely tall for a three-year-old.
b. Jeffrey is extremely tall for a professional basketball player.

On this basis, we expect modified forms of *likely* and *probable* to be sensitive to alternatives, as indeed they are. However, we do not expect alternative-sensitivity to affect the extreme values *certain* and *impossible* since these are maximum-standard adjectives, a type which do not combine felicitously with comparison classes and do not have a vague contextually fixed threshold. This is a very plausible prediction, and is compatible with the results of Windschitl & Wells (who include them among their test items).

Teigen (1988) runs experiments using "likely" and "probable" as well as a wider range of expressions of uncertainty: "reasonable hope", "good/great/small chances", and "doubtful", and "high/low probability". He finds alternative-sensitivity with essentially all of them. On the basis of the linguistic characteristics of this sample, these results are not surprising: every single one of these expressions contains a relative adjective, and is therefore expected to be grammatically sensitive to alternatives, all else being equal.

On the other hand, my account makes the prediction that it should not be possible to find robust alternative-sensitivity with modal expressions such as *certain, impossible, possible, clear, evident, indubitable, ...* that are not relative adjectives. It is interesting to note in this light that these items almost never appear in reasoning experiments. These items are not contextually variable to nearly the extent that relative adjectives are. If I am right, it will in many cases not be possible to reproduce the "fallacies" that have been discovered using relative adjectives if we replace the test items with minimum- and maximum-standard adjectives.

4.7 Conclusion

Grammatical alternative-sensitivity is the usual behavior of relative adjectives, and we should expect it, ceteris paribus, for *likely* and *probable* as well. Most relative adjectives, like *tall*, are semantically sensitive to alternatives. For proposition-embedding relative adjectives such as *likely* and *probable*, the domain-restricting role of the comparison class can be played by a set of propositions provided by the partition induced by the current Question Under Discussion in the sense of Roberts (1996). This explains why these expressions are sensitive to focus, and also explains why the usual equation of "probable" with "more probable than not", although not generally correct, makes the right prediction in neutral contexts. It also predicts the findings of alternative-sensitivity by Teigen (1988) and Windschitl & Wells (1998) in a straightforward way.

As a result, evidence for the alternative-sensitivity of verbal probability judgments does not provide evidence for the claim that humans behave irrationally in making probability judgments that are sensitive to the distribution of alternatives. Likewise, these experimental results provide no support for the claim that humans do not represent or reason about uncertainty probabilistically. On the contrary, once independently motivated semantic facts about gradable adjective and focus are taken into account, the probabilistic approach does an extremely good job of explaining the facts. This conclusion is direct evidence for one of the core hypotheses of this dissertation, the claim that probability plays a crucial role in the semantics of modality. It also provides indirect support for the hypothesis that uncertainty is indeed represented probabilistically in human cognition.

CHAPTER 5

Five Problems for Quantificational Semantics for Deontic and Bouletic Modality

This chapter describes a number of problems which arise for quantificational semantics for expressions of obligation, desire, needs, and requirements. After a review of the motivations and core features of Kratzer's and related approaches in §5.0, §§5.1-5.5 describe five problems which can be attributed to the assumption that the semantics of obligation and desire is appropriately modeled by quantification over possible worlds. The puzzles in §5.1 call into question, in various ways, the assumption that modals are upward monotonic which is built into quantificational theories. §5.2 describes several scenarios which suggest that quantificational semantics does not interact with knowledge in a sufficiently fine-grained way. §§5.3-5.4 highlight the severe problems that gradability and comparison pose for quantificational semantics. Finally, §5.5 discusses conflicts of obligation, which are ruled out as logical impossibilities on quantificational theories. With one exception (the problem of excessive incomparabilities), the objections to quantificational semantics for deontic and bouletic modals apply to other standard theories just as well as Kratzer's.

There is a large literature on the paradoxes of deontic logic, including some of the problems that I will describe (although several of the puzzles in this chapter come from very recent literature, and the ones in §5.3 are new here). Some of the issues that I will discuss could be dealt with in any of a number of ways. I will not try to discuss all of the possible analyses here. Instead, I will argue that these are not isolated leaks in the theory which can be patched up as they arise, but indicate that the main body of deontic logic has been founded on a mistaken assumption that modals are appropriately modeled by first-order quantifiers $\forall$ and $\exists$, usually with some complicated additional apparatus to provide restrictions for the quantifiers. I will argue that, taken together, the puzzles indicate that obligations, needs, requirements, and desires are non-monotonic and information-sensitive, and that they have a closely related semantics which is built around scales. In Chapter 6 I will give a detailed semantic proposal which has these characteristics and fits neatly with the discussion of scale types and the treatment of concatenation as join in Chapter 2.

5.0 Ideal- and Best-Available-Worlds Semantics

A traditional idea in possible-worlds semantics for deontic modals and neighboring concepts is that we want to derive truth-conditions like these: *You ought to do your homework* is true if and only if, in all deontically accessible worlds, you do your homework; *I want to go to the Disneyland* is true if and only if, in all of my bouletically accessible worlds, I go to Disneyland; *You may have a piece of candy* is true if and only if, in some deontically accessible world, you have a piece of candy; and so on.

Semantics for modals of this sort comes in a variety of forms, but, with respect to the logical properties that we are interested in here, they are all equivalent to a construction like the following: context determines a set of accessible worlds, possibly subject to constraints imposed by the lexical semantics of the modal in question; we then quantify over this set using some combination of $\exists$, $\forall$, and $\neg$. Let **Deo** be a function from contexts c and worlds w to the set of worlds which are deontically accessible from c and w, and let **Boul** be the same for bouletically accessible worlds. Traditional modal semantics gives us:

(5.1) a. $[\![\text{You ought to do your homework}]\!]^{\mathcal{M},w,g,c} = 1$ if and only if:
$\forall w' \in \textbf{Deo}(c)(w) : [\![\text{You do your homework}]\!]^{\mathcal{M},w',g,c} = 1$.

b. $[\![\text{You may have a piece of candy}]\!]^{\mathcal{M},w,g,c} = 1$ if and only if:
$\exists w' \in \textbf{Deo}(c)(w) : [\![\text{You have a piece of candy}]\!]^{\mathcal{M},w',g,c} = 1$.

c. $[\![\text{I want to go to Disneyland}]\!]^{\mathcal{M},w,g,c} = 1$ if and only if:
$\forall w' \in \textbf{Boul}(c)(w) : [\![\text{I go to Disneyland}]\!]^{\mathcal{M},w',g,c} = 1$.

I will call this analysis of deontic and bouletic modals **Ideal Worlds Semantics**, since the accessible worlds are to be thought of as the best possible worlds from the relevant perspective.

There are a number of well-known problems with ideal-worlds semantics for deontic modals. One of the most widely discussed in deontic logic are the problem of contrary-to-duty obligations. For example, if little Mary steals a piece of candy from the corner store then, as a good parent, I ought to ground her as punishment. Supposing that *if* is the material conditional, this gives us:

(5.2) If Mary steals, she ought to be grounded.

a. $= 1$ iff: $[\![\text{Mary steals}]\!]^{\mathcal{M},w,g,c} \supset [\![\text{Mary ought to be grounded}]\!]^{\mathcal{M},w,g,c}$

b. $= 1$ iff: $[\![\text{Mary steals}]\!]^{\mathcal{M},w,g,c} \supset \forall w' \in \textbf{Deo}(c)(w) : [\![\text{Mary is grounded}]\!]^{\mathcal{M},w',g,c} = 1$

It is generally thought that (5.2b) is not an adequate analysis of this sentence, however. Intuitively, (5.2) suggests a situation in which the ideal worlds are those in which Mary does not steal, and is not grounded. If this is the case, ideal-worlds semantics predicts that both *It is not the case that Mary steals* and *It is not the case that Mary ought to be grounded* are true. (5.2) is arguably true in this case, but trivially: semantically it is on a par with *If Mary steals, she ought to be taken to Disneyland* and *If Mary steals, she ought to be executed*. This is surely the wrong result, though. What (5.2) intuitively gives us is a recommendation for action if we find ourselves in one of the sub-ideal worlds in which Mary **does** steal. In this contingency, it is indeed optimal that she be grounded, even though this is not globally optimal.

The most common response to this problem (in deontic logic as well as Kratzer's semantics) is to modify both the conditional and the deontic selection function. One way to do this by treating **Deo** as a function from worlds and contexts to **preference orders** rather than unstructured sets. (Note that the term "preference order" can be used ambiguously to a binary order over propositions, where $\phi \succcurlyeq \psi$ is interpreted as "It is at least as good/desirable if ϕ occurs as it is if ψ occurs", or a binary order over worlds, where $w \succcurlyeq w'$ is read "world w is at least as good/desirable as world w'". I will try to say explicitly which type of ordering is relevant in each case, since conflating the two can lead to serious confusion.)[1] Despite the name, the preference orders do not have to

1 Different authors have different ideas about how the preference order is determined. For instance, some authors take the ordering over worlds or events as a primitive of the model, while (as we saw in chapters 1 and 3) Kratzer derives it an ordering over worlds $\succcurlyeq_{\mathbf{g}(w)}$ from an unstructured set of propositions, and then derives an ordering over propositions from the derived order $\succcurlyeq_{\mathbf{g}(w)}$. However, these differences will not be crucial in most of this chapter, as we are concerned primarily with the interpretation of modals using quantifiers. §5.4 will discuss some specific problems that arise for Kratzer that are related to her two-step procedure for inducing an order over propositions, however.

be tied to any person's preference: they can in principle be determined by group preferences, or God's preferences, or whatever other procedure you like as long as it yields a binary order. As Lewis (1978) puts it: "The semantic analysis tells us what is true (at a world) under an ordering. It modestly declines to choose the proper ordering. That is work for a moralist, not a semanticist".

The advantage of using preference orders is that, instead of treating modals as quantifiers over a single set of (deontically, bouletically, etc.) optimal worlds, we can relativize evaluation of modals to the set of worlds which are optimal given some particular condition — for instance, the worlds which are optimal under the assumption that Mary does steal something. Even though there are no absolutely ideal worlds in which Mary steals, the worlds in which she steals and is grounded are still better than worlds in which she steals and is not grounded. Suppose for illustration that we have four worlds as in (5.3).

(5.3) w_1: Mary does not steal and is not grounded
w_2: Mary does not steal and is grounded anyway
w_3: Mary steals and is not grounded
w_4: Mary steals and is grounded

Assuming that there is no other reason to ground Mary, a plausible order is:

(5.4) $w_1 \succ w_4 \succ w_3 \approx w_2$

(Note that there is no need for preference orders to be connected as (5.4) is: for instance, Kratzer's theory normally does not generate connected orders, cf. §5.4.1 below.)

Rather than interpreting (5.2) as in (5.2b) we now treat it as meaning that, in the best worlds in the preference order which satisfy the antecedent *Mary steals*, it also holds that *she is grounded*. This is derived compositionally using the analysis of conditional antecedents as restrictors of the modal ordering, as sketched in ch.1, §1.6:

(5.5) $[\![\textit{If } \phi \textit{ then } \psi]\!]_{\mathbf{h}}^{\mathcal{M},w,g,c} = [\![\psi]\!]_{\mathbf{h}'}^{\mathcal{M},w,g,c}$, where

a. For all w, $\mathbf{h}(w)$ is the relevant deontic ordering on worlds;
b. For all w, $\mathbf{h}'(w) =_{df} \mathbf{h}(w) \upharpoonright \{w' \mid [\![\phi]\!]^{\mathcal{M},w',g,c} = 1\}$.

On this interpretation, the preference order still verifies *Mary should not steal*, since, in all of the best worlds, she does not steal. Equally, it verifies *If Mary steals, she should be grounded*, this time for non-trivial reasons: eliminating worlds in which the antecedent does not hold from (5.4), we wind up with the derived preference order

(5.6) $w_4 \succ w_3$

We then ask whether the consequent *Mary is grounded* is true at all w' which are optimal relative to the preference order in (5.6); finding that it is, we conclude that *If Mary steals, she ought to be grounded* is true.

This simple example illustrates the general character of more sophisticated deontic and bouletic logics, including Kratzer's: modals are interpreted as existential or universal quantifiers with restrictions determined in some complicated fashion. I will call this style of semantics for deontics and desire verbs **Best-Available-Worlds** semantics.

Best-Available-Worlds analyses are the most prominent in the literature on deontic modals, and their basic properties are shared by Kratzer's theory. For most of the discussion in this chapter the choice between Ideal- and Best Available-Worlds semantics will not be crucial, since I am mostly interested in posing problems for the assumption that the relevant expressions pick out existential and universal quantifiers, however complicated the derivation may be otherwise. However, when it is relevant, my main stalking-horse will be Best-Available-Worlds analyses and Kratzer's in particular.

5.1 Problem One: Deontic and Bouletic Modals are not Upward Monotonic

UPWARD MONOTONICITY is a property of operators defined as:

(5.7) An operator $\mathcal{O}$ is **upward monotonic** iff$_{df}$ $[\phi \rightarrow \psi] \vDash [\mathcal{O}(\phi) \rightarrow \mathcal{O}(\psi)]$.

The first-order quantifiers $\forall$ and $\exists$ are both upward monotonic; this means that the inference schemas in (5.8) and (5.9) are valid.

(5.8) a. $\exists x \phi$
b. $\phi \vDash \psi$
c. $\therefore \ \exists x \psi$

(5.9) a. $\forall x \phi$
b. $\phi \vDash \psi$
c. $\therefore \ \forall x \psi$

Likewise, the counterparts of the universal and existential quantifiers in English are upward monotonic in their nuclear scope, as the intuitive validity of the inferences in (5.10) shows:

(5.10) a. Mary has a red car. So, Mary has a car.
b. All of the boys are in Atlanta. So, all of the boys are in Georgia.

An important consequence of the upward monotonicity of the quantifiers is the validity of inferences involving conjunction and disjunction. (5.11) follows from the fact that $(\phi \wedge \psi) \vDash \phi$ in propositional logic, and (5.12) from the fact that $\phi \vDash (\phi \vee \psi)$.

(5.11) a. $\exists x[P(x) \wedge Q(x)] \vDash \exists x[P(x)]$
b. $\forall x[P(x) \wedge Q(x)] \vDash \forall x[P(x)]$

(5.12) a. $\exists x[P(x)] \vDash \exists x[P(x) \vee Q(x)]$
b. $\forall x[P(x)] \vDash \forall x[P(x) \vee Q(x)]$

The prediction is that the inferences in (5.13)-(5.14) are also valid:

(5.13) a. A friend of mine lives in Houston and has a cat. So, a friend of mine lives in Houston.
b. Everybody knows Bill and everybody hates Fred. So, everybody knows Bill.

(5.14) a. Some employee is in Atlanta. So, some employee is in Atlanta or Pittsburgh.

b. All of the boys are in Atlanta. So, all of the boys are in Atlanta or Pittsburgh.

The validity of the inferences in (5.13) is uncontroversial. Although the inferences in (5.14) are somewhat less natural than (5.10), they seem to be valid nevertheless. To the extent that they are odd, this can be attributed to general pragmatic principles (Grice 1989, though cf. Zimmermann 2000; Geurts 2005; Simons 2005 for a different diagnosis).

There is a third problematic inference, (5.15a), whose validity is restricted to the universal quantifier. The counterpart of this inference in English is clearly valid as well, as (5.15b) illustrates.

(5.15) a. $\forall x[P(x)] \wedge \forall x[Q(x)] \vDash \forall x[P(x) \wedge Q(x)]$

b. Everybody has a cat, and everybody has a dog. So, everybody has a cat and a dog.

Since the universal and existential quantifier are both upward monotonic, quantificational theories predict that all modals should be upward monotonic as well. In addition, modals which are modeled as universal quantifiers should validate the world-quantifying counterpart of (5.15b). However, I will argue that all of these predictions admit of counterexamples: for many deontic and bouletic modals $\mathcal{D}$,

(5.16) a. $\mathcal{D}(\phi) \nvDash \mathcal{D}(\phi \vee \psi)$.

b. $\mathcal{D}(\phi \wedge \psi) \nvDash \mathcal{D}(\phi)$.

c. $\mathcal{D}(\phi) \wedge \mathcal{D}(\psi) \nvDash \mathcal{D}(\phi \wedge \psi)$.

(The third inference is a problem only for what I will call mid-scalar $\mathcal{D}$-modals, for reasons which will emerge in chapter 6.)

5.1.1 Ross' Paradox

Ross (1944) posed a problem involving the interaction between imperatives and disjunction which is equally problematic for deontic modals and desire verbs; I'll use the term **Ross' Paradox** to refer to the entire class of problematic cases. The validity of the inferences in (5.12) is unaffected when quantification over individuals is replaced with quantification over worlds; as a result, if modals are quantifiers over worlds, inferences of this form should be valid with modals, subject to similar caveats about pragmatic oddity. However, the inferences in (5.17) have a rather different flavor from those in (5.13).

(5.17) a. The boss wants you to go to Atlanta. So, the boss wants you to go to Atlanta or Boston.

b. Mary needs to work harder. So, Mary needs to work harder or quit her job.

c. We are required to drive less than 70mph. So, we are required to drive less than 70mph or more than 100mph.

d. You should wash the dishes. So, you should wash the dishes or break them.

There is something pathological about the inferences in (5.17). In each case, the second sentence indicates that there is a way of satisfying the boss' desires, Mary's needs, your dish-cleaning duties, etc. which is not made available in the first sentence. More vividly, perhaps, it is clear that someone who responds to *You should wash the dishes* by breaking the dishes cannot claim to have been acting logically.

If the inferences in (5.17) are not valid, however, then the claim that *want*, *need*, *require*, *should*, etc. are upward monotonic is in jeopardy. They are all instances of the schema in (5.18), which is valid. (I use **Acc** as a placeholder for the set of worlds being quantified over, however this is determined.)

(5.18) $[\forall w' \in \textbf{Acc} : [\![\phi]\!]^{\mathcal{M},w',g} = 1] \vDash [\forall w' \in \textbf{Acc} : [\![\phi \vee \psi]\!]^{\mathcal{M},w',g} = 1]$

Unless we can find some alternative explanation, we are forced to question the assumption that *want, need, require*, and *should* are universal quantifiers over worlds. The same point applies to other strong and weak deontic modals, including *ought* and *may*.

It has sometimes been suggested (e.g., Hare 1967; Wedgwood 2006) that the intuitive invalidity of the inferences in (5.17) can be chalked up to Grice's Maxim of Quantity, as I suggested the cases in (5.13) can. To my ear, the inferences in (5.17) are markedly worse than those in (5.13), though of course the raw intuition is not in itself a compelling argument. More convincingly, there are several clear semantic/pragmatic contrasts between the inferences in (5.13) and those in (5.17) which calls this analysis into question. For example, Cariani (2011) notes that deontic modals differ from both universal quantifiers and epistemic modals with respect to their behavior in downward-entailing contexts.

(5.19) a. I doubt that Lynn ought to either wear a tie or a scarf. In fact, I'm reasonably certain that she ought to wear a scarf.[2]

b. # I doubt that everyone is Italian or French. In fact, I'm reasonably certain that everyone is Italian.

(5.19a) seems to be coherent, with or without the special intonation that indicates metalinguistic negation, as Cariani points out. In contrast, (5.19b) is extremely hard to make sense of even with heavy emphasis on *or*. If *ought* is a universal quantifier over worlds, however, we expect that these texts should have the same status.

The acceptability of retractions provides an even clearer contrast between deontic modals and universal quantifiers. It is perfectly reasonable for Sam to retract his claim in (5.20c) by admitting incorrectness, but strikingly odd to do the same with *everyone* instead of a deontic modal.

(5.20) (Playing Bridge)

a. Sam: According to the rules, you have to follow suit or play the king of trumps.

b. Joan: No, the rules quite explicitly say you ought to follow suit, no matter what.

c. Sam: I guess what I said was wrong. (Cariani 2011, slightly modified)

(5.21) a. Sam: Everyone followed suit or played the king of trumps.

b. Joan: Nobody had the king of trumps. Everyone followed suit.

c. Sam: # I guess what I said was wrong.

2 To be sure, there is one reading on which (5.19a) is unquestionably incoherent, with *or* taking scope between *doubt* and *ought*. However, Cariani's point goes through if there is also a reading on which this is a coherent sequence, with *doubt* > *ought* > *or*; (5.19a) is not predicted to be coherent on this reading either by standard semantics, but it does appear to be.

Deontic modals also differ from epistemic modals in this respect. Not coincidentally, epistemic modals **are** upward monotonic in the semantics that I proposed in chapter 3.

(5.22) a. Sam: Joe must have followed suit or played the king of trumps.
b. Joan: Jane had the king of trumps. He must have followed suit.
c. Sam: # I guess what I said was wrong. (Cariani 2011)

These differences between *everyone* and deontic *must* indicate that deontic modals and universal quantifiers do not display the same logical behavior in this context. This militates against a Gricean account of the data in (5.17) and in favor of the widely held opinion that Ross' Paradox is a real problem. The reason that the inferences in (5.17) are intuitively invalid, I suggest, is that they **are** invalid: unlike the overtly similar cases with a universal quantifier, there is (at least sometimes) a real semantic conflict between *must/ought*(ϕ *or* ψ) and *must/ought*(ϕ). If this is right, we have good reason to doubt that *must* and *ought* are universal quantifiers of any stripe. Furthermore, the contrast with epistemic *must* — which is not a universal quantifier, as I argued in chapter 3, but is upward monotonic — suggests a diagnosis. Deontic modals and desire verbs simply do not have the property of upward motononicity, contrary to what is nearly always assumed.[3]

5.1.2 Professor Procrastinate

The reader may still be tempted by a Gricean account of Ross' Paradox, or a theory which modifies the semantics of *or* to achieve the desired result. I think that this would be a mistake: a parallel argument due to Jackson (1985); Jackson & Pargetter (1986) shows that the inference from $\mathcal{D}(\phi$ *and* $\psi)$ to $\mathcal{D}(\phi)$ is not intuitively valid either. Neither of the moves designed to deal with Ross' Paradox would help with this parallel but less-known problem.

Jackson & Pargetter (1986) describe the case as follows:

> Prof. Procrastinate is invited to review a book on which he is the only fully qualified specialist on the planet. Procrastinate's notable character flaw, however, is his inability to bring projects to completion. In particular, if Procrastinate accepts to review the book, it is extremely likely that he will not end up writing the review. In the eyes of the editor, and of the whole scientific community, this is the worst possible outcome. If Procrastinate declines, someone else will write the review — someone less qualified than him, but more reliable.

According to Jackson & Pargetter, (5.23) is true in this scenario:

(5.23) Professor Procrastinate ought to accept and write the review.

However, they judge that (5.24) is false here:

(5.24) Professor Procrastinate ought to accept.

3 Numerous proposals have been made with regard to Ross' paradox, a number of which suggest modifying modal semantics while others argue for a non-Boolean *or*. I do not want to discuss this literature in depth here; the solution that I will propose seems to me to be better motivated than changing the semantics of disjunction, in light of the numerous independent arguments that I give in favor of the scalar semantics for $\mathcal{D}$-modals in this chapter and the next.

Intuitively, the reason for this is that *ought* in (5.24) takes into account not only what will happen if Prof. Procrastinate accepts and writes — the best possible worlds — but also what is much more likely if he accepts, that he will not write. Since this is the worst possible outcome, the likelihood of its occurrence somehow outweighs the fact that the worlds in which he accepts and writes are optimal.

It is not too hard to construct similar cases with the other verbs we are interested in. Take *want*, for example:

> My co-worker Sam is a kind and friendly person, but he becomes loud and aggressive when he drinks. Unfortunately, if he is with people who are drinking, he almost always drinks a great deal and behaves badly.

Given this scenario, (5.25) could reasonably be true —

(5.25) I want Sam to come to my birthday party and stay sober.

while (5.26) is false:

(5.26) I want Sam to come to my birthday party.

As in the previous case, the intuitive basis for this judgment is the fact that if the wish in (5.26) is fulfilled then a bad outcome will probably result, but if the more specific wish in (5.25) is fulfilled a very good outcome will result. However, a quantificational semantics for *want* predicts that (5.25) entails (5.26); this is of course related to the upward monotonicity of the universal quantifier. As a result, it should be impossible for me to coherently hold the desire in (5.25) without also holding the desire in (5.26) if this analysis is correct.

Neither the Gricean account nor a modified semantics for *or* will help with Prof. Procrastinate's dilemma; instead, the example shows that the relevant operators simply are not upward monotonic. In other words, if Jackson & Pargetter are right, then $ought(\phi \text{ } and \text{ } \psi)$ does not entail $ought(\phi)$, even though $(\phi \text{ } and \text{ } \psi)$ entails ϕ. More generally, $\mathcal{D}(\phi \text{ } and \text{ } \psi)$ does not seem to entail $\mathcal{D}(\phi)$ in every case.

5.1.3 Chicken

Jackson (1985) gives an example which shows that the inference from $\mathcal{D}(\phi) \wedge \mathcal{D}(\psi)$ to $\mathcal{D}(\phi \wedge \psi)$ is not unrestrictedly valid either.

> Attila and Genghis are driving their chariots towards each other. If neither swerves, there will be a collision; if both swerve, there will be a worse collision (in a different place, of course); but if one swerves and the other does not, there will be no collision. Moreover if one swerves, the other will not because neither wants a collision. Unfortunately, it is also true to an even greater extent that neither wants to be 'chicken'; as a result what actually happens is that neither swerves and there is a collision. (Jackson 1985: 189)

As Jackson points out, all of the following are intuitively true in this scenario:

(5.27) a. Atilla ought to swerve.
b. Genghis ought to swerve.
c. It's not the case that Atilla and Genghis ought to both swerve.

Again it is not difficult to construct a similar example using *want*:

> I have two tickets for a lottery, ticket 1 and ticket 2. There are 1 million tickets in circulation. Five winning tickets will chosen, and are worth $\$400,000$ each. I live in a country with a rather odd tax system: people who earn less than $\$750,000$ in a year pay no taxes, while people who earn more than that are taxed 90% of their earnings. The numbers of the winning tickets will be chosen tonight.

In this case, it would not be unreasonable for me to simultaneously have all of the desires expressed by the three sentences in (5.28).

(5.28) a. I want ticket 1 to be chosen.
b. I want ticket 2 to be chosen.
c. I don't want tickets 1 and 2 to both be chosen.

According to the usual assumption that *ought* and *want* are universal quantifiers over some set of worlds (however determined), it should not be possible for the judgments in (5.27) and in (5.28) to be even coherent, much less intuitively true. This is a serious problem for this semantics: the inference from $\forall x[P(x)] \wedge \forall x[Q(x)]$ to $\forall x[P(x) \wedge Q(x)]$ is secure, as is its English counterpart. Of course, it is always possible to stipulate that the deontic or bouletic ordering somehow changes while we are reading these sentences, shifting the domain of quantification with it. To my knowledge, though, no one has even offered a theoretically or empirically motivated account of how or why such a shift would occur. On face it looks as if we simply have a counterexample to the analysis of *ought* and *want* as universal quantifiers over worlds.

5.1.4 Upshot

Taken together, Ross' Paradox, Professor Procrastinate, and Chicken present a major challenge to standard semantics for deontic modals and desire verbs. Furthermore, the Professor Procrastinate example cannot be patched up by changing the semantics of *or*, or by invoking Grice in any obvious way. Rather, the problem is that verbs of obligation, need, requirement, and desire are not in fact upward monotonic as the quantificational semantics predicts. Since both of the quantifiers standardly employed in semantic treatments of modality — $\exists$ and $\forall$ — are upward monotonic, and no plausible non-montonic quantificational treatment is on offer, these two puzzles strongly suggest that we have been wrong to assume that modals are quantifiers over possible worlds.[4]

4 A further argument against upward monotonicity is Nouwen's Puzzle, which is generated by the fact that sentences such as *The minimum required speed is 50mph* are not trivially false as quantificational semantics predicts (Nouwen 2010a,b). Nouwen shows that this puzzle is closely related to the monotonicity of both degrees and modals. However, the issues involved are rather intricate, and dealing with it here would require an extended digression. See Lassiter (2011a) for a detailed discussion and an argument that the semantics proposed in chapter 6 resolves this puzzle as well.

5.2 Problem Two: Fine-Grained Interactions with Probability

5.2.1 Medicine and Insurance

Viewed from a certain perspective, all three of the puzzles in the last section are instances of a broader class of puzzles for quantificational semantics of desire verbs and deontics. The problem, in a nutshell, is that obligation and desire interact with graded belief in a more fine-grained way than quantificational semantics can capture.

To see the connection, consider: how can *I want Sam to come to my birthday party* be false, even though in all of the best possible worlds according to my preferences, Sam does come (and stays sober)? This is possible because it is highly **likely** that, if Sam does come, he will get drunk and we will not be in one of the optimal worlds. Why is it that *Prof. Procrastinate ought to accept the review* is false even though the best worlds are ones in which he does accept the review (and writes it too)? Because, assuming that he accepts the review, there is a high **probability** that he will not write it, and we will find ourselves in one of the worst possible situations instead of one of the best. How can it be that Atilla ought to swerve and Genghis ought to swerve, even though it would be disastrous if both of them did? Because, as part of the story, we are told that it is highly **unlikely** that either of them will in fact do what they ought.

The lesson here is that, at least with respect to *ought* and *want*, probability matters. We cannot simply look at the optimal worlds (as Ideal-Worlds Semantics tell us) or even the best worlds that are epistemically possible (as Best Available Worlds Semantics would presumably recommend). Rather, we need some kind of mechanism which tells us how to weigh good and bad outcomes against each other, taking probability into account in some way.

This, at least, is Goble's (1996) conclusion regarding *ought*, and van Rooij's (1999) and Levinson's (2003) regarding *want*. Goble's story, simplified considerably, goes like this. Suppose that a doctor must choose which medicine to give to a critically ill patient, *A* or *B*. *A* has a small chance of producing a total cure, and a large chance of killing the patient. Meanwhile *B* will save the patient's life, but will leave him slightly debilitated.

What should the doctor do? Intuition suggests that the doctor ought to choose *B*, since *A* is very risky. However, standard quantificational semantics for *ought* unhesitatingly recommends choosing *A*, because all of the **best** accessible worlds in this scenario are worlds in which the doctor gives medicine *A*. This is evidently the wrong recommendation. Again, the problem relies on the fact that, in quantificational semantics, the fact that all of the best worlds are *A*-worlds is enough to render true *The doctor ought to give A*, regardless of the improbability of these worlds.

Levinson (2003) gives an example which makes the same point involving *want*. Consider the following four worlds, representing the possible outcomes I must consider in making insurance-buying decisions as a homeowner:

(5.29) w_1: I do not buy insurance and my home burns down
w_2: I do not buy insurance and my home does not burn down
w_3: I buy insurance and my home does not burn down
w_4: I buy insurance and my home burns down

It seems clear that, as a homeowner, if my house burns down, I would prefer to have fire insurance:

$w_4 > w_1$. I also do not like to spend money pointlessly, and so, assuming my home does not burn down, I prefer a state in which I do not buy insurance: $w_2 > w_3$. Finally, I prefer a state in which my home does not burn down to a state in which my home burns down, no matter what: $w_2, w_3 > w_1, w_4$. The only consistent preference order meeting these constraints is:

(5.30) $w_2 > w_3 > w_4 > w_1$

Assuming that all of w_1-w_4 are real epistemic possibilities, we should be able to conclude on the basis of (5.30) that

(5.31) I want not to buy insurance,

since all of the top-ranked worlds in (5.30) are worlds in which I do not buy insurance. Similarly, my financial advisor should feel comfortable telling me:

(5.32) You ought not to buy insurance.

But neither of these sentences expresses an inference which is appropriate to draw in this situation. Even though I would presumably prefer not to buying insurance if I **knew** that my house would not burn down, I may still want to buy insurance because I am uncertain whether it will. In particular, if I think that there is a decent chance that my house will burn down at some point, I may want to buy insurance even though all of the worlds in which I buy insurance are (by (5.30)) suboptimal. Or again, if my financial advisor thinks that there is a significant risk of fire, he would presumably advise me that I ought to buy insurance, knowing full well that the *best possible* worlds (relative to my financial health) are ones in which I do not.

These examples indicate that in considering what we ought to do, or what we want to do, we do not just look at the best possible worlds. Instead, whether or not I want to or ought to buy insurance will presumably depend on my judgment about how likely it is that my house will burn down, as well as factors such as the cost of insurance vs. the value of the house. Simply put, non-optimal worlds matter, and probability matters. Standard quantificational semantics for deontic modals and desire verbs are not able to capture these facts.

5.2.2 The Miner's Paradox

Another puzzle, the Miner's Paradox, demonstrates the relevance of uncertain information to obligation in a dramatic way. This puzzle was recently popularized by Kolodny & MacFarlane (2010), who attribute it to Regan (1980). I quote from Kolodny & MacFarlane (2010) (example numbers have been changed):

> Ten miners are trapped either in shaft A or in shaft B, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed.
>
> We take it as obvious that the outcome of our deliberation should be
>
> (5.33) We ought to block neither shaft.

Still, in deliberating about what to do, it seems natural to accept:

(5.34) If the miners are in shaft *A*, we ought to block shaft *A*.

(5.35) If the miners are in shaft *B*, we ought to block shaft *B*.

We also accept:

(5.36) Either the miners are in shaft *A* or they are in shaft *B*.

But (5.34), (5.35), and (5.36) seem to entail

(5.37) Either we ought to block shaft *A* or we ought to block shaft *B*.

And this is incompatible with (5.33). So we have a paradox.

As Kolodny & MacFarlane (2010) point out, if indicative conditionals are modeled as sentential connectives, the fact that (5.33)-(5.36) are consistent and true in the scenario at hand is inexplicable. We are able to make a little bit of headway by moving to what I called "Best-Available-Worlds" semantics earlier, i.e. a quantificational theory built around binary orders and armed with Kratzer's (1986) restrictor analysis of conditionals. The story invites us to imagine a ranking of worlds which is determined by the number of lives saved in each world: if n miners are saved in w and $m < n$ miners are saved in w', then $w > w'$. If this is the case, the best possible outcomes are of two types; either they are actually in *A* and we block *A*, or they are actually in *B* and we block *B*. In either case, all ten miners are saved.

Best-Available-Worlds semantics correctly predicts truth for the conditional sentences (5.34) and (5.35) in this scenario. Since the semantic effect of an *if*-clause is to restrict the deontic ordering relation to worlds in which the antecedent holds, the truth-conditions are informally:

(5.38) a. (5.34) is true iff all of the best worlds in which the miners are in fact in *A* are worlds in which we block *A* (and all ten are saved).

b. (5.35) is true iff all of the best worlds in which the miners are in fact in *B* are worlds in which we block *B* (and all ten are saved).

On this count, we are in the clear. However, the semantics fails to verify (5.33) (*We ought to block neither shaft*), because the best worlds in the unrestricted order are **not** worlds in which we block neither shaft. Worlds in which we block neither shaft are, without exception, worlds in which nine lives are saved, and as such are strictly dominated by worlds in which we block the shaft that they are in, whichever this happens to be.

It should be clear at this point that the problem of the Miner's Paradox is closely related to the puzzles involving insurance and medicine in the previous subsection: in each case, we have a situation where what the action that we intuitively *ought* to do is not the action that we take in the best possible worlds. Instead, what we ought to do is something that is globally sub-optimal but safe: we ought to block neither shaft (buy insurance, choose medicine *B*, etc.). Formally, the Miner's Paradox is very similar to the insurance and medicine cases, as well as other similar cases well-known in meta-ethics (cf. especially Jackson 1985; Jackson & Pargetter 1986 for several more).

What Kolodny & MacFarlane (2010) add to the mix here are two useful points. First, they consider the semantics of conditionals in interaction with deontic modals in some detail; their discussion focuses on a modified ("shifty") version of the restrictor analysis of conditionals which they present, and in particular on the fact that modus ponens is not unrestrictedly valid on this semantics. (Kratzer's analysis, though more or less standard in linguistic semantics, is not widely known in philosophy — for example, Bennett's (2003) otherwise thorough survey volume on conditionals does not mention or cite Kratzer.)

The second crucial feature of Kolodny & MacFarlane's account is their innovative but problematic notion of *serious information-dependence*. As I pointed out above, standard quantificational semantics for *ought* in combination with a restrictor analysis of conditionals predicts that (5.33) should be false, although intuitively it is true. Such an account also predicts that (5.39) should be true, though it is in fact clearly false:

(5.39) We ought to either block shaft A or block shaft B.

(5.40) is wrongly predicted true because in all of the best worlds (where ten miners are saved), we either block A (and they are in A) or we block B (and they are in B). Within a quantificational semantics for *ought*, it is very difficult to see how this result can be avoided.

Kolodny & MacFarlane propose to resolve this problem while retaining a quantificational treatment by allowing that the information that we have can actually influence the deontic ordering over worlds. Let $i_1, i_2, \ldots$ be variables over information states (sets of epistemically accessible worlds). Their crucial innovation is to allow that the deontic selection function d — which chooses a set of "best" worlds for *ought* to quantify over — can vary non-monotonically with information gain.

(5.40) A deontic selection function d is **seriously information-dependent** iff for some information states $i_1, i_2 \subseteq i_1$, there is a world $w \in i_2$ such that $w \in d(i_1)$ but $w \notin d(i_2)$.

Serious information-dependence allows that a world can be among the best worlds if we are in some information state i, but then fail to be among the best worlds once we acquire new knowledge: that is, $w \in d(i) \wedge w \notin d(i')$ for $i' \subseteq i$.

Serious information-dependence is actually ruled out *a priori* by a restrictor analysis of conditionals, since the only effect of a conditional on this analysis is to restrict the deontic ordering to pairs of worlds both of which satisfy the antecedent; it cannot change the ordering. What Kolodny & MacFarlane are proposing, essentially, is a semantics for conditionals where *gaining information can reverse the deontic ordering between pairs of worlds*: it may be that w_1 is better than w_2 (from our perspective in $w_@$), but, if we knew more, w_2 might turn out to be better than w_1. Their semantics for conditionals making use of selection functions has the effect of the restrictor analysis, but makes room for (5.40) essentially by remaining non-committal about the relationship between information gain and deontic ideality. This renders the crucial examples (5.33)-(5.36) logically compatible, as standard quantificational semantics cannot: it is possible, in principle, that worlds where we block neither shaft are deontically ideal with respect to an uninformed state, but fail to be ideal once we gain information about the miners' location.

Although Kolodny & MacFarlane's (2010) account works on a technical level, Charlow (2011) points out two methodological worries, arguing convincingly that we should seek a different route to achieve their results. The first issue is simply that nothing in their semantics explains why the crucial example *We ought to block neither shaft* is intuitively **true** in the scenario at hand. All that Kolodny & MacFarlane give us is a way to achieve **consistency**, and they say nothing about how information gain actually affects the deontic selection function. In effect, they block the reductio in (5.33)-(5.37) by seriously weakening the logic of deontic modals: in principle information can influence the deontic ordering in any arbitrary way. This is a considerable loss in predictive power relative to the restrictor analysis that we started out with. In that semantics, information gain was related to deontic ideality in a straightforward fashion, essentially via the $\upharpoonright$ operation restricting a binary order to a subset of its domain, and the account made strong predictions about the relationship between information gain and the truth-conditions of deontic modals.

The second problem is simply that serious information-dependence means allowing that information gain can reverse our preferences among fully-specified states of affairs; this is methodologically dubious for several reasons. As Charlow (2011) points out (where "Stability" is equivalent to "lack of serious information-dependence"):

> If a possibility has enough (with respect to other possibilities in a set p) good-making features, then it does not cease having enough good-making features with respect to a contraction of p. Contracting p, if anything, *reduces the possibility's competition*. Denying Stability is, at first glance, rather like denying that the best restaurant in Manhattan ... must also be the best restaurant in Soho [a neighborhood in Manhattan].

Charlow also points out that serious information-dependence is an implicit rejection of the choice-theoretic notion of "Independence of Irrelevant Alternatives", a "very basic requirement of rational choice" (Sen 1969: 384). Kolodny & MacFarlane's proposal is really a quite radical departure from standard assumptions, and presumably requires a compelling motivation rather than the (rather nonchalant) one-paragraph discussion that they give.

In addition, Kolodny & MacFarlane's independent justification for writing information-dependence into the semantics does not hold up to scrutiny. The single motivation that the provide for giving up Stability is the intuition that "a world in which both shafts are left open may be more ideal than one in which shaft *A* is closed relative to a less informed state, but less ideal relative to a more informed state" (Kolodny & MacFarlane 2010). At first glance this is clear enough, but its plausibility relies on intuitions about the ideality of *propositions* relative to an information state, not the ideality of individual *worlds*. The proposition that both shafts are left open is indeed more ideal in our uninformed state than the proposition that shaft *A* is blocked: this is just what (5.33) tells us. However, the relative ideality of a *world* (a fully specified state of affairs, with no remaining uncertainty) in which both shafts are left open, and a *world* in which shaft *A* is blocked *depends on the miners' location in that world*. Some worlds in we block *A* are better — ten lives are saved — and some are much worse: all die if we block the wrong shaft. Nevertheless, it is hard to see how anyone could deny that a world in which ten lives are saved is better than a world in which nine lives are saved, even if we did something risky and foolish in order to get there.[5]

Just as in the insurance and medicine cases discussed in the last subsection, what the Miner's Paradox demonstrates is that we need a semantics for deontic modals (and desire verbs) which allows that, in certain situations of uncertain decision-making, the action that we ought to take is sometimes one which is *guaranteed* to lead to a sub-optimal outcome. Intuitively, this will be the case when the action(s) which *might* lead to a globally optimal outcome also carry substantial risk: that is, the same actions might lead to a very bad outcome with some substantial probability. As the reader will have guessed, the diagnosis that I propose is very simple: the problem is generated by the false assumption that *ought* and other modals express quantification over a set of "best" worlds.

Three more specific desiderata emerge from the considerations of this section. First, we would like to have a semantics which gives us a concrete account of why the sentences in (5.33)-(5.36) are all true — not just consistent — in the scenario at hand. Second, we want a semantics that does so without weakening the logic to the point that anything goes in the interaction between information and obligation, as Kolodny & MacFarlane's does. Third, we would like to do so without making use of the philosophically and methodologically problematic notion of serious information-dependence. The fact that it is necessary to abandon Stability and weaken the semantics in a quite counter-intuitive way in order to deal with the Miner's Paradox within a quantificational theory of modality is an indication of how deeply problematic information-sensitivity is for such theories.

In contrast, the scalar semantics that I give in the next chapter gives us a simple and straightforward account of the Miner's Paradox with all three of the desired properties: it gives a simple and compelling explanation of why the crucial examples are true in the case at hand; it makes strong predictions about how information gain relates to the semantics of deontic modals; and it does allows changing information to influence the deontic ordering over *propositions*, but not over *worlds*. On this approach, the goodness of a proposition is calculated by considering both the goodness of the

5 I suppose that certain ethical systems could allow us to avoid this conclusion: for example, we might think that blocking one of the shafts in ignorance of the miners' location would violate an important moral norm (do what you ought to do), and that the fact of being in a world in which this core ethical principle has been violated outweighs any possible positive result. This is reminescent of the moral theory of Kant (1797), who famously claimed that it would be wrong to lie to a murderer who has come to your house to kill your friend, regarding your friend's whereabouts. We might think, along similar lines, that a state in which some ethical norm has been violated (e.g., we acted recklessly by blocking one of the shafts despite not knowing that everything would turn out OK) is bad simply by virtue of the fact that an ethical norm has been violated.

This approach might be viable, but it would involve building some fairly heavy-duty ethical commitments in our semantics. In addition, unless some independent account of how obligations are determined can be given, it is quite circular to claim that (a) the worlds where we block neither shaft are best, given our current information, because we do what we ought to do in those worlds, and that (b) *We ought to block neither shaft* is true because that is that what we do in all of the best worlds relative to our information.

Even if some such account could be given, though, the tactic would not seem to work at all for structurally similar cases involving desire verbs, such as the insurance puzzle. Surely the goodness of worlds relevant to expressions of desire is determined by the consequences of my actions, and not by whether I acted in accordance with what I thought best in a limited state of information: that is, I prefer to be in a world in which I get what I want most to one in which I get something I want less, even if I have made a risky or even a stupid decision in that world. The problem in making practical decision, both in the Miners' Paradox and the insurance puzzle, is in knowing which actions will lead to which states with which probabilities, not in the ranking of states themselves.

worlds it contains and the probability that these worlds will be realized. In other words, the ordering over worlds is not information-sensitive; instead, it is the mapping from a deontic ordering over worlds to a deontic ordering over propositions which is information-dependent. This feature allows us to maintain the methodologically desirable stability constraint while also capturing the data.[6]

5.3 Problem Three: Gradability and Scalarity

We have already seen considerable evidence that at least some modals are gradable. Here I will discuss two sorts of evidence of this type with an eye to deontics and desire verbs specifically: evidence that they take part in gradability and comparison, and evidence that they come in minimum, relative, and maximum-standard/high-degree varieties just like gradable adjectives.

In this section I give a number of naturally-occurring examples of *need, want, require*, and other deontic and desire verbs occurring with degree modifiers and in comparatives, and discuss why this is a challenge to quantificational semantics for modals.

5.3.1 The Data

Recapping discussion in chapter 3 briefly, recall that the modals that occur most freely in degree modification and comparison structures are main verbs (*require, want*, and main verb *need*) and adjectives (*good, permissible, obligatory, desirable*). *Should*, *must*, *may*, and auxiliary *need* are more restricted, although *should* in particular does occur in comparatives and with degree modifiers in some cases. *Ought* — which behaves syntactically like a main verb in some ways and like an auxiliary in others — is intermediate in this way as well, showing up more frequently in degree constructions than the auxiliaries but less frequently than the main verbs and adjectives. The correlation between syntactic category and participation in modification and comparison may suggest that the limited gradability of auxiliaries is due to syntactic rather than semantic factors.

The verbs *require, need*, and *want* occur frequently in comparatives, as does *ought*. Probably the most common degree modifier with these items, as with other main verbs such as *like*, is the intensifier *very much*. Although corpus searches for these strings return many false hits and cases in which it is not clear whether the modal or another item is involved, the examples below have been chosen as clear examples of cases in which the only plausible interpretation requires the deontic or desire verb to be graded. For example, (5.41a) — taken from a moral philosophy paper — is clearly intended to compare the degree to which Constance ought to help George (to whom she has

6 A brief note on Charlow's (2011) positive proposal: although his criticism of Kolodny & MacFarlane (2010) is compelling, I have doubts whether the proposed alternative gets the facts right in several places. Most notably, Kolodny & MacFarlane's first example (5.33) (*We ought to block neither shaft*) is intuitively true in this scenario, but Charlow's official proposal makes it false, deriving only the weaker *It's not the case that we ought to block A or block B*. (Charlow acknowledges this in a footnote and suggests a fix, but the account strikes me as rather *ad hoc*: see his fn.28.) Charlow also predicts that *We must not block A or block B* is unambiguously false in this scenario, but it strikes me as pretty clearly true(*We must not block A or block B. That would be reckless, since we don't know where they are*).

More generally, the positive account of Charlow (2011) continues to treat deontic modals as quantifiers and so encounters most of the other problems discussed in this chapter. It also relies heavily on the proposal of von Fintel & Iatridou (2008), for which I raise some independent problems in §5.3.4 below and give a quite different alternative in ch.6. [Modified to correct an error, 12 Jan 2012. Thanks to Nate Charlow for discussion.]

previously done wrong) to the degree to which Constance ought to help others who she has not harmed.

(5.41) *Ought*:

a. [O]nce the damage is done, Constance ought to help George — or, at least, she ought to help him more than she ought to help anyone else similarly situated.[7]

b. A war between Great Britain and the U.S. ought very much to be deprecated.[8]

(5.42) *Good*:

a. I think it is very good that Stephen King came out with an honest opinion. Not too many celebrities would do that.[9]

b. It is better for children to grow up in the countryside than in a big city.[10]

(5.43) *Require*:

a. The members of a literary group are required to have a blazer, more than they are required to have ever actually read a book.[11]

b. Thus, you are very much required to have a good credit record to prove yourself as a reliable client to the insurance providers.[12]

(5.44) *Want*:

a. [M]any library officials want more to intimidate than to really change an institutional culture that has squelched feedback.[13]

b. I am an American and I want very much to travel to Cuba.[14]

(5.45) *Need*:

a. [W]e need very much to have additional requirements until such time as Mexican carriers meet the standards that prevail in the USA. [15]

b. But really emails need to be timely more than they need to be amazing.[16]

c. [M]en need to disparage women more than women need to disparage men. (Horney 1942)

7 From Driver (1997: 853). The example is actually ambiguous in isolation between the relevant reading and one in which what is being compared is the amount of helping that Constance owes to George vs. other people. However, the context of this paper makes it clear that it is the degree of obligation that is under discussion.
8 Army/Navy Chronicle, 27 February 1939.
9 http://starrynite45.xanga.com/691585421/stephen-king-blasts-on-stephanie-meyer-shes-not-very-good/
10 http://www.urch.com/forums/ielts/129157-it-better-children-grow-up-countryside-than-big-city.html
11 http://www.thefashioniste.com/28.html
12 http://www.play-it-forward.org
13 Massachusetts Board of Library Commissioners.
14 http://www.wordtravels.com/forum/comments.php?DiscussionID=1531&page=1
15 *Congressional Record*, 7/25/2001.
16 http://blog.penelopetrunk.com/2005/12/26/is-your-email-out-of-control-test-yourself/

Good examples of comparison and gradability with deontic auxiliaries are harder to find, but I have located a few clear examples with *should*.

(5.46) *Should*:

a. I don't think he [UFC fighter Phil Davis] should be compared to Rosholt as much as he should be to Houston Alexander.[17]

b. If you desperately need to change an old post then PM one of the moderators ... This should very much be considered the exception though. The normal edit window should usually be enough.[18]

5.3.2 The Problem

The fact that these deontic modals and desire verbs are semantically gradable is a serious problem for standard quantificational semantics of modality. As we saw in some detail in chapters 1-2, semantics for gradable expressions typically proceeds by creating a partial order over objects of the appropriate type (with or without the intervention of degrees). In straightforward cases, we then look for a threshold value in this order with which to compare the object in question, and return True just in case the object in question exceeds the threshold value in the relevant order. As we saw, there is some variability among items in whether the threshold is typically associated with a minimum, a maximum, a context-dependent mid-range value, or a context-dependent high degree, and degree modifiers, comparatives, and equatives are generally supposed to be operators which manipulate this threshold value.

Ideal-Worlds analyses of deontic and bouletic modality look nothing like this, of course: instead context simply determines an unordered set of worlds over which we quantify. Best-Available-Worlds analyses such as Kratzer's are somewhat more like analyses of gradable expressions of other categories in that they make use of partially ordered domains. However, the rest of the semantics is quite different: the role of the partial order over worlds in Kratzer's theory is (at least in the simplest case) to provide a more flexible way of arriving at the set of "best" worlds which can be existentially or universally quantified over. (If there are not any best worlds, we look for a set of "good enough" worlds beyond which there is no variation with respect to the value of the proposition in question.)

In the details, the quantificational approach to modality looks quite different from semantics for gradability in other categories. For example, to evaluate *Sam is tall* we start by establishing an ordering over the relevant x of which *x is tall* could be predicated, and then comparing Sam's position in this order to the position of the others. In contrast, for Kratzer *Mary should leave* is not evaluated by establishing an order over relevant ϕ for which *should* ϕ could be predicated, and then comparing the position of *Mary leaves* in the order to that of other ϕ. Instead, we look at an order over objects one type lower — worlds, not propositions — and equate the truth or falsity of *Mary should leave* to the truth or falsity of universal quantification over the set of worlds which exceed some independently established threshold value (given by the ordering $\succcurlyeq_{\mathbf{g}(w)}$, in Kratzer's theory).

17 http://www.bloodyelbow.com/2009/12/17/1205767/ufc-signs-hot-prospect-phil-davis

18 http://www.adventuregamers.com/forums/showthread.php?p=546770

5.3.3 The Exceptions

There are two notable exceptions to the observation that deontic modals and desire verbs are gradable: it is difficult to find examples of gradability and comparison with *may* and *must*. Interestingly, these are the same items for which we had this difficulty in the case of epistemic modals in chapter 3. As in that case, we are faced with two possible analyses. First, we might conclude that *may* and *must* really are quantificational, while the other deontics and desire verbs are scalar. Second, we might surmise that there is an independent grammatical motivation for the apparent ungrammaticality of the examples below:

(5.47) a. ?* You may leave more than John may.
b. ?* Sam may very much be at home.

(5.48) a. ?* Sam must stay home more than Bill must go out.
b. ?* Sam must very much stay home.

As before, I am inclined to push the scalar semantics as far as it can go, but I must acknowledge the possibility that there may turn out to be important semantic differences among modal expressions. However, there are two reasons to favor the scalar approach as a general theory and ascribe (5.57) and (5.58) to syntactic restrictions relating to the auxiliary position.

First, the fact that most modal expressions are gradable means that the scope of a quantificational theory of modality is highly limited — indeed, if we endorse a non-scalar quantificational semantics for *may* and *must* these may well turn out to be the only items for which it is appropriate. It strikes me as undesirable to maintain the heavy-duty semantic apparatus required for quantificational semantics for modality just for a few items.

More importantly, *may* and *must* participate in a number of the semantic and pragmatic puzzles which are the subject of this chapter, notably the arguments for non-monotonicity and for information-sensivity. These puzzles pose a problem for quantificational semantics for *must* and *may* just as much as they do for *want, need*, and other items. Since the scalar alternative that I will advocate resolves these issues, this provides indirect evidence that the restrictions on degree modification and comparison with *may* and *must* are due to grammatical restrictions rather than indicating that these items have a non-scalar semantics.

Summing up, there is clear evidence that *ought, need, want, require, obligatory, permissible* and *should* are gradable. Quantificational semantics is not equipped to deal with these facts in a way compatible with the best available semantics of gradability in general. Further, even though the deontic modals *may* and *must* do not appear in degree constructions to my knowledge, this does not show conclusively that their semantics is not scalar, and there is indirect evidence suggesting that they are.

5.3.4 Intermediate-Strength Modals and Neg-Raising

Quantificational semantics for modals generally allows us to make two types of distinctions among modal expressions. They can be either existential or universal quantifiers, and they can have different underlying orders.[19]

However, modals show more than two distinctions of logical strength, and resemble the three-way typology of gradable adjectives considerably. In particular, as discussed by Horn (1972, 1989); Copley (2006); von Fintel & Iatridou (2008), *ought* and *should* are semantically weaker than *must* and *have to*:

(5.49) a. You ought to wash your hands, but I guess I can't say that you have to.
b. # You have to wash your hands, but I guess I can't say that you ought to.

The best-known account of this fact in recent formal semantics is due to von Fintel & Iatridou (2008). As they point out, Kratzer's official semantics does not have the resources to deal with this phenomenon. They note, following Sloman's (1970) "'Ought' and 'Better' ", that the difference in meaning between *ought/should* and *must/have to* can be summarized as follows:

(5.50) a. *Ought/should* "picks out the best means without excluding the possibility of others";
b. *Must/have to* "implies that no other means exists" (Sloman 1970: 390-1).

In von Fintel & Iatridou's (2008) proposal, this difference is captured by adding a second ordering source to Kratzer's semantics. Basically, the idea is that *must* and *have to* quantify over a larger set of "best" worlds than *ought* and *should*: $must(\phi)$ means that all of the best worlds (according to the first ordering source) satisfy ϕ, while $ought(\psi)$ means that, among the best worlds according to the first ordering source, all of the best ones according to the **second** ordering source satisfy ψ.

This analysis accounts for Sloman's observation neatly. $Must(\phi)$ is true only if there are no $\neg\phi$-worlds among the best worlds (according to the first ordering source), but $ought(\psi)$ can be true even if there are some $\neg\psi$-worlds among these, as long as the ψ-worlds are better than the $\neg\psi$-worlds according to some additional measure. It also predicts the fact that $must(\phi)$ asymmetrically entails $ought(\phi)$: both are universal quantifiers, but the set of worlds that *ought* quantifies is a subset of the set of worlds over which *must* quantifies. Logically, then, the relationship between *must* and *ought* is similar to the relationship between *everyone* and *everyone in this room*: if everyone is happy, then everyone in this room is happy. Likewise, if $must(\phi)$ is true (because ϕ holds throughout the best worlds according to the first ordering source), then $ought(\phi)$ is true as well (because ϕ holds throughout the subset of these worlds which are also best according to the second ordering source).

However, there are some problems. Theoretically, the analysis is somewhat cumbersome: the primary motivation for adding a second ordering source to Kratzer's theory — which already places a heavy burden on "context" to provide us with fairly complicated theoretical machinery — is to explain how there can be further distinctions among modal strengths without giving up the assumption that *ought*, *should*, *must*, and *have to* are all appropriately modeled by $\forall$. (A secondary motivation for the second ordering source is von Fintel & Iatridou's observation that many languages

19 Other quantificational forces are in principle possible — for instance, Copley (2006) attributes to Horn (1972, 1989) the idea of treating *ought* and *should* as MOST-quantifiers over accessible worlds. This approach is really a non-starter, though: as von Fintel & Iatridou (2008) point out, counting worlds does not seem likely to give us a plausible modal semantics. Horn does not really advocate this idea, though; he merely notes that *ought* and *most* occupy similar mid-range positions on their respective scales. Incidentally, Goble's (1996) analysis, which I will endorse in a slightly modified form in the next chapter, makes sense of the semantic similarity between *ought* and *most* noted by Horn without engaging in dubious world-counting activities.

express deontic *ought* by combining deontic *must* with counterfactual morphology. They suggest that the counterfactual morphology brings in the second ordering source; but they have no story about how counterfactuality is relevant, and no account of which propositions would be in the second ordering source or why.)

Empirically, von Fintel & Iatridou's account encounters difficulty because it fails to capture the connection between modal strength and neg-raising pointed out by Horn (1989): *ought* and *should* participate in neg-raising while *must* and *have to* do not.

(5.51) a. I don't think you ought to leave. $\rightsquigarrow$ I think you ought to stay.
b. I don't think you should do that. $\rightsquigarrow$ I think you should not do that.

(5.52) a. I don't think Mary has to clean her room. $\not\rightsquigarrow$ I think Mary has to not clean her room.
b. I don't think Sam must wash the dishes. $\not\rightsquigarrow$ I think Sam must not wash the dishes.

As Horn (1989) also points out, *ought* and *should* pattern in this respect with both the quantifier *most* and what are now known as relative-standard adjectives.

(5.53) I don't think most of my friends would like this music. $\rightsquigarrow$ I think most of my friends would dislike this music.

(5.54) a. I don't think Mary is happy. $\rightsquigarrow$ I think Mary is unhappy.
b. I don't think Topeka is far. $\rightsquigarrow$ I think Topeka is close.

In contrast, *must* and *have to* pattern with the universal quantifiers *all* and *every* as well as high-degree and maximum-standard adjectives, which do not lead to an inference that the mirror image holds: compare (5.52) to (5.55) and (5.63).

(5.55) a. I don't think Mary ate all of the cookies. $\not\rightsquigarrow$ I think Mary ate none of them.
b. I don't think Sam likes every girl in his class. $\not\rightsquigarrow$ I think Sam likes none of them.

(5.56) a. I don't think Jaffrey is enormous. $\not\rightsquigarrow$ I think Jaffrey is tiny.
b. I don't think the glass is full. $\not\rightsquigarrow$ I think the glass is empty.

For von Fintel & Iatridou (2008), the strong and intermediate-strength modals alike are modeled as universal quantifiers, and the only difference is the restrictor. It is a mystery on this account why *ought* and *should* would pattern with *most* rather than *all* and *every* with respect to neg-raising. The neg-raising data suggest instead that the difference in logical strength among these items is connected either with the difference between the quantifiers *most* vs. *all*, or with the relative/maximum-standard adjective distinction. Since the former suggestion involves counting worlds, though, it seems unattractive as discussed above.

To be sure, this point does not stand as a knock-down objection to von Fintel & Iatridou's (2008) account on its own. However, in combination with other logical and empirical arguments discussed in this chapter which cast doubt on the assumption that deontic modals are quantifiers, their analysis looks like a rather *ad hoc* attempt to force a recalcitrant data set into the assumption that all modals express either $\exists$ or $\forall$. If another semantics could capture these relationship between *must/have to* and *ought/should* in a simple and well-motivated way without treating them all as universal quantifiers, this would be a strong point in its favor.

The semantics I will propose for these items in the next chapter treats them as scalar items, closely related to relative-standard and high degree adjectives. In addition to accounting for the neg-raising facts and the difference in logical strength, Sloman's (1970) observation about the difference in meaning between *must* and *ought* summarized in (5.50) is built into the semantics in a simple fashion. As it happens, the semantics that I will give is much closer to Sloman's original proposal than the quantificational semantics of von Fintel & Iatridou: the differences in meaning between *must* and *ought* are accounted for by treating both as semantically related to *better*.

5.4 Problem Four: Deontic and Bouletic Comparatives

For theorists who endorse Best-Available-Worlds semantics for deontic modals and desire verbs, the obvious analysis of deontic and bouletic comparatives is to make use of the binary order which underlies this theory. Kratzer (1991) makes an explicit proposal along these lines. This section discusses two empirical shortcomings of this approach; the first is specific to Kratzer's theory, while the second afflicts any straightforward attempt to give semantics to deontic and bouletic comparatives in this way.

5.4.1 Kratzer Generates Too Many Incomparabilities

One approach to the gradability of deontic modals and desire verbs might be to modify Kratzer's theory in order to make it look more like a degree semantics, as Portner (2009) suggests and as we discussed in chapter 3. The idea is to identify the set of degrees with the set of equivalence classes of propositions under Kratzer's (1991) derived order

(5.57) $\succcurlyeq^{s}_{\mathbf{g}(w)} = \{(p,q) \mid \forall w' \in q\ \exists w'' \in p : w'' \succcurlyeq_{\mathbf{g}(w)} w'\}$

which, recall, is defined in terms of the order $\succcurlyeq_{\mathbf{g}(w)}$ over worlds.

(5.58) $\succcurlyeq_{\mathbf{g}(w)} = \{(w',w'') \mid \{p : p \in \mathbf{g}(w) \wedge w' \in p\} \supseteq \{q : q \in \mathbf{g}(w) \wedge w'' \in q\}\}$

Since $\succcurlyeq_{\mathbf{g}(w)}$ is built around a superset relation, it is a quasi-order: reflexive and transitive, but not connected. $\succcurlyeq^{s}_{\mathbf{g}(w)}$ inherits this feature. By taking the reduction of $\succcurlyeq^{s}_{\mathbf{g}(w)}$ to equivalence classes, we arrive at a partially ordered set of degrees — reflexive, transitive, and antisymmetric — rather than a linear order.

While mere partial orders may be plausible for other types of scales (cleverness and bigness, for instance; see Bierwisch 1989; van Rooij 2010 for discussion), the fact that degrees of modality form a mere partial order in Kratzer's theory is a particular problem for the semantics of modality. The way that Kratzer constructs the ordering using a subset relation necessarily predicts that any two propositions which violate disjoint sets of norms will not be deontically or bouletically comparable. While I already pointed out the difficulty of this prediction in the case of epistemic modals in chapter 3, it is worth returning to it here because the descriptive failing is even more acute with deontic modals and desire verbs.

For a simple example, consider what happens if the ordering source $\mathbf{g}(w)$ includes both the propositions *Norm1=There is no trespassing* and *Norm2=There is no murder*. Suppose the modal base contains (among others) world w_1, where someone trespasses but no one commits murder, and world w_2, where someone commits murder but no one trespasses. Since w_1 violates *Norm1* but not *Norm2*, while w_2 violates *Norm2* but not *Norm1*, it follows from (5.58) that neither $w_1 \succcurlyeq_{\mathbf{g}(w)} w_2$ nor $w_2 \succcurlyeq_{\mathbf{g}(w)} w_1$: they are deontically incomparable.

Now, given the way that the ordering over propositions is derived in (5.57), it does not follow immediately that murder and trespassing are deontically incomparable. To get this result we need to make sure that the modal base is rich enough and the propositions in the ordering source are logically independent. Nevertheless, unless by accident the modal base and ordering source are limited in some peculiar ways[20], Kratzer's theory predicts that both (5.59a) and (5.59b) should be without truth-value if *p is better than q* is true iff $p \succ^s_{\mathbf{g}(w)} q$:

(5.59) a. It is better to trespass than it is to murder.
b. It is better to murder than it is to trespass.

Essentially, because the ordering on worlds is built on a superset relation, Kratzer's theory makes it impossible in principle to make deontic comparisons unless there happens to be a (contextual) entailment relation between the norms which the propositions being compared violate. Judging by (5.59), this seems to be an excessive restriction: in a society like ours in which both trespassing and murder are prohibited, it is still possible — indeed obviously correct — to assign truth to (5.59a) and falsity to (5.59b).

For the same reasons, a semantics for bouletic comparatives built on Kratzer's theory would predict that I cannot want ϕ more than I want ψ if I would have to compromise different and logically independent desires in order to get them. Suppose my desires include having a sandwich for lunch and going to a movie this afternoon. As long as there are worlds in the modal base in which I have a sandwich but don't go to a movie, worlds in which I go to a movie but don't have a sandwich, worlds in which I do both, and worlds in which I do neither, (5.60) is predicted to be truth-value-less:

(5.60) I want to go to a movie more than I want to have a sandwich for lunch.

This is a very restrictive set of assumptions to build into the semantics of deontic modals and desire verbs, to put it mildly. I am quite sure that I can want to go to a movie more than I want to have a sandwich, even if the best possible scenario is one in which I do both.

5.4.2 Lack of Quantitative Information

The idea of building a degree semantics on top of Kratzer's theory also encounters some of the same problems involving quantitative comparisons that we saw for epistemic modals in chapter 3. Recall

20 That is, unless for each $p, q \in \mathbf{g}(w)$, p either contextually entails or is contextually entailed by q. (p "contextually entails" q in the relevant sense just in case every p-world in the modal base is a q-world.) This would be a rather strange scenario with no obvious applicability, though. I will assume that we are interested in cases in which the propositions in the ordering source are independent relative to the modal base.

that Kratzer's $\succcurlyeq^s_{\mathbf{g}(w)}$ is a pre-order; when we define a Kratzer-structure $\mathcal{K}$ using this relation we end up with a very weak scale, even weaker than an ordinal scale (the weakest scale type standard in RTM). It also produces a good deal of incomparability — which, translated into RTM terms, means that non-increasing transformations of admissible μ are permitted in many cases.

The use of a scale which allows for non-increasing transformations leads to the now-familiar prediction that *x is at least as P as y* will also not be interpretable in many cases (since it is true in some μ and false in others). We have already seen that this prediction is problematic; however, several of the properties that are shared with stronger ordinal scales are also problematic. Suppose, as above, that ϕ *is better than* ψ is true iff $\phi \succ_{\mathbf{g}(w)} \psi$, which implies that $\mu(\phi) > \mu(\psi)$ in all $\succ_{\mathbf{g}(w)}$-admissible μ. How about (5.61)?

(5.61) ϕ is twice as good as ψ.

As usual, (5.61) will be true iff, under all admissible μ, $\mu(\phi) = 2 \times \mu(\psi)$. Being weaker than an ordinal scale, our Kratzer-structure $\mathcal{K}$ will obviously not be able to make sense of (5.61). Here, unlike the case of epistemic modals discussed in chapter 3, this prediction seems to be correct.

However, $\mathcal{K}$-structures are too weak to support **any** quantitative comparisons, including ones as uninformative as (5.62):

(5.62) ϕ is much better than ψ.

Supposing that ϕ is indeed better than ψ, we still have the problem that all monotone increasing transformations are admissible. As a result, no matter how small the difference in goodness required to make ϕ *much* better than ψ, there will be admissible measure functions for which the difference between ϕ and ψ fails to exceed this threshold.

The prediction, then, is that (5.62) should be as hard to make sense of as (5.61). It seems clear that this prediction is wrong: not only is (5.63) not uninterpretable, it is even true.

(5.63) It is much better to give your money to charity than to gamble it on sports.

If we limit our attention to the scale types that are standard in RTM, it is clear that, in order to get sentences like (5.63) to come out as interpretable, deontic modals (and desire verbs) need to be associated with orderings on propositions that are at least as structured as interval scales. As a result, the criticisms leveled at Kratzer's theory here also affect other proposals to interpret deontic modals and/or desire verbs with respect to binary orders, e.g. Lewis (1973); Heim (1992); Villalta (2008). All of these authors make use of ordinal or even weaker scales, which are do not carry any quantitative information.

A final point regarding the attempt to devise a degree semantics built on Kratzer's theory — not so much an objection as an observation — is that implementing this approach in enough detail to account for the data involving gradability adduced here would basically mean abandoning quantificational semantics for modality to a considerable degree. For example, we might well be able to account for the gradability of, say, *need* and *want* by treating *I need/want* ϕ as true just in case ϕ exceeds some highish threshold value in the relevant Kratzerian order. This would make *need* look much like a maximum-standard or high-degree adjective; perhaps the same trick would work for the other gradable modals (putting other problems aside for the moment).

While I am entirely in favor of something like this move, it should be clear that implementing it would constitute an abandonment of the core ideas of the quantificational theory of modality: *need* and other items would not be quantifiers over worlds any longer. Furthermore, this modification in itself would not help with most of the other problems including excessive incomparabilities, lack of quantitative information, and the evidence reviewed in this chapter for non-monotonicity and the need for a more fine-grained interaction with factual information. The scalar semantics that I will propose in the next chapter will have the general character of the move toward a scalar analysis just discussed, but will make much more radical changes to the method of ordering propositions which allow us to avoid these problems.

5.5 Problem Five: Deontic Conflicts

I've promised to pick up my sister at the airport, and I've also promised to go to my friend's concert. I've just discovered that my sister is arriving during the concert. What should I do? According to many moral theorists, this is a situation in which I have a conflict of duty: as a general moral rule I *ought* to fulfill my promises, from which it follows that I *ought* to pick up my sister and I *ought* to go to the concert. Unfortunately, I can't do both.

Deontic conflicts are real, and yet, if standard quantificational theories are correct, it is an *a priori* truth that they cannot exist. Suppose, with many in deontic logic, that *ought A* is true if and only if some world satisfying *A* is better than any world satisfying *not A*. van Fraassen (1973) calls this the "axiological thesis", and it a version of what I will call "Possibilism" in the next chapter. (I will come back to the relationship with Kratzer's semantics momentarily.) One of the several ways to prove that there are no deontic conflicts is:

> [S]uppose *A* and *B* are incompatible. Then if it ought to be the case that *A*, higher values attach to some outcomes satisfying *A* than to any that satisfy *not A*. But, because of the assumed incompatibility, all outcomes that satisfy *B* satisfy *not A*. Hence it is better to opt for *A* than for *B*. So, whenever *A* and *B* are mutually incompatible, it cannot be that both ought to be the case — either we ought to opt for *A*, or we ought to opt for *B*, or the matter is indifferent (morally indifferent, that is). (van Fraassen 1973: 8)

van Fraasen gives a second way to derive the incompatibility, which is also a valid argument on standard assumptions. He phrases it in terms of *ought(A)* and *ought(not A)* both being true, but the argument also extends to the case of *ought(A)* and *ought(B)* for incompatible *A* and *B*, because quantificational semantics validates the argument *ought B, B implies not-A, therefore ought not-A*.

> It is asserted that "it ought to be the case that" implies "it is permitted (morally unobjectionable) that it be the case that", and that similarly "ought not" implies "not permitted". But then it follows that if it ought to be the case that *A*, then it is permitted that *A*, and hence it cannot be true that it ought not to be the case that *A*. Hence, *A* and *not A* can never both be such that they ought to be the case. (van Fraassen 1973: 12)

A third way to derive this result is by using the widely accepted principle that *ought* implies *can*, and the fact that in quantificational semantics *ought* A and *ought* B together imply *ought* A *and* B. Quite clearly, if A and B are incompatible, you cannot bring them both about. It follows by modus tollens that it is not the case that *ought* A *and* B, and so it is not true that both *ought* A and *ought* B.

It seems strange, however, to rule out the possibility of a genuine conflict of duty simply because our semantic theory cannot make sense of it; on face this looks like a problem with the semantic theory, rather than the concept of a moral conflict. van Fraassen (1973: 8) argues along similar lines that

> If the axiological thesis is accepted, then certain tenable ethical positions are ruled out. From this ... I conclude that the axiological thesis is itself an ethical doctrine, not a thesis of metaethics. (And if this is so, deontic logic should not be founded on it ...)

Kratzer (1991: 647-9) has an interestingly different perspective on the problem, although her account is ultimately problematic as well. In her semantics, obligations can be thought of as propositions which are included in the relevant deontic ordering source $\mathbf{g}(w)$: "You pick up your sister" and "You go to your friend's concert" are added to $\mathbf{g}(w)$ when you make the relevant promises, and from here they exert their semantic influence by affecting the pre-order $\geqslant_{\mathbf{g}(w)}$ in terms of which modal expressions like *ought* are defined.

Recall that $w \geqslant_{\mathbf{g}(w)} w'$ if and only if every proposition in $\mathbf{g}(w)$ which contains w' also contains w, and that the semantics of *ought* and *should* is given by universal quantification over $\geqslant_{\mathbf{g}(w)}$-undominated worlds. When there are inconsistencies in the modal base, Kratzer's semantics does not yield a contradiction: even though there are no worlds in which both of these requirements are fulfilled, there are $\geqslant_{\mathbf{g}(w)}$-undominated worlds in which you pick up your sister (ϕ), and $\geqslant_{\mathbf{g}(w)}$-undominated worlds in which you go to the concert (ψ), although there are of course none in which you do both.

Kratzer (1991) takes this as one of the main arguments in favor of her semantics for modality: unlike standard deontic logic, the system does not collapse when there are incompatible requirements, and so does not render everything permissible. In one sense, this theory is indeed able to model incompatible requirements, since the propositions in the ordering source can be inconsistent. In another sense, though, this characterization is misleading: in Kratzer's theory it is still a contradiction to say that *ought* ϕ and *ought* ψ are both true if ϕ and ψ are inconsistent.

The reason is that, given that the ordering source contains inconsistent propositions ϕ and ψ, the set of "best" worlds over which *ought*, *should*, *may*, etc. quantify contains worlds of each type: ϕ-worlds, $\neg\phi$-worlds, ψ-worlds, and $\neg\psi$-worlds. As a result both of the following are predicted to be true in the situation at hand:

(5.64) a. It's not the case that you should pick up your sister at the airport.
$(\neg\forall w' \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) : w' \in \phi)$

b. It's not the case that you should go to your friend's concert.
$(\neg\forall w' \in \mathbf{BEST}(\mathbf{f}(w))(\mathbf{g}(w)) : w' \in \psi)$

But this is wrong: the whole problem is that you *should* pick up your sister, and you *should* go to the concert.

Even though Kratzer's theory does not collapse into absurdity in case of conflicting propositions in $\mathbf{g}(w)$, it still does not give us the correct result: it remains true that, if ϕ and ψ are incompatible, then necessarily at least one of *ought* ϕ and *ought* ψ is false. If both ϕ-worlds and ψ-worlds appear in $\bigcap \mathbf{f}(w)$, the situation is even worse: both $ought(\phi)$ and $ought(\psi)$ are false. The entire problem of moral conflict is that there are sometimes conflicting *ought*-statements which are simultaneously **true**; what Kratzer gives us instead is a semantics which renders them both **false**. As a result (5.64a) and (5.64b) are unacceptable consequences of the theory. We wanted to find a logic for *ought* and *should* that makes it possible to model moral conflicts, but what we have here is one which simply ignores them.

5.6 Summary and Preview of Chapter 6

In this chapter we saw Ideal- and Best Available-Worlds semantics for deontic and bouletic modality, and surveyed a number of problems which affect theories of this type. Best Available-Worlds approaches are superior to Ideal-Worlds theories, but still encounter what I believe to be insuperable problems, including problems involving

- **Monotonicity**: Ross' Paradox, Chicken, and the Professor Procrastinate puzzle call into question the upward monotonicity of deontic modals and desire verbs, which follows directly from any plausible quantificational theory.

- **Importance of Probability**: Quantificational semantics has no hope of capturing fine-grained effects of probabilistic information on our judgments of the truth of statements of obligation and desire, in particular the fact that suboptimal outcomes can apparently make these statements true if it would be risky to pursue the optimal outcomes. A coarse-grained distinction between epistemically possible and epistemically impossible worlds is not sufficient.

- **Gradability**: Deontic and desire verbs and adjectives accept degree modifiers, as does *should*. They also resemble gradable adjectives in making a three-way distinction in strength, and in the fact that the intermediate-strength items participate in neg-raising. Quantificational semantics does not seem to have any way to capture these facts.

- **Deontic and Bouletic Comparatives**: The most prominent approach to deontic and bouletic comparatives and equatives within quantificational semantics, due to Kratzer (1991), predicts far too many incomparabilities and cannot make sense of even weak quantitative comparisons such as ϕ *is much better than* ψ.

- **Moral Conflicts**: Conflicts of obligation clearly exist, but quantificational theories standard in deontic logic declare them to be logically impossible, while Kratzer's theory deals with the problem by eliminating the conflict altogether.

In Chapter 6 I will propose a scalar theory of deontic and bouletic modals which incorporates information about both preference and probability, and show that it accounts for each of the puzzles noted here. Monotonicity puzzles are explained because obligation and desire are non-monotonic; probability is built into the semantics directly; and deontic modals and desire verbs are scalar, with important points of connection with the semantic typology of gradable adjectives. Scales built on expected utility are concatenative interval scales, and thus also make sense of deontic and bouletic comparatives, including quantitative comparisons. The theory also predicts the existence of genuine moral conflicts in a certain restricted class of cases.

CHAPTER 6

Scalar Semantics for Deontic Modals and Desire Verbs

6.1 Introduction

Already in the Port-Royal Logic of 1662, Antoine Arnauld & Pierre Nicole pointed out the basic problem that befalls standard approaches to the semantics of deontic and bouletic modals, which only pay attention to maximal values.

> Many people fall into an illusion which is more deceptive the more reasonable it appears to them. They only consider the greatness of the consequences of the advantage that they wish for, or of the inconvenience that they fear, without considering in any way the probability that that advantage or inconvenience will occur or not.
>
> So, when it is a great evil which they are considering, such as the loss of their lives or their goods, they think that it is prudent not to neglect any precaution to guarantee these; and if it is some great good, such as the gain of a hundred thousand crowns, they think they are acting wisely in trying to obtain it if the attempt costs them little, however small the likelihood may be that they will succeed.[1]

Nearly all of the problems discussed in the last chapter can, in one way or another, be traced back to this feature of standard modal semantics: when you are considering the desirability of a course of action, or the degree of good or harm it may cause (moral or otherwise), it is not enough to look at the extreme values, e.g. the desirability of the state of affairs given by the *best* worlds that instantiate that state of affairs.

My counter-proposal is rather close to the positive claim that Arnauld & Nicole make as well: in addition to considering the desirability of the best possible states, is necessary to consider the likelihood that these states will actually come about, and likewise for the non-optimal states.

> To judge what one must do in order to obtain a good or avoid an evil, one must consider not only the good and the evil in itself, but also the probability that it will occur or not, and to view geometrically the proportion that all these things have together. (*Ibid.*)[2]

I will argue for one way of fulfilling Arnauld & Nicole's desideratum for a system for reasoning about good, evil, and the desirability of various courses of action and states of affairs: essentially, you can determine how good or desirable ϕ is by finding out how good or desirable each world

1 Antoine Arnauld & Pierre Nicole, *La Logique ou l'Art de Penser*, 1662, Fourth Part, Chapter XVI; this quote and the next are my translations from pp.331-2 of the 1992 Gallimard edition, edited by Charles Jourdain.

2 The latter quote is the epigram of chapter 1 of Jeffrey (1965b). Jeffrey seems to have been the first to notice that the basic ideas of modern decision theory were anticipated in the Port-Royal Logic. As far as I know no one has noticed before that another passage in the same chapter describes the corresponding problem with the main body of deontic logic, as well as Kratzer's semantics.

$w \in \phi$ is, and then combining these values in proportion to how likely w is to be the actual world on the assumption that ϕ holds.

As I will develop it, this approach turns out to be equivalent to a well-understood construct known under the name "expected utility", but which I will call **Probability-Weighted Preference** to emphasize its continuity with more standard preference-based deontic and bouletic logic (and its conceptual independence from rational choice theory). As I will show, we can adopt a quite standard theory of desire and obligation based on preference orders, and — with little modification — make use of the same probabilistic information which we need to account for epistemic modals to construct scales for deontic and bouletic modals that allow us to avoid the paradoxes of quantificational deontic and bouletic semantics. Furthermore, these scales fit neatly into the typology of scales that was developed in chapter 2: they are interval scales, and are intermediate with respect to concatenation.

After introducing the technical foundations, I will proceed to the main result: each of the problems described in the last chapter has a straightforward solution in the present theory, with no revisions to other standard assumptions of formal semantics. The key features which make the resolution possible are that obligation and desire, as developed here, are scalar, non-monotonic, and sensitive to probabilistic information.

6.2 Scales and Two Kinds of Preference

6.2.1 Ordinal and Interval Scales

As we already saw in chapter 5, it is frequent in formal semantics to think of desire and obligation in terms of binary orders, often referred to as **preference orders**. As these are usually formulated, they are **ordinal scales**: structures $\langle W, \succcurlyeq \rangle$ where W is some set of possible worlds and $\succcurlyeq$ is a binary order over W. This representation has the virtue of simplicity, but it does not carry much information: an ordinal scale can tell us whether one world is better than another, but it is silent about *how much* better.

However, as we saw in the final sections of chapter 5, it is quite natural to talk about degree of difference, rather than simple comparison, with both deontic and bouletic modals.

(6.1) a. I want to go to Rome <u>much</u> more than I want to go to Paris.
b. I need to sleep <u>far</u> more than I need to eat.
c. It is <u>much</u> better to give your money to charity than it is to gamble it on sports.

Since *want, need*, and *better* operate over propositions rather than worlds, these examples are not entirely conclusive as to the nature of any relevant preference order on worlds. However, they at least make a good first case that the representation of preference should be richer than an ordinal scale, since a preference order over propositions derived from an ordinal scale of this type would almost certainly render these examples uninterpretable.

The next stronger scale type that is standard in the Representational Theory of Measurement is an **interval scale**. If these sentences are built on an interval-order preference relation, then we may be able to make sense of quantitative comparisons like those in (6.1). Recall as well that interval scales are needed for other areas of natural language semantics: for instance, in chapter 2 I argued that interval scales are the relevant scale type for properties such as temperature and danger. These scales support quantitative comparisons similar to those in (6.1) but do not allow as many quantitative comparisons as ratio-scale expressions like *tall* and *expensive*.

Formally an interval scale is a structure $\langle X, Y, \succcurlyeq_P \rangle$, where X is the domain of property P, Y is a set of pairs of objects in X, and $\succcurlyeq_P$ is a weak order on Y which satisfies several axioms (see ch. 3, §2.1.2.3). We can think of $\succcurlyeq_P$ as comparing the relative size of intervals on P, as in:

(6.2) $(a,b) \succcurlyeq_P (c,d)$ iff a exceeds b with respect to property P by more than c exceeds d.

So, for example, if $\succcurlyeq_{time}$ is the interval-valued relation underlying clock time, then $(a,b) \succcurlyeq_{time} (c,d)$ means "the interval $[a,b]$ — i.e., the length of time between point a and point b — is at least as great as the interval $[c,d]$".

An admissible measure function μ_P on an interval scale $\mathcal{S}_P$ satisfies the condition:

(6.3) $(a,b) \succcurlyeq_P (c,d)$ if and only if $\mu_P(a) - \mu_P(b) \geq \mu_P(c) - \mu_P(d)$.

This is clear again in the case of clock time: there may be reasonable ways of measuring time that differ from ours in various ways, but they should preserve the truth or falsity of statements about the relative length of intervals.

Another way to think of interval scales is that it is they have the minimal structure which allows us to measure the size of intervals along a scale without having a fixed minimum point. This intuition corresponds to the formal fact that all admissible measure functions on an interval scale $\mathcal{S}_P$ are related to all others by some positive linear function (Krantz et al. 1971).

(6.4) a. If $\mathcal{S}_P$ is an interval scale and μ_P is an admissible measure function on $\mathcal{S}_P$, then, for all $\alpha \in \mathbb{R}^+$ and all $\beta \in \mathbb{R}$, $f(\mu(x)) = \alpha \times \mu_P(x) + \beta$ is also a $\mathcal{S}_P$-admissible measure function.

b. If $\mathcal{S}_P$ is an interval scale and μ_P, μ'_P are both $\mathcal{S}_P$-admissible measure functions, then, for some $\alpha \in \mathbb{R}^+$ and some $\beta \in \mathbb{R}$, $\mu'_P(x) = \alpha \times \mu_P(x) + \beta$.

For the purposes of a degree semantics, we can characterize an interval scale equivalently by using the qualitative structure $\mathcal{S}_P = \langle X, Y, \succcurlyeq_P \rangle$ or by considering only the statements which have constant truth-value across all $\mathcal{S}_P$-admissible μ_P. Furthermore, Krantz et al. (1971) prove that, if we have a class of measure functions which contains all and only μ_P satisfying these conditions, then we can construct an equivalent qualitative representation which is an interval scale $\mathcal{S}_P$.

6.2.2 Preference over Worlds

Our first step toward a scalar semantics of obligation and desire capable of resolving the problems noted in chapter 5 is simply this: instead of treating the preference orders over worlds relevant for desire and obligation as ordinal scales, treat them using the next weakest scale type, interval scales. Again I will use the placeholder $\mathcal{D}$ for "deontic or bouletic modal" (or just "$\mathcal{D}$-modal"), and superscript a W to emphasize that this is a scale whose domain is a set of *worlds* rather than propositions. We have schematically:

(6.5) A deontic/bouletic preference order $\mathcal{S}^W_{\mathcal{D}}$ is a structure $\langle W, Y, \succcurlyeq_{\mathcal{D}} \rangle$, where

a. $Y \subseteq W \times W$;

b. $\succcurlyeq_{\mathcal{D}}$ is a weak order on Y;

c. $\mathcal{S}^W_{\mathcal{D}}$ obeys the usual interval scale axioms.

The relation $(w_1, w_2) \succcurlyeq_{\mathcal{D}} (w_3, w_4)$ reads "w_1 is morally better/more desirable/etc. than w_2 by more than w_3 is than w_4". As usual we can extract a weak order over individual worlds corresponding to the simpler "w_1 is morally better/more desirable/etc. than w_2" from $\succcurlyeq_{\mathcal{D}}$:

(6.6) $w_1 \succcurlyeq^W_{\mathcal{D}} w_2$ if and only if, for some $w_3 \in W$, $(w_1, w_3) \succcurlyeq_{\mathcal{D}} (w_2, w_3)$.

The superscripted binary relation $\succcurlyeq^W_{\mathcal{D}}$ is hereby reserved for the binary order on worlds which is constructed from the more basic relation on pairs of worlds using (6.6).

6.2.3 Preference over Propositions

Interval orders over worlds provide an important starting point, but they are not sufficient in themselves: goodness, obligation, and desirability are not really something that we predicate of worlds, but of events, states, actions, and propositions. I will assume here (along with many, though not all, in formal semantics, deontic logic, and decision theory) that deontic and bouletic modals can be treated uniformly as taking propositional arguments.[3] In order to relate the preference order just discussed to the semantics of modals, then, we need some way of relating an ordering over propositions to a preference order over the worlds which compose the proposition.

One way of doing this is closely related to what van Fraassen (1973) called the "Axiological Thesis", probably better known under the name POSSIBILISM: ϕ is better than ψ if and only if there is a (relevant) ϕ-world which is as good as or better than all (relevant) ψ-worlds. This is also very close to Kratzer's theory, and is what her semantics would give us if her ordering $\succcurlyeq_{\mathbf{g}(w)}$ were connected. Formally, a possibilist scale has the form:

(6.7) $\succcurlyeq^P_{\mathcal{D}} =_{df} \{(\phi, \psi) \mid \exists w \in \phi \forall w' \in \psi : w \succcurlyeq^W_{\mathcal{D}} w'\}$

Superscripted P is to remind us that this is an ordering on propositions, not worlds.

Even though (6.7) is an interval order capable of making the sentences in (6.1) interpretable — and so avoids a few of the problems from the last chapter — it encounters essentially all of the other problems for quantificational semantics described there. However, it is useful to consider Possibilism here because it gives us a simple picture of one way of constructing a preference order over propositions from a preference order over worlds, which will serve as the jumping-off point for my proposal.

Essentially, Possibilism tells us to line up all the worlds in ϕ according to their position in $\succcurlyeq^W_{\mathcal{D}}$, and associate ϕ with the maximum of this ordering. As a result, if $\succcurlyeq^W_{\mathcal{D}}$ is an interval order over worlds, the induced order over propositions $\succcurlyeq^P_{\mathcal{D}}$ will be an interval order as well. Each vertical line in the following graphic represents a proposition χ, with the maximum vertical extent marking out the position of the best world(s) in χ and the minimum extent marking out the position of the worst world(s) in χ.

3 That is, I'm assuming that "It is good for x to do A", "x wants to do A","x ought to do A", etc. can be treated as meaning that it is good/ought to be the case that x does A, x wants it to be the case that x does A, etc. This equation goes back at least to Chisholm (1964), and is pretty standard in formal semantics, although controversial in philosophy; see Ross (2010) for a recent objection (which, however, relies on assumptions about the logic of *ought* that I will reject here).

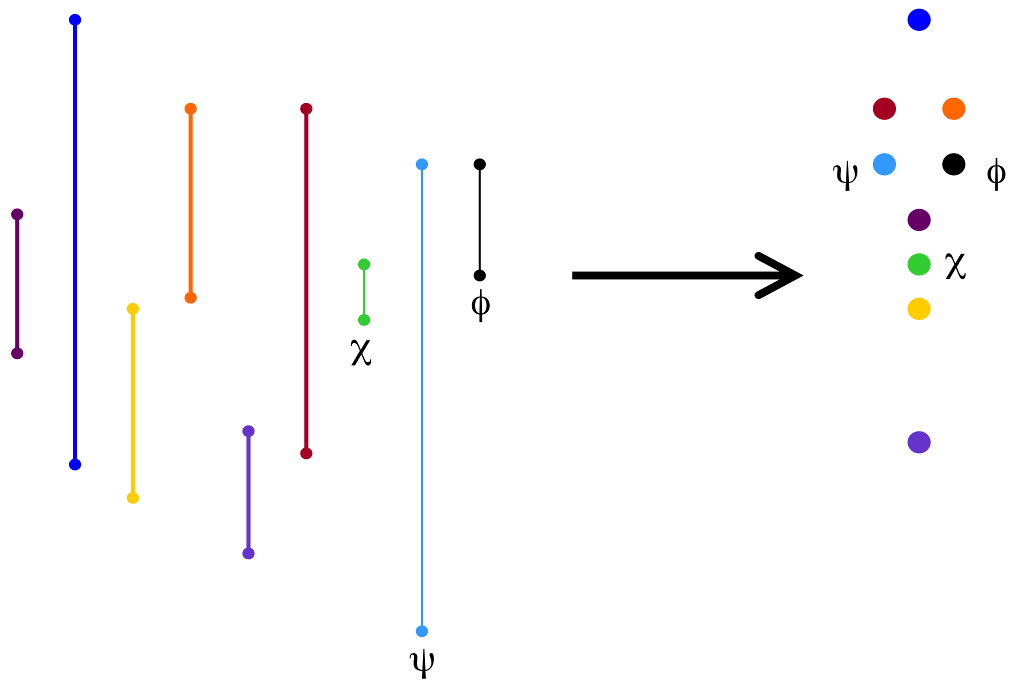


Fig. 6.1. Possibilist construction of an ordering on propositions $\succcurlyeq^P_{\mathcal{D}}$ from $\succcurlyeq^W_{\mathcal{D}}$.

One important feature of this construction is that it does not matter at all how many worlds are in each of these propositions, or how likely they are to be instantiated: all that matters is the relative order of the highest-ranked world(s) in each. Another striking fact is that the order over propositions $\succcurlyeq^P_{\mathcal{D}}$ does not care if there are very bad worlds in a proposition, whether they are many or few. For instance, take the two propositions on the right edge of the left graphic; call the one on the right edge ϕ and the one to its left ψ. According to Possibilism, ϕ and ψ are equally ranked in $\succcurlyeq^P_{\mathcal{D}}$, even though all of the ϕ-worlds are pretty good, and all of the worst worlds in the model are ψ-worlds. Similarly, Possibilism entails that ψ is much better/more desirable than the proposition to its left (χ), but intuitively this is not at all clear: χ seems to be quite indifferent, while, if ψ comes to pass, there is a substantial risk of ending up in one of the worst possible worlds.

This feature of Possibilism (and Kratzer's theory) is essentially a formalization of the approach to reasoning about desirability and obligation that Arnauld & Nicole criticize in the opening quote of this chapter. As they point out, it **should** matter when comparing ϕ and ψ whether there are some very bad worlds in one of these propositions, and it should likewise matter whether the highest-ranked worlds in ϕ are very likely or very unlikely to come about. A great deal of relevant information is being lost in the Possibilist method of constructing an order over propositions from an order over worlds. Jackson (1985: 179) puts the point clearly:

> There is a certain arbitrariness in the standard semantics. It seeks to capture what ought to be the case by identifying it with what is the case in the best worlds. But what is the case in the best worlds will also be the case in many far from best worlds, indeed some things that are the case in the best worlds will also be the case in the

worst worlds. Suppose *A* is the case in both the best and the worst worlds. *A* will count as something that ought to be the case on the standard semantics. But that is to treat its appearance in the best as special, for it is just as true that it appears in the worst. That seems arbitrary. A more even-handed approach would regard *A*'s being the case in the best as at least possibly cancelled out by its appearance in the worst. Or suppose *B* is the case in the best worlds and also the case in some very ordinary worlds; while *C* is not the case in the best, but is the case in very many very good worlds, and moreover is never the case in any worlds less than very good. Does *C*'s good, even record outweigh *B*'s decidedly patchy one? On the standard semantics it is no contest, *B*'s appearance in (all) the best worlds settles the matter once and for all — *B* ought to be, *C* ought not to be. Again it seems a more even-handed and less arbitrary approach is called for, on which the matter is open given what has been said so far, and settlable given more information.

In the next section I will describe an alternative method of constructing a binary order over propositions, which we might call "Probabilism" in contrast to "Possibilism" (as Goble 1996 suggests). This method makes better use of information about the distribution of worlds in $\succcurlyeq^W_{\mathcal{D}}$, and also takes into account information about the probability that various worlds will be actual if the proposition in question is instantiated. The crucial difference is that, instead of looking just at the MAXIMUM position in the deontic ordering of the worlds in a proposition, we look at the EXPECTED POSITION. Essentially, this is a point which tells us where on the scale the actual world is on average most likely to fall if the proposition is instantiated, given some body of probabilistic information. This method returns an interval order over propositions as well, and — as I will show in later sections — forms the basis of a scalar semantics for desires, obligations, needs, and requirements which avoids the paradoxical features of quantificational semantics noted in chapter 5.

6.3 Expectation and Weighted Preference

A different way to construct an interval order over propositions from an interval order over worlds was developed in its modern form by von Neumann & Morgenstern (1944) and extended by many since, notably Savage (1954); Jeffrey (1965b). This approach is widely known as **expected utility** and has been highly influential in economics, decision theory, Bayesian statistics, psychology, and elsewhere, but has had very little impact in deontic logic and linguistic semantics. A connection was suggested in Jeffrey (1965b,a) and then largely ignored in the main body of deontic logic, with the exception of some work in consequentialist ethics (e.g. Harsanyi 1978; Broome 1995, 1999; Jackson 1985, 1991) and a handful in logic and formal semantics (Goble 1996; van Rooij 1999; Levinson 2003).

As I will present it — following Goble (1996) in a number of respects — this construct is an attractive alternative to Possibilism (including Kratzer's theory) as a method for constructing interval orders over propositions from interval orders over worlds, which can be shown to have numerous other advantages for the analysis of deontic and bouletic modals. I want to emphasize that, despite the associations that the name "expected utility" conjures up, the abstract theory that I am giving has no special connection to economic behavior, subjective utility, subjective probability, or even decision-making. Questions of how the relevant ordering over worlds and propositions is determined in context, what the relevant probability distribution is, and whether and how agents actually use this information to make choices can safely be left to the side here. All we are interested in here is the right method of constructing scales for deontic and bouletic modals, and in particular what constraints the semantics of these expressions place on the relationship between preference over worlds, preference over propositions, and probability distributions.

In order to avoid these distracting associations, however, I will generally avoid using the term "expected utility", referring to the method of constructing scales as **probability-weighted preference** or just **expectation** in order to emphasize the generality of the concept and its applicability to non-subjectivistic concepts, in particular moral obligation.

6.3.1 Weight and Expectation

Suppose that you are going to combine a number of different containers of water $x_1, x_2, \ldots, x_n$ into a single container, and you want to know what the average temperature in degrees Celsius of the result will be (ignoring the possibility of heat loss). You might be able to get a rough estimate by calculating the average temperature of the containers $[\sum_{i=1}^{n} temp(x_1)]/n$, but this will usually give the wrong result unless all of the containers have the same volume. If one of the containers is much larger than the others, the temperature of this container will dominate the result in reality, but not in the naïve average computed without taking volume into account.

A better way to compute this quantity is, of course, is to take a **weighted average**, where the weights are provided by the amount of water in each container, and the result is normalized by dividing by the total amount of water in all of the containers (to ensure that the answer is comparable with the temperature measurements that we started out with). If there are n containers, then the result that we want is just

$$\frac{\sum_{i=1}^{n} [temp(x_i) \times vol(x_i)]}{\sum_{k=1}^{n} vol(x_k)}$$

This is of course equivalent to calculating the proportion of the total amount of water that each x_i will provide, and using this quantity as the **weight** of w_i:

$$\sum_{i=1}^{n} \left[temp(x_i) \times \frac{vol(x_i)}{\sum_{k=1}^{n} vol(x_k))} \right]$$

where the term on the right gives the weight accorded to $temp(x_i)$, here the volume of x_i as a proportion of the total volume of water in question. The total of the weights must sum to 1, or be made to by adding a normalizing term (as we did in dividing by $\sum_{i=1}^{n} vol(x_i)$ in the first equation).

More generally, a weighted average of n values can be represented by a $2 \times n$ matrix, where the top row is a list of values $val(x_i)$ and the bottom row is a list of weights $weight(x_i)$.

$$\begin{pmatrix} val(x_i) & val(x_2) & val(x_3) & val(x_4) & \dots & val(x_n) \\ weight(x_i) & weight(x_2) & weight(x_3) & weight(x_4) & \dots & weight(x_n) \end{pmatrix}$$

The weights can be anything you like, but if they do not total 1 we must divide by the sum of the weights in order to get a quantity that is meaningful relative to the values of individual variables.

Weighted averages are frequently important when we have occasion to look for the expected value $\mathbb{E}$ of some variable. In these cases the weighting function is usually (though not necessarily) given by a probability measure. For a simple example, suppose you are playing a game which involves rolling a die and betting on which numbers will come up. Someone offers you a bet where the die is rolled, and you win \$1 if the number of dots is 3 or less, but pay \$1 otherwise. The number of dots that comes up on a roll of the die is a variable $\mathcal{X}$. To find out whether this is a bet worth taking, you can take the average of the number of dots on each side x_i, weighted by the probability that side x_i will come up.

Supposing that it is a fair die, the probability of each side coming up is $\frac{1}{6}$, and no normalization is needed because the probabilities sum to 1. This gives us:

$$\mathbb{E}(\mathcal{X}) = \left(1 \times \frac{1}{6}\right) + \left(2 \times \frac{1}{6}\right) + \left(3 \times \frac{1}{6}\right) + \left(4 \times \frac{1}{6}\right) + \left(5 \times \frac{1}{6}\right) + \left(6 \times \frac{1}{6}\right) = 3.5$$

or more generally:

$$\mathbb{E}(\mathcal{X}) = \sum_{i=1}^{n} [val(x_i) \times prob(\mathcal{X} = x_i)]$$

where $[x_1, ..., x_n]$ is a vector of the n possible outcomes of the roll, *val* is a function from each outcome x_i to the number of dots that will appear if x_i comes up, and *prob* is the weight function, here a probability measure telling us how likely it is that we will see each x_i when we roll.

This calculation allows you to infer that the bet you have been offered is not a good one: since the expected value of the roll of the die is 3.5, you are more likely to lose money than to win on an even bet that the value of the roll of the die will be ≤ 3. Note that the expected value of a variable does not need to be a value that the variable can actually take on — no side of the die has 3.5 dots on it.

The assumption that the die is fair is important: if the die were weighted, the expected value would be different. For example, if the die were heavily weighted, the probability function might be

$$prob(\mathcal{X} = x_i) = \begin{Bmatrix} x_1 \longrightarrow & .7 \\ x_2 \longrightarrow & .1 \\ x_3 \longrightarrow & .1 \\ x_4 \longrightarrow & .05 \\ x_5 \longrightarrow & .05 \\ x_6 \longrightarrow & 1 \end{Bmatrix}$$

and the expected value would then be

$$\mathbb{E}(\mathcal{X}) = (.7 \times 1) + (.1 \times 2) + (.1 \times 3) + (.1 \times 4) + (.05 \times 5) + (.05 \times 6) = 2.15$$

In this case, the bet you are being offered is a good one, since the expected value is less than 3 (so there is a greater than 50% chance that you will win).

6.3.2 Scale Type and Expectation

We want to see if we can apply the concept of expectation to the transition from a preference order $\succsim_{\mathcal{D}}^{W}$ over worlds to an ordering $\succsim_{\mathcal{D}}^{P}$ over propositions. Now, if preference is given by an ordinal scale, this is not possible. The notion of a weighted average (or any average) is not well-defined for ordinal scales, and so we cannot apply the notion of an expected value of a set because it is not interpretable in the RTM sense.

To see this, recall that, if $\mathcal{S}_P$ is ordinal, then any order-preserving (monotone increasing) transformation of an admissible μ_P is also admissible. It follows that admissible measure functions on an ordinal scale can disagree on the expectation of a variable. For example, let $\mathcal{S}_P = \langle X, \succsim_P \rangle$, where $X = \{a,b,c,d,e\}$ and $a \succ_P b \succ_P c \succ_P d \succ_P e$. Two admissible measure functions are:

$$\mu_1 = \begin{Bmatrix} a \longrightarrow & 5 \\ b \longrightarrow & 4 \\ c \longrightarrow & 3 \\ d \longrightarrow & 2 \\ e \longrightarrow & 1 \end{Bmatrix} \qquad \mu_2 = \begin{Bmatrix} a \longrightarrow & 10 \\ b \longrightarrow & 6 \\ c \longrightarrow & 3 \\ d \longrightarrow & 1 \\ e \longrightarrow & 0 \end{Bmatrix}$$

The expected values of X relative to μ_1 and μ_2 will not in general be related in any systematic way. For instance, if each of *a*-*e* is equiprobable, then the expectation $\mathbb{E}_1(X)$ (determined by reference to μ_1) is 3, but the expectation $\mathbb{E}_2(X)$ (determined by reference to μ_2) is 4. This is enough to make the point, since, if expectation were preserved under the order-preserving transformation $f(\mu_1(x)) = \mu_2(x)$, then the value 3 in μ_1 would have to be mapped to its counterpart in μ_2, namely 3. Since the claim that $\mathbb{E}(X) = \mu(c)$ is true relative to μ_1 but false relative to μ_2, we have to declare the statement uninterpretable in the RTM sense — and in general, where ordinal scales are concerned.

I already claimed that it is better to associate preference with an interval scale, however. To see that expectation is a stable quantity with interval scales, suppose that μ' is an admissible measure function on an ordinal scale with domain X. (This is actually clear from the first example of a weighted average above, since temperature is an interval scale. It will still be useful to see why it works in general for interval scales, though, since it will provide the proof that our construction of measure functions on propositions yields an interval scale in the next subsection.) Then all and only transformations of the form $\mu''(x) = \alpha \times \mu'(x) + \beta$ are also admissible measure functions, for $\alpha > 0$. Letting *weight* be a normalized weight function on some subset Z of X, the expectation $\mathbb{E}'$ of Z relative to μ' is:

$$\mathbb{E}'(Z) = \sum_{i=1}^{n} \left[\mu'(x_i) \times weight(x_i)\right]$$

Transforming each $\mu'(x)$ into $\mu''(x) = \alpha \times \mu'(x) + \beta$, where $\alpha > 0$, we can calculate the expectation $\mathbb{E}''$ of Z relative to μ'' as

$$\mathbb{E}''(Z) = \sum_{i=1}^{n} \left[\mu''(x_i) \times weight(x_i)\right] = \sum_{i=1}^{n} \left[(\alpha \times \mu'(x_i) + \beta) \times weight(x_i)\right]$$

$$= \alpha \times \sum_{i=1}^{n} \left[\mu'(x_i) \times weight(x_i)\right] + \left(\beta \times \sum_{i=1}^{n} weight(x_i)\right)$$

Since the weight function is normalized, $\sum_{i=1}^{n} weight(x_i)$ is equal to 1; so the equation simplifies to

$$\mathbb{E}''(Z) = \alpha \times \sum_{i=1}^{n} [\mu'(x_i) \times weight(x_i)] + \beta$$

But the latter formula is equivalent to $\mathbb{E}''(Z) = \alpha \times \mathbb{E}'(Z) + \beta$, which is just the same transformation that we applied to each value of μ' in order to create μ''. Since this was an arbitrary admissible transformation of an interval scale, we can conclude that, unlike ordinal scales, expected value is stable across all admissible μ. As a result statements making reference to expectation are interpretable on interval scales (and stronger scales, including ratio scales).

6.3.3 Obligation, Desire, and Probability-Weighted Preference

The proposal, then, is this: the basic semantics of obligation, desire, and related concepts (requirements, needs, etc.) is given by an interval order $\succcurlyeq_{\mathcal{D}}^{P}$ of propositions which is constructed from two components. The first is an interval order $\succcurlyeq_{\mathcal{D}}^{W}$ on worlds; the second is a probability measure *prob*, which, as we saw in chapter 3-4, is needed independently to account for the semantics of epistemic modals. *prob* provides the weight function which is used to calculate the expectation $\mathbb{E}(\phi)$ of a proposition ϕ, and this calculation relates the degree of obligation or desirability of a proposition to the degree of obligation or desirability of its component worlds: the scales of deontic and bouletic modals are given by **probability-weighted preference**.

The degree of obligation/desire/etc. attached to a proposition by $\succcurlyeq_{\mathcal{D}}^{P}$ is the weighted average of the degrees of obligation/desire/etc. attached to the individual worlds in ϕ by $\succcurlyeq_{\mathcal{D}}^{W}$, where the weights are given by the probabilities of the individual worlds. As usual, the function must be normalized by dividing by the total probability of the worlds in ϕ. Let $\mu_{\mathcal{D}}^{W}$ be an admissible measure function on an interval scale $\mathcal{S}_{\mathcal{D}}^{W} = \langle W, Y, \succcurlyeq_{\mathcal{D}} \rangle$, where W is a set of worlds. $\mu_{\mathcal{D}}^{P}$, the corresponding measure function on propositions, is given by the equation:[4]

$$\mu_{\mathcal{D}}^{P}(\phi) = \frac{\sum_{w \in \phi} \left[\mu_{\mathcal{D}}^{W}(w) \times prob(w)\right]}{\sum_{w' \in \phi} prob(w')} = \sum_{w \in \phi} \left[\mu_{\mathcal{D}}^{W}(w) \times \frac{prob(w)}{\sum_{w' \in \phi} prob(w')}\right]$$

Noticing that the effect in this equation of normalizing by the total probability of the worlds in ϕ is the same as taking the weight function to be a conditional probability measure, we can express this equation a bit more compactly as (6.8), which is my official proposal for the structure of the scales underlying deontic and bouletic modality.[5]

(6.8) $\mu_{\mathcal{D}}^{P}(\phi) = \sum_{w \in \phi} \left[\mu_{\mathcal{D}}^{W}(w) \times prob(w|\phi)\right]$

4 I have to confess to a slight abuse of notation: officially *prob* takes propositional arguments, and so $prob(w_i)$ should be $prob(\{w_i\})$, the probability that the actual world is in $\{w_i\}$. I don't think that this practice will lead to any confusion, though.

5 I am simplifying by assuming that W is finite. Allowing for infinite W would not affect the proposal in any significant way, but would require getting a bit more mathematically involved and would distract from the main point here.

(The conditional probability measure $prob(\psi|\chi)$ is equal to $(prob(\psi \wedge \chi)/prob(\chi))$, the ratio of the probability that ψ and χ both hold to the probability that χ alone holds. Roughly this is the probability that we would assign to ψ if we were to assume, or find out, that χ is definitely true.)

Since expectation is preserved across interval-scale measure functions, as we saw in the last subsection, it follows that $\mu_{\mathcal{D}}^{P}$ is associated with an interval scale. This is because $\mu_{\mathcal{D}}^{P}$ is constructed by taking a weighted average of the values of the worlds in a set, relative to some measure function on worlds $\mu_{\mathcal{D}}^{W}$ which is admissible for some interval scale $\mathcal{S}_{\mathcal{D}}^{W}$. Since it is an interval scale, any positive linear transformation $f(\mu_{\mathcal{D}}^{W}(w)) = \alpha \times \mu_{\mathcal{D}}^{W}(w) + \beta$, with $\alpha > 0$, is also admissible for $\mathcal{S}_{\mathcal{D}}^{W}$. It follows from this and the proof in the previous subsection that the expectation of any set of worlds ϕ relative to an admissible transformation of $\mu_{\mathcal{D}}^{W}$ will be related to $\mu_{\mathcal{D}}^{P}(\phi)$ by the same α and β. Since α and β were arbitrary (with $\alpha > 0$), we can conclude that all transformations $f(\mu_{\mathcal{D}}^{P}(\phi)) = \alpha \times \mu_{\mathcal{D}}^{P}(\phi) + \beta$ are admissible measure functions on propositions corresponding to some $\mathcal{S}_{\mathcal{D}}^{W}$-admissible measure function on worlds.

Now, as discussed in chapter 2, the fact that all and only transformations of this type are admissible for $\mu_{\mathcal{D}}^{P}$ provides necessary and sufficient conditions for $\mu_{\mathcal{D}}^{P}$ being associated with an interval scale. As a result, we can be sure that any $\mu_{\mathcal{D}}^{P}$ as defined in (6.8) uniquely characterizes an interval scale $\mathcal{S}_{\mathcal{D}}^{P} = \langle \Phi, \mathbf{Y}, \succsim_{\mathcal{D}}^{P} \rangle$, where Φ is a set of propositions; $\mathbf{Y} \subseteq \Phi \times \Phi$; and $\succsim_{\mathcal{D}}^{P}$ is required to obey the interval scale axioms. The interval scale of preference on worlds and the scale of preference on propositions are related in a straightforward way: because the scale $\mathcal{S}_{\mathcal{D}}^{W} = \langle W, Y, \succsim_{\mathcal{D}} \rangle$ has as its domain a set of worlds W, the propositional scale $\mathcal{S}_{\mathcal{D}}^{P}$ defined from it has as its domain the power set of W.[6]

6.4 Semantic and Logical Properties of Expectation: Some Puzzles Resolved

The enduring influence of probability-weighted preference in many fields under many names — "desirability", "expected utility", "expected loss", and various formal and informal guises in moral philosophy, to name a few — is a testament to the usefulness of the concept in various practical applications. I suspect that this is because probability weighting provides an optimal way of combining three kinds of information: information about preferences among states (worlds); information about the distribution and density of the states making up a particular proposition along the preference order; and information about the likelihood that a given state has of being actual, and thus how much we should have hope — or fear — that this state will come to be.

Probability-weighting does not only pay attention to extreme values, as possibilism and Kratzer's semantics do, but it does not ignore these values either: if some of the worlds in a proposition

6 A brief note on the relationship between the approach adopted here and traditional axiomatizations of expected utility. von Neumann & Morgenstern (1944) and much following work derive preference orders on "gambles" which are equivalent to the preference orders on propositions that are characterized by (6.8). Essentially they show that, on various plausible assumptions about the structure of a rational agent's preferences (as well as some Archimedean and solvability assumptions), an agent's preferences among gambles over states — equivalent to propositions along with probability distributions — will obey (6.8) and will be an interval order. I've avoided operationalizing the concept of probability-weighted preference like this, however, because we can construct a semantics along these lines without it, and because I don't want to tie it to the idea of any individual's preferences or the theory of choice more generally; expected utility is a concept with broader applicability than the traditional construction might suggest.

are very good, this will skew the desirability of the proposition as a whole positively, to a degree proportional to the goodness of the worlds in question. However, if these worlds are very remote possibilities, they may will not receive much weight even if they are very good.

These are some of the essential features which allow us to avoid several paradoxes of quantificational semantics noted in the previous chapter, as I will show. However, there is more: because of the way that weighted preference is constructed, its logical properties differ considerably from quantificational semantics in various ways. This section will show that an expectation-based scalar semantics for deontic and bouletic modals is

- **Non-monotonic**, and makes specific and intuitively correct predictions about the puzzles which led us to question the monotonicity of obligation and desire in chapter 5;
- **Sensitive to information** in a way which — I will show — fits the data regarding information-related puzzles, including the Miner's Paradox;
- **Scalar**, and compatible with a good compositional semantics of deontic and bouletic comparisons, as well as a principled approach to the difference between weak and strong "necessity" modals.

6.4.1 Monotonicity

The semantics that I will develop for deontic modals and desire verbs essentially treats them, like gradable adjectives, as establishing a threshold value and returning the value True if and only if the proposition is mapped to a point on the scale which exceeds the threshold. All deontic modals and desire verbs, then, have the schematic form

(6.9) $\mathcal{D}(\phi)$ is true if and only if $\mu^{P}_{\mathcal{D}}(\phi) \geq \theta_{\mathcal{D}}$, where $\theta_{\mathcal{D}}$ is a threshold value determined by the lexical semantics of $\mathcal{D}$ in interaction with the context.

For the moment, I want to be relatively non-committal about exactly how the various lexical items constrain the value of $\theta_{\mathcal{D}}$, since many of the puzzles in Chapter 5 can be resolved by looking at the structure of deontic and bouletic scales at this level of abstraction, without worrying too much about lexical semantics just yet.

As a preview, though, I will use entailment data and various tests developed in previous chapters to argue that deontic and bouletic modals fall into three groups:

- **High scalar** $\mathcal{D}$-modals such as *require, need, must* and *have to*, which have a high threshold and resemble maximum-standard and high-degree adjectives in various ways;
- **Mid-scalar** $\mathcal{D}$-modals such as *want, ought, supposed to, should*, and *good* which resemble relative-standard gradable adjectives, *likely*, and *probable* in setting the value of $\theta_{\mathcal{D}}$ in a way that is sensitive to contextual alternatives;
- **Weak scalar** $\mathcal{D}$-modals such as *allowed, permitted*, and *may* which have a relative low threshold and resemble minimum-standard adjectives in certain respects.

I will give "high" and "low" more content in what is to come later; for now these characterizations should be sufficient.

Note that these items also display some variation as to how the underlying preference orders are determined: roughly, *want* is associated with subjective preference, *ought* and *must* often associate with moral preference, and so forth. However, for the purpose of investigating the logic of these notions and the entailments that they license, this particular parameter of variation does not matter too much: the scales underlying their semantics have the same basic structure, I claim.

If something like (6.9) is right, then it's straightforward to show that probability-weighted preference is NON-MONOTONIC — that is, neither of the inference schemata in (6.10) and (6.11) is valid.

(6.10) UPWARD MONOTONICITY

a. $\mathcal{D}(\phi)$

b. $\phi \vDash \psi$

c. $\therefore\ \mathcal{D}(\psi)$

(6.11) DOWNWARD MONOTONICITY

a. $\mathcal{D}(\phi)$

b. $\psi \vDash \phi$

c. $\therefore\ \mathcal{D}(\psi)$

It is uncontroversial that deontic and bouletic modals are not downward monotonic: nobody would endorse a semantics which validates inferences like "You ought to go home, so you ought to go home and burn your house down". Upward monotonicity, on the other hand, is a feature of most deontic logics, including Kratzer's.

We saw several reasons to be skeptical of upward monotonicity in chapter 5: for example, I argued that the arguments (6.12) and (6.13) admit of counter-examples (in this form and also with *want* replacing *ought*).

(6.12) a. $ought(\phi)$

b. $\phi \vDash (\phi \vee \psi)$

c. $\therefore\ ought(\phi \vee \psi)$

(6.13) a. $ought(\phi \wedge \psi)$

b. $(\phi \wedge \psi) \vDash \phi$

c. $\therefore\ ought(\phi)$

The failure of these inference patterns suggests that upward monotonicity is not a general property of $\mathcal{D}$-modals. The Chicken example from Jackson (1985) also suggested that the following related inference is invalid:

(6.14) a. $ought(\phi) \wedge ought(\psi)$

b. $\therefore\ ought(\phi \wedge \psi)$

Here I show that a semantics built on weighted preference correctly predicts the counter-examples.

6.4.1.1 Ross' Paradox

Ross' Paradox is the puzzle which motivated us to reject (6.12) as a valid inference; a variant of the classic example is (6.15).

(6.15) a. You must mail this letter.
b. ∴ You must mail this letter or burn it.

It is not hard to see that this inference will not be valid if *must* has a scalar semantics along the lines of (6.9). According to that proposal, $must(\phi)$ is true if and only if the expectation of ϕ exceeds some threshold θ_{must}. Let ϕ be the proposition that you mail the letter, and ψ be the proposition that you burn the letter. On the present theory the premise (6.15a) has the schematic truth-conditions

(6.16) $must\ \phi$ is true iff $\mathbb{E}(\phi) \geq \theta_{must}$.

Expanding the right side of (6.16), we have

(6.17) $must\ \phi$ is true iff $\sum_{w\in\phi}\left[\mu^W_{\mathcal{D}}(w) \times prob(w|\phi)\right] \geq \theta_{must}$.

Note that the expectation is calculated by weighting each world w by the conditional probability of *w given that* ϕ *is true* (rather than the unconditional probability of w).

On the other hand, the conclusion *You must mail this latter or burn it* has the truth-conditions in (6.18) and (6.19):

(6.18) $must\ (\phi\ or\ \psi)$ is true iff $\mathbb{E}(\phi \vee \psi) \geq \theta_{must}$.

(6.19) $must\ (\phi\ or\ \psi)$ is true iff $\sum_{w\in(\phi\vee\psi)}\left[\mu^W_{\mathcal{D}}(w) \times prob(w|\phi \vee \psi)\right] \geq \theta_{must}$.

In (6.19), unlike (6.17), we are considering the weighted average of **all** the worlds in $(\phi \vee \psi)$, and the weights are given by the conditional probability of a world *w given that* ϕ *is true* ***or*** ψ *is true*. Effectively, this means that, if there is a substantial probability that ψ may happen, then this will shift the expected value of the disjunction away from the desirable ϕ-worlds and toward the less desirable ψ-worlds.

Another way to see this is to note that, if ϕ and ψ are incompatible, we can calculate $\mathbb{E}(\phi \vee \psi)$ as the probability-weighted average of ϕ and ψ:

$$\mathbb{E}(\phi \vee \psi) = \frac{\mathbb{E}(\phi) \times prob(\phi) + \mathbb{E}(\psi) \times prob(\psi)}{prob(\phi) + prob(\psi)}$$

or, equivalently, taking the weight of each proposition to be given by its conditional probability *on the assumption that one or the other is true* (cf. Jeffrey 1965b: ch.5).

(6.20) $\mathbb{E}(\phi \vee \psi) = \mathbb{E}(\phi) \times prob(\phi|\phi \vee \psi) + \mathbb{E}(\psi) \times prob(\psi|\phi \vee \psi)$

It follows immediately from (6.20) that $\mathbb{E}(\phi \vee \psi)$ will not always be greater than or equal to $\mathbb{E}(\phi)$, and so $\mathbb{E}(\cdot)$ is not upward monotonic. Instead, as long as $prob(\phi)$ and $prob(\psi)$ are both non-zero,

- $\mathbb{E}(\phi \vee \psi)$ is greater than $\mathbb{E}(\phi)$ if and only if $\mathbb{E}(\psi)$ is greater than $\mathbb{E}(\phi)$.

- $\mathbb{E}(\phi \vee \psi)$ is less than $\mathbb{E}(\phi)$ if and only if $\mathbb{E}(\psi)$ is less than $\mathbb{E}(\phi)$.
- $\mathbb{E}(\phi \vee \psi) = \mathbb{E}(\phi)$ if and only if $\mathbb{E}(\psi) = \mathbb{E}(\phi)$.

As a result, the fact that $\mathbb{E}(\phi) \geq \theta_{must}$ (i.e., that you must mail the letter) tells us nothing at all about whether $\mathbb{E}(\phi \vee \psi) \geq \theta_{must}$ (whether you must mail or burn it). At best, this inference will hold only we add the premise $\mathbb{E}(\psi) \geq \mathbb{E}(\phi)$ — that burning the letter is at least as good as mailing it, a condition which is clearly very implausible here.

This is why probability-weighted preference is non-monotonic. It is also why Ross' Paradox does not arise for the present theory: the paradox gets its force from the fact that the expectation of ψ = *You burn this letter* is intuitively very low in any normal context, and surely less than $\mathbb{E}(\phi)$. Since the requirement of disjointness is clearly fulfilled as well, we apply (6.20) and get the result that (6.15a) does not entail (6.15b). This is the right result, it seems.

A quick note on this solution. It is true that, if we add the premise $\mathbb{E}(\psi) \geq \mathbb{E}(\phi)$, then we do have a valid inference: for example, if you must mail the letter, then you must mail the letter or cure cancer. This is not too troubling, though. In the specific semantics for *must* that I will propose below, these premises are not consistent unless ψ is extremely improbable (cf. §6.5.5.2). In the latter case, the fact that this inference is odd can be explained along standard Gricean lines: since curing cancer is extremely improbable, the inference is valid but misleading, since it implicates strongly that curing cancer is a realistic option. As a result, *must* (ϕ) entails *must* $(\phi$ *or* $\psi)$ iff ψ is very good and it is not a realistic possibility in context. This is in sharp contrast to quantificational accounts, on which the inference from *must* (ϕ) to *must* $(\phi$ *or* $\psi)$ is always valid, even when ψ is a realistic option and is very bad.

6.4.1.2 Aside on Disjunction and Concatenation

In fact, we can get a bit more specific than we have about the relationship between the expectation of a disjunction and the expectation of the individual disjuncts. It follows from (6.20) that the expectation of a disjunction will fall somewhere between the expectation of its disjuncts, inclusive:

(6.21) If $\mathbb{E}(\phi) \geq \mathbb{E}(\psi)$, then $\mathbb{E}(\phi) \geq \mathbb{E}(\phi \vee \psi) \geq \mathbb{E}(\psi)$.

This property of expectation should look familiar: it is the same as the property of intermediacy with respect to concatenation which I discussed in chapter 2 (§§2.2.1-2.2.2, especially (2.33)).

(6.22) A scale $\mathcal{S}_P$ is **intermediate** with respect to concatenation if and only if the following (equivalent) conditions hold:

a. If $x \succcurlyeq_P y$, then $x \succcurlyeq_P (x \circ y) \succcurlyeq_P y$.

b. For all admissible μ_P, if $\mu_P(x) \geq \mu_P(y)$, then $\mu_P(x) \geq \mu_P(x \circ y) \geq \mu_P(y)$.

I argued there that the properties of danger and temperature are intermediate with respect to concatenation. According to the present proposal, deontic and bouletic modals also have this property: the degree of goodness of desirability of a concatenation (disjunction) of propositions is intermediate between the individual goodnesses/desirabilities of the disjuncts.

6.4.1.3 Professor Procrastinate

A related argument shows that this proposal resolves Jackson's (1985) Professor Procrastinate puzzle. Recall that Professor Procrastinate, the world expert in some obscure subject, has been asked to review a book; if he accepts good-naturedly, he almost certainly will not write it; but if he declines, someone less qualified but also less forgetful will do it. Jackson (1985); Jackson & Pargetter (1986) judge that both of the following are true in this scenario:

(6.23) a. Procrastinate ought to accept and write the review.
b. It's not the case that Procrastinate ought to accept the review.

Similarly, I argued that there are situations in which both of the sentences in (6.24) can be true:

(6.24) a. I want Sam to come to my birthday party and stay sober.
b. I don't want Sam to come to my birthday party.

In both cases, the crucial aspect of the situation which seemed to make these pairs compatible is the fact that, if the professor accepts the review and Sam comes to the party there is a very high **probability** that we will be in one of the worst possible situations, where no review is written, and Sam is drunk and belligerent.

Each of the pairs in (6.23) and (6.24) characterizes a set of sentences of the form $\{\mathcal{D}(\phi \wedge \psi), \neg\mathcal{D}(\phi)\}$. I will show that, on the present account, a set of this form is logically consistent just in case (a) the conditional probability of $\neg\psi$ given ϕ is sufficiently high, and (b) $\phi \wedge \neg\psi$ is less desirable than $\phi \wedge \psi$ to a sufficient degree. This seems to be a good characterization of what is going on in the Professor Procrastinate scenario and its variants (modulo certain special features of the high-degree modals *must*, *require*, etc.; see §6.5.5 below).

(6.23b) and (6.24b) have the schematic truth-conditions

(6.25) $\mathcal{D}(\phi \wedge \psi)$ is true iff $\mathbb{E}(\phi \wedge \psi) \geq \theta_{\mathcal{D}}$

where $\theta_{\mathcal{D}}$ is the relevant threshold value. Likewise, (6.23a) and (6.23b) have the schematic truth-conditions

(6.26) $\neg\mathcal{D}(\phi)$ is true iff $\neg(\mathbb{E}(\phi) \geq \theta_{\mathcal{D}})$, which is true iff $\mathbb{E}(\phi) < \theta_{\mathcal{D}}$.

Making use of the equivalence between ϕ and $(\phi \wedge \psi) \vee (\phi \wedge \neg\psi)$, we can rewrite $\mathbb{E}(\phi)$ as $\mathbb{E}((\phi \wedge \psi) \vee (\phi \wedge \neg\psi))$. Since the disjuncts are incompatible, we can use the formula for calculating disjunctions of expectations in (6.20), yielding

$$\mathbb{E}(\phi) = \mathbb{E}((\phi \wedge \psi) \vee (\phi \wedge \neg\psi)) = \mathbb{E}(\phi \wedge \psi) \times prob(\phi \wedge \psi|\phi) + \mathbb{E}(\phi \wedge \neg\psi) \times prob(\phi \wedge \neg\psi|\phi)$$

Since by assumption $\mathcal{D}(\phi \wedge \psi)$ is true while $\mathcal{D}(\phi)$ is false, $\mathbb{E}(\phi \wedge \psi) > \mathbb{E}(\phi)$. Plugging into this inequality the expression for $\mathbb{E}(\phi)$ just derived and simplifying the conditional probability statements, we have:

$$\mathbb{E}(\phi \wedge \psi) > \mathbb{E}(\phi \wedge \psi) \times prob(\psi|\phi) + \mathbb{E}(\phi \wedge \neg\psi) \times prob(\neg\psi|\phi)$$

Substituting $(1 - prob(\neg\psi|\phi))$ for $prob(\psi|\phi)$ and rearranging gives us

$$\mathbb{E}(\phi \wedge \phi) \times prob(\neg\psi|\phi) > \mathbb{E}(\phi \wedge \neg\psi) \times prob(\neg\psi|\phi)$$

which is equivalent to (6.27):

(6.27) $\mathbb{E}(\phi \wedge \psi) > \mathbb{E}(\phi \wedge \neg\psi)$

In other words, on a purely logical level there is no difficulty at all in accommodating the Professor Procrastinate example and its variants. In effect, the minimal condition for a model to verify "He should accept and write" and "He should not accept" simultaneously is that it must be better for the Professor to accept and write than to accept and not write. By construction, this condition is fulfilled in the relevant examples.

It may seem at this point that the solution is overkill. Isn't it too easy to fulfill the condition in (6.27), and won't too many pairs of sentences of the form in (6.25) come out satisfiable as a result? After all, part of what makes the Professor Procrastinate examples interesting is that relatively unusual scenarios seem to be needed before we judge it reasonable for someone to endorse both $ought(\phi\ and\ \psi)$ and $not(ought(\phi))$. (6.27), on the other hand, is a condition which will be fulfilled in many scenarios.

This objection is not too troubling, though. What (6.27) gives us is just the **minimal** condition under which a Procrastinate case could ever be constructed in any model, in particular in the case in which $\mathbb{E}(\phi \wedge \psi)$ is equal to or just barely greater than θ_{ought}. However, the situations for which examples like this are compelling are generally ones in which $\mathbb{E}(\phi \wedge \psi)$ is much greater than θ_{ought}. It turns out that, if this condition holds, then $ought(\phi)$ will fail only if $\mathbb{E}(\phi \wedge \neg\psi)$ is less than θ_{ought} by a correspondingly large amount, or if $\neg\psi$ is more likely than ψ if ϕ holds. This means that, in cases in which $ought(\phi \wedge \psi)$ is clearly true, then a quite special set of circumstances must hold in order for $ought(\phi)$ to fail.

It's worth pausing to be quite precise about this feature of the semantics, since the detailed rationale will pop up again several times in discussing other puzzles. Suppose that $\mathbb{E}(\phi \wedge \psi)$ exceeds the *ought* threshold by a difference of η: $\mathbb{E}(\phi \wedge \psi) = \theta_{ought} + \eta$ for some $\eta > 0$. It can be shown that that (6.25) and (6.26) are consistent if and only if[7]

$$\mathbb{E}(\phi \wedge \neg\psi) \leq \theta_{ought} - \eta \times \left[\frac{prob(\psi|\phi)}{prob(\neg\psi|\phi)} \right]$$

In words, if the goodness/desirability of $\phi \wedge \psi$ exceeds θ_{ought} by some amount η, the only way that $ought(\phi)$ can fail to be true is if the goodness/desirability of $\phi \wedge \neg\psi$ is less than θ_{ought} to a degree proportional to (a) the difference in expectations η, and (b) the odds of ψ conditional on ϕ.[8]

Essentially, what this means is that there are two cases we need to consider. If the odds that the Professor will complete the review if he accepts it are very low, as in Jackson & Pargetter's (1986) story, then the difference between θ_{ought} and $\mathbb{E}(\phi \wedge \neg\psi)$ does not need to be too great. On the other hand, if the odds that he will complete the review if he accepts it are highish — say, there is a 75% chance that he will do so, so that the odds are three to one — then $ought(\phi)$ will fail to be

7 The derivation is straightforward, but would take up more space than it deserves here; it just involves expanding the formulas and doing some algebraic manipulation.

8 The **odds** of an event A are the ratio of the probability of A to the probability of its negation: $odds(A) = \frac{prob(A)}{prob(\neg A)} = \frac{prob(A)}{1-prob(A)}$. The **conditional** odds relative to some event B are the same, with the unconditional probability measure $prob(\cdot)$ replaced by the conditional probability measure $prob(\cdot|B)$.

true only if $\mathbb{E}(\phi \wedge \neg\psi)$ is far worse than θ_{ought}, with a difference of at least 3η. This corresponds to the intuition that, if ϕ is unimportant in itself but it is of world-shattering importance that ψ occur (and disastrous if it does not), then even a small chance of $\neg\psi$ given ϕ may be enough for us to conclude that $ought(\phi)$ is false.

In the story as Jackson & Pargetter (1986) tell it, we have the most favorable case we could find: η is large and the conditional odds of ψ given ϕ are low. These two conditions combine to make it very easy for $ought(\phi \wedge \psi)$ and $not(ought(\phi))$ to be true simultaneously, rendering an otherwise remote scenario quite intuitive: he clearly ought to accept the review and write it, but equally clearly he ought not to accept it since he probably won't write it if he does.

6.4.1.4 Chicken

In Jackson's (1985) Chicken puzzle, we have Atilla and Genghis driving their chariots toward each other. There will be a collision if neither swerves or if both do, but a collision will be averted if only one swerves. Both are proud and, most likely, neither will swerve. The problem is to explain the intuition that all three of the sentences in (6.28) are simultaneously true in this scenario.

(6.28) a. Atilla ought to swerve.
b. Genghis ought to swerve.
c. It's not the case that Atilla and Genghis ought to both swerve.

It is very hard to account for this situation if *ought* is a universal quantifier over possible worlds, but it has a simple solution in the proposed semantics. We have possibilities represented by the following four worlds, along with probabilities and utilities which are in line with the story as Jackson tells it. (The probabilities are assigned assuming that each player makes the choice whether to swerve independently, and that each will swerve with probability 0.1. The actual numerical values that the utility function μ_D^W outputs are not meaningful, of course, but only the relative distance between values.)

World w	Description of w	Collision?	$prob(w)$	$\mu_D^W(w)$
w_1	Neither swerves	yes	.81	$^-100$
w_2	Atilla swerves, Genghis does not	no	.09	$^+50$
w_3	Genghis swerves, Atilla does not	no	.09	$^+50$
w_4	Both swerve	yes	.01	$^-100$

Table 2 Worlds, probabilities, and utilities in the Chicken game.

Suppose for simplicity's sake that $ought(\phi)$ is true if and only if $\mathbb{E}(\phi)$ is greater than $\mathbb{E}(\neg\phi)$ (as in Goble 1996, Levinson 2003; I will propose a slightly more general analysis below, but the result is the same in this case).

Atilla swerves in worlds w_2 and w_4, so (6.28a) is equivalent to $ought(\{w_2, w_4\})$. On the proposed semantics for *ought* this is true if and only if the average utility of these worlds, weighted by their conditional probability given that the actual world is in $\{w_2, w_4\}$, is greater than the related quantity computed for the complement of $\{w_2, w_4\}$.

$$[\![(6.28a)]\!]^{\mathcal{M},w,g} = 1 \text{ iff } \mathbb{E}(\{w_2, w_4\}) > \mathbb{E}(\{w_1, w_3\})$$

The right-hand side of the biconditional expands to

$$\begin{aligned} & \mu_{\mathcal{D}}^{W}(w_2) \times prob(\{w_2\}|\{w_2, w_4\}) + \mu_{\mathcal{D}}^{W}(w_4) \times prob(\{w_4\}|\{w_2, w_4\}) \\ > \; & \mu_{\mathcal{D}}^{W}(w_1) \times prob(\{w_1\}|\{w_1, w_3\}) + \mu_{\mathcal{D}}^{W}(w_3) \times prob(\{w_3\}|\{w_1, w_3\}) \end{aligned}$$

which, following Table 2, is true if and only if

$$50 \times .9 - 100 \times .1 > -100 \times .9 + 50 \times .1.$$

Atilla ought to swerve is true, since $45 - 10 = 35$ is greater than $-90 + 5 = -85$. The situation with *Genghis ought to swerve* is precisely parallel, and (6.28b) is also correctly predicted to be true.

Unlike quantificational theories, however, these facts are compatible with the truth of (6.28c) on the proposed semantics. The latter sentence comes out true if and only if $\mathbb{E}(\{w_4\})$ is **not** greater than $\mathbb{E}(\{w_1, w_2, w_3\})$, i.e. if and only if

$$\begin{aligned} \mu_{\mathcal{D}}^{W}(w_4) \times prob(\{w_4\}|\{w_4\}) \not> \; & \mu_{\mathcal{D}}^{W}(w_1) \times prob(\{w_1\}|\{w_1, w_2, w_3\}) \\ + \; \mu_{\mathcal{D}}^{W}(w_2) \times prob(\{w_2\}|\{w_1, w_2, w_3\}) + & \mu_{\mathcal{D}}^{W}(w_3) \times prob(\{w_3\}|\{w_1, w_2, w_3\}) \end{aligned}$$

which, consulting Table 2 and rounding off a few of the conditional probabilities, means that

$$-100 \times 1 \not> .82 \times -100 + .09 \times 50 + .09 \times 50$$

which is true, since -100 is not greater than -73. This is the result that we needed here: all three of the sentences in (6.28) are true in this scenario, and so the problem is dissolved.

In addition to getting the intuitively correct result here, the scenario also demonstrates what is generally necessary for $ought(\phi)$ and $ought(\psi)$ to both be true while $ought(\phi \wedge \psi)$ is false: essentially, $\phi \wedge \psi$ has to both very unlikely and very bad. The scenario that Jackson gives us has this special feature, but usually these conditions will not be fulfilled. In cases in which $ought(\phi)$ and $ought(\psi)$ are true, it is often likely that $ought(\phi \wedge \psi)$ is true as well. This accounts for the widely held feeling that this inference is a reasonable one: usually it is, but it is not logically valid.

6.4.2 Information-Sensitivity

In chapter 5 I argued that quantificational semantics for modals does not allow for sufficiently fine-grained interactions with probabilistic information. The present theory, being built around probability-weighted preference orders, is designed to do just this. Here I will briefly go through the puzzles noted there and show that the semantics I am proposing delivers the right results.

6.4.2.1 Medicine, Insurance, etc.

The puzzles involving a doctor's choice of medicine and a homeowner's choice of insurance in chapter 5 were taken from Goble (1996) and Levinson (2003) respectively. Both of these authors make proposals closely related to the one I have given for their respective domains; I will briefly describe how the solutions go, referring the reader to these papers for more details.

In the medicine scenario, we had two mutually exclusive options, medicine *A* and medicine *B*. If *A* was given, there was a 10% chance of a complete recovery and a 90% chance of death, while giving *B* would result in a partial recovery with certainty. The problem is that intuition suggests that the doctor ought to give medicine *B*, but quantificational semantics recommends *A* because all of the best worlds are worlds in which he gives *A*.

This is the classic form of a decision-theoretic problem, and it has a simple solution: the doctor should take the action with a higher probability-weighted preference. Suppose for simplicity that there are just three worlds. w_1, where the patient has a complete recovery, has the value +100; w_2, where the patient dies, has the value -100. w_3, where the patient has a partial recovery, has some middling value, say +20. Given the probabilities above, the expectations of giving *A* and giving *B* are

(6.29) a. $\mathbb{E}(\textit{Doctor gives A}) = +100 \times .1 + -100 \times .9 = -80$

b. $\mathbb{E}(\textit{Doctor gives B}) = +20 \times 1 = +20$

The expectation of giving *B* is much greater than the expectation of giving *A*, even though the best worlds that are not ruled out by our knowledge are worlds in which the doctor gives *A*. This is the result we want.

Now, it is not yet clear how this fits with the proposed semantics for $\mathcal{D}$-modals. What, for example, is to prevent θ_{ought} from being lower than -100, giving us the result that the doctor both ought to give *A* and ought to give *B* (although he can't)? Goble discusses several options, but the simplest is to suppose (as we did in the discussion of Chicken) that *ought* ϕ is true if and only if $\mathbb{E}(\phi)$ is greater than $\mathbb{E}(\neg\phi)$. Since the doctor cannot give both *A* and *B*, and (we are supposing) will definitely give one or the other, the effect is that *The doctor ought to give A* comes out false, and *The doctor ought to give B* comes out true. As a result θ_{ought} is constrained to be greater than $\mathbb{E}(\textit{Doctor does not give B}) = \mathbb{E}(\textit{Doctor gives A}) = -80$ in this context.

Goble also suggests that θ_{ought} is in some cases determined by comparing the expectation of a proposition not to the expectation of its negation, but to the expectations of a set of relevant alternatives. This is close to the analysis of Sloman (1970) (mentioned in connection to von Fintel & Iatridou (2008) in chapter 5) and to the detailed semantics that I will propose later in this chapter. It also ties in closely with the alternative-based semantics for *likely* and *probable* from chapter 4, as we will see in some detail below. In this case, however, the two proposals are equivalent: since there are only two options, *A* and *B*, not choosing *A* means choosing *B* are vice versa.

Levinson's (2003) solution to the insurance paradox has essentially the same form. Similarly to Goble's first proposal regarding *ought*, Levinson argues that *x wants* ϕ is true if and only if $\mathbb{E}(\phi) > \mathbb{E}(\neg\phi)$ according to *x*'s personal utility and probability functions. I will also argue for an alternative-based formulation of this semantics later, but let's see first how the simpler proposal accounts for the puzzle.

Recall that in the insurance puzzle we had the relevant options (6.30) and the intuitive constraints on reasonable preference (6.31):

(6.30) w_1: I do not buy insurance and my home burns down
w_2: I do not buy insurance and my home does not burn down
w_3: I buy insurance and my home does not burn down
w_4: I buy insurance and my home burns down

(6.31) $w_2 \succ w_3 \succ w_4 \succ w_1$

The problem was that quantificational semantics predicts that *I want to buy insurance* is false in any model where the preference order in (6.31) is satisfied, regardless of how probable the various events are or how strong the agent's preferences are.

The essential feature of quantificational semantics for *want* that generates this problem is that it does not take into account *how much* the agent prefers not spending money to spending money, and how the size of this interval compares to *how much* he prefers having insurance to not having insurance if his house burns down. Intuitively, if the latter preference is much stronger than the former, then the agent might want to buy insurance if he thinks that the risk of a home fire is sufficiently great.

Suppose that the probability of a fire is .05. (We don't have to worry about conditional vs. unconditional probabilities here, because buying insurance and having a home fire are probabilistically independent.)[9] For now familiar reasons, both the difference in utility and the odds having a fire will be relevant the the expectation of the various propositions. In particular, the odds of having no fire are 19-to-1. As a result the sentence *I want to buy insurance* will come out true, on Levinson's proposal, just in case the difference in utility between having insurance and not having insurance if there is a fire ($\mu(w_4) - \mu(w_1)$) is more than 19 times as great as his preference for not spending money if there is no fire ($\mu(w_2) - \mu(w_3)$).[10]

One model which fulfills these constraints has:

(6.32) $\mu(w_1) = -200$
$\mu(w_2) = +100$

---

9 One of the nice features of probability-weighted preference is that the agent generally doesn't need to bother assigning a probability to the event of his buying insurance; if this is the choice under consideration, all of the probabilities will be conditional probabilities where the conditioning event is the choice in question, and the unconditional probability of buying insurance drops out.

The insurance scenario is further simplified by the fact that whether or not the agent buys insurance has no effect on whether there will be a fire; or at least there is no causal connection, which is what we are interested in here. I am assuming, as in causal decision theory, that it is not relevant to expectation that the probability of a fire and the agent's choices with regard to insurance-buying might be probabilistically related via some third event, e.g. a personality feature of the agent. See Nozick (1969); Gibbard & Harper (1978) among many others for examples showing that failing to require a causal connection leads to absurd results in certain cases.

Incidentally, the claim that causality is relevant might seem to be at odds with my earlier claim that obligations etc. uniformly attach to propositions, given that causal decision theorists generally distinguish acts from propositions. However, causal decision theory can also be formulated without making this distinction rigid: see Joyce 1999 (especially ch.5) for discussion and arguments that causal decision theory should treat acts as a special case of propositions.

10 Here and in the next section I suppress the super- and subscripts on the measure functions for readability; unless otherwise noted these μ represent $\mathcal{S}^W_{\mathcal{D}}$-admissible measure functions $\mu^W_{\mathcal{D}}$.

$\mu(w_3) = +95$
$\mu(w_4) = +70$

Here we are supposing that buying insurance is worse than not buying insurance if there is no fire, so that w_2 is better than w_3. However, the loss of having an uninsured house burn down is far greater than the gain of not paying for insurance. The expectations of *I buy insurance* and *I do not buy insurance* are then

(6.33) a. $\mathbb{E}(\textit{I buy insurance}) = \mathbb{E}(\{w_3, w_4\}) = 95 \times .95 + 70 \times .05 = +93.75$
b. $\mathbb{E}(\textit{I do not buy insurance}) = \mathbb{E}(\{w_1, w_2\}) = 100 \times .95 + -100 \times .05 = +85$

If *x wants* ϕ means that $\mathbb{E}(\phi) > \mathbb{E}(\neg\phi)$, as Levinson argues, then *I want to buy insurance* will come out true in this scenario even though the best worlds are worlds in which I do not. This is the intuitively correct result: it is reasonable for me to want to buy insurance even if I know that it probably will not do me any good, if I also think that it would be disastrous to have a fire and no insurance.

In §6.5 I will propose a refinement of Goble's and Levinson's proposals which adds alternative-sensitivity and a significance parameter for mid-scalar deontic items such as *want* and *ought*, making them semantically close to *likely* and *probable* as these items were analyzed in chapter 4. For the cases at hand, however, the analysis is essentially the same: *ought*(ϕ) and *x wants*(ϕ) are not quantifiers, but expressions which establish a threshold value and compare the probability-weighted preference of their propositional argument to this threshold. This account allows us to resolve a deep problem with quantificational semantics caused by the coarse-grained way in which it interacts with information.

6.4.2.2 The Miner's Paradox

The Miner's Paradox discussed in Kolodny & MacFarlane (2010) can be given essentially the same analysis as Goble's and Levinson's puzzles. The solution has the added benefits of being more general and theoretically motivated than Kolodny & MacFarlane's proposal, and doing without the problematic features identified in the chapter 5: our solution does not abandon the methodologically desirable constraint enforcing Stability in world-orderings, and it accounts for the intuition that the crucial sentences are true (not just consistent) in the scenario described.

In the mining disaster we have the following situation –

> Ten miners are trapped either in shaft A or in shaft B, but we do not know which. Flood waters threaten to flood the shafts. We have enough sandbags to block one shaft, but not both. If we block one shaft, all the water will go into the other shaft, killing any miners inside it. If we block neither shaft, both shafts will fill halfway with water, and just one miner, the lowest in the shaft, will be killed.

and we want to find a semantics which makes sentences (6.34)-(6.36) true.

(6.34) We ought to block neither shaft.

(6.35) If the miners are in shaft A, we ought to block shaft A.

(6.36) If the miners are in shaft *B*, we ought to block shaft *B*.

The structure of this example is essentially similar to the examples in medicine and insurance scenarios. In each case, it would be easy to choose which action to take (buy insurance or don't, give medicine *A* or *B*, block shaft *A* or *B*) if we knew what conditions hold. However, the structure of the situation makes it better to adopt a course of action which is guaranteed to be sub-optimal, but is somehow still the best available course of action given our state of knowledge. The puzzle is essentially how to incorporate our imperfect knowledge into the semantics of *ought* in a way that returns the intuitively correct result. From the perspective I have adopted, though, this is no puzzle at all: it is a textbook decision theory problem, and admits of a textbook solution.

The scenario invites us to consider the following possibilities:

(6.37) w_1: We block *A* and they are in *A*
w_2: We block *A* and they are in *B*
w_3: We block *B* and they are in *A*
w_4: We block *B* and they are in *B*
w_5: We block neither shaft and they are in *A*
w_6: We block neither shaft and they are in *B*

Assuming that each life is individually and equally valuable, and that no other relevant facts have been omitted from the story, the value of each world is a positive linear function of the number *n* of lives saved in that world: for all admissible μ', $\mu'(w) = \alpha \times n + \beta$, with $\alpha > 0$. One such function (with $\alpha = 1$ and $\beta = 0$) is $\mu(w) = n$, the function that assigns to each world the number of miners whose lives are saved.

According to the story, this μ assigns the following value to each of our six worlds:

(6.38) $\mu(w_1) = 10$
$\mu(w_2) = 0$
$\mu(w_3) = 0$
$\mu(w_4) = 10$
$\mu(w_5) = 9$
$\mu(w_6) = 9$

Since we have no idea which shaft they are in, presumably the probability that they are in *A* and that probability that they are in *B* are both equal to .5. That is,

(6.39) $prob(\{w_1, w_3, w_5\}) = prob(\{w_2, w_4, w_6\}) = .5$

Since the miners' location is independent of our actions, the probability that the miners are in *A* is the same — 0.5 — whether we block *A*, block *B*, or do nothing. So, for example, $prob(w_1|\{w_1, w_2\}) = prob(\{w_1, w_3, w_5\}) = 0.5$.

We can calculate the expectations of the various possible actions as follows:

(6.40) a. $\mathbb{E}(Block\ A) = \mu(w_1) \times prob(w_1|\{w_1, w_2\}) + \mu(w_2) \times prob(w_2|\{w_1, w_2\})$
$= .5 \times 10 + .5 \times 0 = 5$

b. $\mathbb{E}(Block\ B) = \mu(w_3) \times prob(w_3|\{w_3, w_4\}) + \mu(w_4) \times prob(w_4|\{w_3, w_4\})$
$= .5 \times 0 + .5 \times 10 = 5$

c. $\mathbb{E}(\textit{Block neither shaft}) = \mu(w_5) \times prob(w_5|\{w_5, w_6\}) +$
$\mu(w_6) \times prob(w_6|\{w_5, w_6\})$
$= .5 \times 9 + .5 \times 9 = 9$

Probability-weighted preference has a very direct interpretation in this case: the expectation of each action is just the expected number of lives saved. If we block *A* or *B*, the expected number of lives saved given our information is five, since there is a 50% chance of saving ten and a 50% chance of saving zero. If we block neither, the expected number of lives saved is 9. So we ought to block neither.

Of course, we still have to say something about what *ought* means in order to be sure that we get this result, but we can be pretty sure that the result will follow on any plausible scalar semantics for *ought*. The miminal constraints seems to be that there is some action which we ought to take here — i.e. that it's not acceptable to choose an action at random — and that the semantics of *ought* obeys the uncontroversial principle in (6.41):

(6.41) If $ought(\phi)$ is true, ϕ and ψ are incompatible, and ϕ is much better than ψ, then $ought(\psi)$ is false.

For example, both of Goble's (1996) proposals for the meaning of *ought* obey (6.41), and both get the right result here. This is obvious if *ought* compares alternatives, and also follows if *ought* ϕ compares ϕ to $\neg\phi$: the negation of *Block neither* is *Block A or Block B* in this context, and the expectation of this proposition is necessarily 5 as well.

(6.35) and (6.36) are also guaranteed to come out true, as long as we adopt a restrictor analysis of conditionals along the lines of Kratzer (1986), which we saw informally at the beginning of chapter 5. This is standard in linguistic semantics, and Kolodny & MacFarlane (2010) adopt a variant of this analysis as well. On this approach, the antecedent of the conditional restricts the body of information relative to which the consequent is evaluated to one in which the antecedent holds throughout. In effect, *If the miners are in A, we ought to block A* is true if and only if *we ought to block A* comes out true when we temporarily ignore worlds in which the miners are not in *A*.

In the case at hand, this is equivalent to finding $\mathbb{E}(\textit{block A}|\textit{the miners are in A})$, the **conditional expectation** of blocking A on the assumption that they are in A. We can do this by expanding the definition of expectation and adding *the miners are in A* as a further conditionalizing factor in the probability measure:

(6.42) $\mathbb{E}(\textit{Block A}|\textit{They are in A}) = \mu(w_1) \times prob(w_1|\textit{we block A and they are in A})$
$= \mu(w_1) \times prob(w_1|\{w_1, w_3, w_5\} \cap \{w_1, w_2\})$
$= \mu(w_1) \times prob(w_1|\{w_1\}) = 10 \times 1 = 10$

That is, if we knew that the miners were in *A*, then the expected number of lives saved if we take the action of blocking *A* would be 10. As long as this semantics for conditionals is viable, then, we have the result that *If the miners are in A, we ought to block A* will come out true on anybody's semantics for *ought*. The same holds, mutatis mutandis, of (6.36); and so we have the result (6.34)-(6.36) are consistent and true in the intended model, as desired.

Naturally, we also want (6.43) to come out false in this scenario, even though all of the best worlds — namely w_1 and w_4 — are worlds in which *we block A or we block B* holds.

(6.43) We ought to either block shaft *A* or block shaft *B*.

Avoiding the prediction that (6.43) is true is a problem for quantificational semantics, and led Kolodny & MacFarlane (2010) to considerable lengths in rejecting the well-motivated Stability constraint on deontic orderings. For us, however, this is no problem at all: assuming the obvious logical form, we assign it the truth-conditions in (6.44), which are clearly false.

(6.44) $[\![(6.43)]\!]^{\mathcal{M},w,g} = 1$ iff $\mathbb{E}(\{w_1, w_2, w_3, w_4\}) > \mathbb{E}(\{w_5, w_6\})$

Both $\mu(w_5)$ and $\mu(w_6)$ are equal to 9, and so $\mathbb{E}(\{w_5, w_6\}) = 9$; but $\mathbb{E}(\{w_1, w_2, w_3, w_4\}) = 5$ — on average blocking either *A* or *B* will save 5 lives — and so (6.43) is false because the expected number of lives saved is not greater than $\mathbb{E}(\{w_5, w_6\}) = 9$.

With respect to each of the three desiderata outlined in chapter 5, the semantics given here improves on Kolodny & MacFarlane's (2010). My theory captures the crucial result that Kolodny & MacFarlane are seeking, in that it renders (6.34)-(6.36) logically compatible. However, it goes further by also showing why these sentences are intuitively **true** in the scenario at hand, rather than simply blocking the inconsistency proof (desideratum one). It does so while also making clear and precise predictions about the way in which information influences the deontic ordering over propositions (desideratum two). Finally, my account of the Miner's Paradox avoids Kolodny & MacFarlane's problematic notion of *serious information-dependence*, according to which information can manipulate the deontic ordering over worlds. I argued in chapter 5, following Charlow 2011, that this feature of their semantics is unmotivated and undesirable. The present theory does well to capture the information-dependence of the deontic ordering over propositions while holding the ordering over worlds constant. The crucial feature that makes it possible to do this is simply that *modals are not quantifiers over possible worlds*.

6.5 Gradability and the Typology of Deontic and Bouletic Modals

The results of §6.4 make a strong case for the expectation-based theory of deontic and bouletic modals. Starting with only an interval-scale preference order and probability measures which are needed for other purposes anyway, we constructed a skeletal threshold-based semantics for $\mathcal{D}$-modals which accounts for each of the first two sets of problems for quantificational semantics that were discussed in chapter 5. In this section we will deal with the other three, and give details of the semantics of the various deontic and bouletic modals.

6.5.1 Gradability and Comparison

In contrast to quantificational theories, the scalar theory has no difficulty in making sense of gradability and comparison in the modal domain. For instance, we can treat proposition-embedding *good* as having a lexical entry very similar to the one we assigned to *likely/probable* in chapter 4: it is a measure function which maps propositions to their proposition-goodness, where the latter is defined as the expected world-goodness worlds of the worlds in the proposition.

(6.45) $[\![good]\!]^{\mathcal{M},w,g} = \lambda p_{\langle s,t \rangle}[\mathbb{E}(p)]$

As in the above, I do not always specify how the preference order which is used to calculate expectation is selected. Presumably, in general with *good* it will be moral or practical goodness, with *want* it will be personal preference, etc. These complex and context-dependent issues are more or less orthogonal to the present discussion, focusing as it does on the structural aspects of these domains. All that is required is that the relevant scales are interval orders over propositions calculated via probability-weighted preference from interval orders over worlds, and it does not matter for our purposes how the latter are derived.

Assuming (6.45), the theory of comparatives that we borrowed from Kennedy (1997, 2007) in chapter 1 predicts the following truth-conditions for sentences of the form ϕ *is better than* ψ (making the obvious adjustments for propositional rather than individual arguments):

(6.46) ϕ *is better than* ψ is true iff $\mathbb{E}(\phi) > \mathbb{E}(\psi)$.

We can treat verbal comparatives with *want* and *need* along the same lines (again taking care to ensure that *more* and other degree operators have appropriate type-polymorphic denotations):

(6.47) $[\![want]\!]^{\mathcal{M},w,g} = \lambda p_{\langle s,t\rangle}\lambda x_e[\mathbb{E}_x(p)]$, where $\mathbb{E}_x(\cdot)$ is calculated with reference to x's beliefs and preferences.

(6.48) x *wants* ϕ *more than* ψ is true iff $\mathbb{E}_x(\phi) > \mathbb{E}_x(\psi)$.

And so on for the other $\mathcal{D}$-modals which occur in the comparative.

Regarding incomparability, recall from chapter 5, §5.4.1 that Kratzer's theory of modality generates an unacceptable amount of modal incomparability, and that this problem is even more clearly problematic with the $\mathcal{D}$-modals than it is with epistemic modals. For the expectation-based theory that I have introduced, the situation is much different: degrees of obligation/desire/etc. are provided by the set of all $\mu^W_{\mathcal{D}}$ which are admissible according to $\mathcal{S}^P_{\mathcal{D}}$. Since the latter is an interval scale, there is, in the simplest case, no deontic incomparability at all. As a result, the semantics does not prevent us from comparing violations of norms: unlike Kratzer's theory, we have no difficulty in assigning a truth-value to *Tresspassing is better than murder*.

However, as in the scalar approach to epistemic comparatives discussed in chapter 3 (§3.6.3), it may well be that there are genuine deontic incomparabilities, and it is possible to modify the semantics slightly to accommodate them if this seems advisable. The device is the same one that was proposed by van Rooij (2009) to account for multidimensional adjectives such as *clever* and *big*, and essentially involves treating $\mathcal{S}^W_{\mathcal{D}}$ as a **set** of interval orders. By this device, we can introduce exactly as much incomparability as the data warrant, and no more.

As for degree modification, the predictions are again straightforward: since $\mathcal{S}^P_{\mathcal{D}}$ is in every case an interval order, it is able to support degree modifiers such as *very (much)*, for the same reason that the interval order underlying temperature allows us to speak of something's being *very hot*. It is noteworthy that the degree modifiers that we saw in chapter 5, and those that show up most frequently in corpora, are indeed ones which carry relatively weak quantitative information (*very, rather, extremely*, etc.). The fact that $\mathcal{S}^P_{\mathcal{D}}$ is an interval order with no inherent upper or lower bound makes further predictions: we do not expect to find modifiers which rely on the presence of upper- or lower-bounds, or proportional modifiers. In general, this prediction seems to be correct (though I have not yet undertaken an exhaustive survey).

(6.49) a. # It is slightly/completely permissible for you to leave.[11]

b. # It is half/completely/almost/70% obligatory for you to leave.[12]

In general, unlike quantificational theories in general and Kratzer's in particular, the expectation-based theory developed here makes reasonable predictions about degree modification; future work will hopefully confirm the predictions made here or allow us to refine the theory appropriately.

6.5.2 A Three-Way Typology of Deontic and Bouletic Modals

I tried to say very little about the lexical semantics of specific $\mathcal{D}$-modals while accounting for the first two sets of puzzles from chapter 5. The reason for doing this was simply that the solution to those problems is mostly independent of the specific proposals regarding these items that I will now make. The crucial feature of my account of the Miner's Paradox, for example, was to give up thinking of modals as quantifiers over the "best" worlds, and to think of them instead as expressions which relate adjectives to points on scales built on probability-weighted preference and compare them to some threshold value. The precise way in which this threshold was not too important for these purposes; all that was needed was the obvious principle that *ought* ϕ and *ought* ψ cannot both be true if ϕ and ψ are mutually exclusive and one is clearly better than the other. As a result, the solution does not rely heavily on the discussion of modal types that we are about to undertake.

In this section I will give more detail about the semantics of modals of various strengths — what we have called, following Horn (1989), WEAK SCALAR, MID-SCALAR, and HIGH SCALAR modals. These first and last of these correspond to what are usually known as "weak" and "strong" modals in modal logic and semantics, and the mid-scalar modals are what von Fintel & Iatridou (2008) call "weak necessity" modals. I will argue that this typology is related to the typology of adjective strengths that is familiar from previous chapters, but with some wrinkles. Most clearly, probability-weighted preference is a totally open scale; as a result we cannot simply identify the weak scalar $\mathcal{D}$-modals with minimum-standard scalar expressions, or the high scalar $\mathcal{D}$-modals with maximum-standard scalar expressions. As we saw, minimum- and maximum-standard adjectives can only associate with scales which have minimum and maximum elements, respectively.

Nevertheless, there is a close intuitive connection between the three modal strengths and the typology of adjectives:

11 *Slightly permissible* occurs on 243 Google pages as of 5 April 2011, which make up 0.0013% of occurrences of *permissible*. By comparison, *slightly tall* occurs about 43,300 times, making up approximately 0.017% of hits for *tall*, or about 10 times as many as a proportion of total hits for the adjective. Whatever factors allow *slightly tall* to be acceptable in certain special contexts, something similar is presumably responsible for the even less frequent acceptability of *slightly permissible*.

12 The strings *almost/completely obligatory* do occur occasionally, but very rarely in a deontic sense. Much more common are examples such as *Growing old is completely obligatory. Growing up is optional* (from http://sanitythief.deviantart.com/), or *It is almost obligatory to decorate casts with signatures or stickers* (http://www.drug3k.com/forum1/Injuries/Is-it-possible-to-put-a-sticker-or-decals-on-an-arm-cast-60223.htm). Both of these involve a different sense of *obligatory*, apparently having to do with inevitability or frequency. Both of these are plausibly associated with upper-bounded scales. The very few examples I have found of *almost/completely obligatory* with a deontic sense are plausibly cases of metalinguistic modification, similarly to attested examples of *almost/completely tall*.

- **Mid-scalar** $\mathcal{D}$-modals such as *want*, *ought*, *should*, and *good* (in the positive form) compare their propositional argument to a threshold determined by the values of contextual alternatives, like relative adjectives, *likely*, and *probable*;
- **High scalar** $\mathcal{D}$-modals such as *must*, *need*, *require*, and *obligatory* (in the positive form) establish a very high threshold, reminiscent of high degree adjectives such as *gigantic* and *ecstatic*;
- **Weak scalar** $\mathcal{D}$-modals measure deviation from a low threshold which is defined in terms of the negation of the corresponding high scalar modals, as in traditional accounts.

Throughout this section, I assume the presence of the silent **pos** morpheme where appropriate, including at least *good, want, need, require*, and perhaps also *ought* and *should*.

6.5.3 Mid-Scalar $\mathcal{D}$-Modals and Alternatives

For mid-scalar $\mathcal{D}$-modals I want to argue for something quite close to Goble's (1996) proposed semantics for *ought*, and to the proposal for *likely* and *probable* that I made in chapter 4. Essentially, the idea is that these modals establish a threshold value on the basis of the distribution of values among a set of contextual alternatives, just like relative adjectives such as *tall*. Evidence for this conclusion will come from the fact that these items are robustly focus-sensitive, and that they fail to license certain entailments that would be expected if the threshold were fixed in some other way.

6.5.3.1 *Want*

Villalta (2008) points out that *want* and several of the other items that I am calling "$\mathcal{D}$-modals" are sensitive to contextual alternatives in a way that is unexpected on standard theories of its semantics (e.g. Heim 1992; von Fintel 1999). Heim pointed out that *want* only pays attention to epistemically possible worlds; however, as Villalta suggests, focus can affect the interpretation of *want* so that it considers only a subset of epistemically possible worlds. Consider:

(6.50) a. **Mother**: I want you to go to the grocery store.
b. **Son**: I don't want to go to the grocery store. I want to go to a movie.
c. **Mother**: Well, I want you to go to a CLEAN movie, then.

The reply in (6.50c) may well indicate weak parenting skills, but it does not contradict the preference stated in (6.50a). On the other hand, suppose that the mother had replied instead:

(6.51) **Mother**: Well, I want you to go to a movie.

Unlike (6.50c), this reply involves blatant self-contradiction. But this is strange, because, on standard quantificational semantics for *want*, (6.50c) is predicted to entail (6.51).

More or less following Villalta, a plausible explanation is that the value of θ_{want} is being calculated on the basis of different alternative sets in these examples.[13] In (6.50a) the mother states

13 The proposal in Villalta (2008) was an important source of inspiration for this section, although her implementation in terms of selection functions and quantification over worlds falls foul of several of the other problems discussed in

her preference for the son to go to the grocery store as opposed to some other set of activities — not going to the grocery store, perhaps, or else activities like staying at home, going to the swimming pool, going to a movie. Either of the alternative sets in (6.52) is reasonable in a context of this sort:

(6.52) a. $\mathbf{ALT}(6.50a) = \{$Son goes to the grocery store, son does not go to the grocery store$\}$

b. $\mathbf{ALT}(6.50a) = \{$Son goes to $X \mid X \in \{$grocery store, swimming pool,movie theater, ...$\}\}$

In contrast, on standard theories of focus, an utterance of (6.50c) will introduce a quite different set of alternatives:

(6.53) $\mathbf{ALT}(6.50c) = \{$Son goes to an X movie $\mid X \in \{$clean, violent, racy, ...$\}\}$

Now, in chapter 4 we adopted a semantics for relative adjectives, *likely*, and *probable* which makes the threshold value sensitive to the distribution of values among in the domain of the adjective. As we saw, this domain can be supplied contextually or can be temporarily restricted by devices such as focus and comparison classes. For instance,

(6.54) *x is* **pos** *tall* is true iff *x*'s height exceeds θ_{tall}, where θ_{tall} is a value significantly greater than the expected height of a member of its domain.

I propose similarly for *want* and other mid-scalar $\mathcal{D}$-modals:

(6.55) *x wants* ϕ is true iff $\mathbb{E}(\phi) \geq \theta_{want}$, where θ_{want} is a value significantly greater than $\mathbb{E}(\bigcup \mathbf{ALT}(\phi))$.

$\mathbb{E}(\bigcup \mathbf{ALT}(6.50a))$ is just the weighted average of the expectations of the propositions in **ALT**(6.50a).[14] Suppose that **ALT**(6.50a) includes all of the places that the son might go. The expectation of this set is the same as the expectation of a tautology (as is the expectation of the set in (6.52a), since $\bigcup\{\phi, \neg\phi\} = \top$).

(6.56) If Φ is a set of propositions such that $prob(\bigcup\Phi) = 1$, then $\mathbb{E}(\bigcup\Phi) = \mathbb{E}(\top)$. (cf. Jeffrey 1965b: 81-2)

The reason is simply that all of the probability mass in W is concentrated in the members of Φ, and so any $\psi \notin \Phi$ receives weight 0 in calculating expectation.

$\top$ also happens to be the point of **indifference**: if $\mathbb{E}(\chi) = \mathbb{E}(\top)$, then $\mathbb{E}(\neg\chi) = \mathbb{E}(\top)$ as well, and so it is completely unimportant which of χ and $\neg\chi$ holds. If the disjunction of the alternative set has probability 1, as in (6.52), we can simplify (6.55):

(6.57) *x wants* ϕ is true iff $\mathbb{E}(\phi)$ is significantly greater than the point of indifference $\mathbb{E}(\top)$.

In unmarked contexts, this does seem to be what *want* means, and it is not too different from the proposals of Heim (1992) and Levinson (2003). However, there is a subtle difference. Heim and

chapter 5. I think that the core insights of that paper can be retained by adopting a more thoroughly scalar semantics and probabilistic representation of information, though, at least with respect to the gradability and focus-sensitivity of desire verbs.

14 In case they are disjoint. It doesn't matter for this calculation whether they are, though: $\mathbb{E}(\cdot)$ returns sensible values for overlapping alternative sets as well, supposing that this is the right way to go (cf. Beaver & Clark 2008). In this case the expectation of the set can't be found by taking a weighted average of the expectations of the propositions in it, though.

Levinson build into the lexical semantics of *want* the requirement that a proposition be preferred to its own negation. As a result, if ϕ and ψ are incompatible then *x wants* ϕ and *x wants* ψ cannot both be true.[15]

Now, the present proposal also predicts that *ought* ϕ entails that ϕ is prefered to $\neg\phi$, as long as the alternative set contains both ϕ and $\neg\phi$. However, since $\mathbf{ALT}(\phi)$ may not contain $\neg\phi$, it is also possible to have a richer alternative set with two or more incompatible ϕ, ψ such that *x* wants both. (6.58) indicates that this is a welcome prediction:[16]

(6.58) I want to go to Paris for the July summer school, and I want to go to Rome for the month-long festival in July. Obviously it's impossible for me to be in both places at the same time. What do you recommend that I do?

Most theories, including Kratzer's (1991) and Villalta's (2008), cannot account for this example. Levinson (2003) has a way to deal with it, but he does so by allowing that an agent can have multiple inconsistent preference orders. The semantics for *want* proposed in (6.55) gives us a more straightforward solution: (6.58) is true just in case going to Paris and going to Rome both have significantly greater expected value than the point of indifference. As long as there are enough alternatives — including other possible destinations which are below the point of indifference — (6.58) is predicted to describe a consistent set of desires.

This same proposal also accounts for the focus-sensitivity of (6.50c) and the fact that it is compatible with (6.50a). The key is to notice that **ALT**(6.50c) does not exhaust all of the epistemically possible scenarios, and so $\mathbb{E}(\bigcup\mathbf{ALT}(6.50c))$ is not equal to $\mathbb{E}(\top)$. In fact, (6.50c) can come out true even if going to a movie is massively dispreferred on a global level. Because of the way that focus alternatives are defined, the disjunction of alternatives *Either Son goes to a violent movie, or he goes to a clean movie, or he goes to a racy movie, or ...* is equivalent to $\exists X$[*Son goes to an X movie*]. The latter, of course, is equivalent to the simpler *Son goes to a movie*.

As a result it follows from (6.55) that the value of θ_{want} relevant for evaluating a sentence like *I want you to go to a CLEAN movie* is set by reference to the expected value of *you go to a movie*. Specifically:

(6.59) *I want you to go to a CLEAN movie* is true iff $\mathbb{E}$(*You go to a clean movie*) is significantly greater than $\mathbb{E}$(*You go to a movie*).

This is a pretty good paraphrase of what the sentence means, and all that is needed for (6.50a) to come out true is that the expectation of *Son goes to a clean movie* be significantly greater than the expectation of *Son goes to a movie*. This can be true even if the latter has a low expectation, as it does in the case at hand.

More generally, in the present semantics *x wants* ϕ can be true when ϕ is dispreferred (i.e. when $\mathbb{E}(\phi) < \mathbb{E}(\top)$), but only if $\mathbb{E}(\bigcup\mathbf{ALT}(\phi))$ is even lower. Since this theory treats *want* as the

15 At least, on current assumptions. Levinson (2003) points out that Heim's (1992) proposal makes room for contradictory desires under some conditions, by complicating some of the usual assumptions of quantificational semantics. However, Heim's proposal still encounters most of the problems discussed in this chapter and the last, and my proposal is able to account for the puzzles involving (lack of) entailments from *want* that she discusses.

16 Note the similarity between this point and the discussion of alternatives with *likely* and *probable* in chapter 4, cf. also Yalcin (2010).

verbal counterpart of a relative adjective, the effect of focus here is comparable to the fact that someone can be tall for a jockey without being tall simpliciter: in each case we are using a device of temporary domain restriction to affect the calculation of the threshold.

There is much more that could be said about *want* here, and for the sake of space I have not tried to show that this semantics accounts for all of the logical problems and missing entailments of standard quantificational theories noted by Heim (1992) and Villalta (2008) (but it does). There are, however, several relevant issues involving *want* that need consideration in future work. First, I have not considered Heim's data involving presupposition projection from the scope of *want*; I think that this issue could be dealt with, but doing so would require more space than we can give it here. Second, I have not said anything about Villalta's (2008) observation that focus-sensitivity correlates closely with the use of the subjunctive mood in Spanish. Villalta argues that the subjunctive morpheme is actively involved in the evaluation of alternatives, and that it is only licensed when there is a focused constitutent. From the current perspective, the evaluation of alternatives does not depend on the presence of a mood marker (although conceivably Spanish and English differ in this respect). My suspicion is that the focus-sensitivity of $\theta_{\mathcal{D}}$ for mid-scalar $\mathcal{D}$-modals and the fact that they select the subjunctive in Spanish have a common cause in the general context-sensitivity of their threshold values. I must leave this issue to future work as well.

However these latter issues are resolved, the present theory has clear benefits over standard analyses of *want*, as well as its sources of inspiration in Levinson (2003) and Villalta (2008). Unlike Levinson, my theory explains focus-sensitivity and other types of context-sensitivity; and unlike either, my approach predicts the possibility of conflicting desires in a natural way. Most importantly, perhaps, the present analysis fits in neatly with the general scalar theory of modality proposed here, the evidence for scalarity, non-monotonicity, and information-sensitivity that we have seen, and an independently motivated alternative-based semantics for relative adjectives.

6.5.3.2 *Good*

The proposal made in §6.5.3.1, that the value of θ_{want} is calculated on the basis of alternatives, is not a special lexical semantic property of this expression (as Villalta 2008 assumes). Rather, this is a general feature of relative-standard expressions, including the epistemic modal adjectives *likely* and *probable* and, I claim, mid-scalar $\mathcal{D}$-modals in general. Here I make the case for *good*, and then turn to *should* and *ought* in the next section.

Good is a relative-standard adjective, just like *tall* and *likely/probable* as discussed in chapters 3 and 4.

(6.60) Degree Modification

a. Jeffrey is ✓very/??completely/#mostly/??slightly/#half tall.

b. It is ✓very/??completely/#mostly/??slightly/#half likely/probable that it will rain.

c. It is ✓very/??completely/#mostly/??slightly/#half good for children to obey their parents.

(6.61) Zone of Indifference

a. ✓Sam is not tall, but he is not short either.

b. ✓It is not likely that we will win, but is is not unlikely either.
c. ✓It is not good to spend money on expensive chocolates, but it is not bad either (it's morally indifferent).

As the existence of the zone of indifference would lead us to expect, *good* is a neg-raiser like *want* and *likely*:

(6.62) a. I don't want to leave. ⇝ I want to stay.
b. Mary is not likely to leave. ⇝ Mary is likely to stay.
c. It is not good for you to leave. ⇝ It is good for you to stay.

Like its fellow proposition-embedding adjectives *likely* and *probable*, *good* does not readily accept *for a NP*-type comparison classes: *# It is good for you to go home for a student* does not mean "It is better for you to go home than it is, on average, for students to go home". Also like these items — and like *want* — *good* allows focus to play the role of a comparison class, as we can see by contrasting (6.63a) and (6.63b).

(6.63) a. It is good that you spilled WHITE wine on the carpet.
b. It is good that you spilled wine on the carpet.[17]

We can account for the fact that (6.63a) does not entail (6.63b) in the same terms that we explained the focus-sensitivity of *want*: in the positive form, ϕ *is good* means ϕ *is significantly better than* $\bigcup\mathbf{ALT}(\phi)$. This is equivalent to the condition that $\mathbb{E}(\phi)$ is significantly greater than $\mathbb{E}(\bigcup\mathbf{ALT}(\phi))$. Since $\bigcup\mathbf{ALT}(\phi)$ is equivalent to the proposition *You spilled wine on the carpet*, the net effect is that (6.63a) compares $\mathbb{E}$(*you spilled white wine on the carpet*) to $\mathbb{E}$(*you spilled wine on the carpet*). This is, indeed, the intuitive meaning of (6.63a).

This predicts correctly that (6.63a) is reasonable, because spilling white wine is much better than spilling, say, red wine or rosé. This can be true even though all three options are sub-optimal relative to not spilling wine at all. Assuming that (6.63b) makes reference to a different alternative set — namely {*You spilled wine, you did not spill wine*} — this analysis accounts for the fact that (6.63a) does not entail (6.63b).

6.5.3.3 Information and Tense

Example (6.63) might appear to make trouble for me in a different way, though. Here is a quick refutation of my theory: when it embeds a finite clause, *good* is factive. So, anyone who utters (6.63a) **knows** that you spilled white wine on the carpet, and as a result the probability of any world in which you spilled any other kind of wine is zero. As a result, these worlds receive zero weight in the calculation of $\mathbb{E}$(*you spilled wine on the carpet*), and all of the probability mass in this proposition resides in the worlds where you spilled white wine on the carpet, with the result that $\mathbb{E}$(*you spilled wine on the carpet*) ends up being equal to $\mathbb{E}$(*you spilled white wine on the carpet*). So (6.63a), and any other sentence of the form *it is good that* ϕ, should be trivially false because its factivity affects the calculation of expectation in a way which guarantees that $\mathbb{E}(\phi) = \mathbb{E}(\bigcup\mathbf{ALT}(\phi))$.

17 As noted in chapter 4, I owe the observation that *good* is focus-sensitive to Dean Pettit.

The argument is valid, but it has a hidden premise which is false: there is an implicit assumption that the relevant probability measure must be tied to the speaker's information at the time of utterance. As von Fintel & Gillies (2008); Yalcin (2011); MacFarlane (2011) among others have discussed with regard to epistemic modals, there is considerable flexibility in determining what information state is relevant to the evaluation of an epistemically modalized sentences. An example from von Fintel & Gillies (2008: 87) makes the point clearly:

> Sophie is looking for some ice cream and checks the freezer. There is none in there. Asked why she opened the freezer, she replies:
>
> (6.64) a. There might have been ice cream in the freezer.
> b. PAST(*might*(ice cream in freezer))
>
> It is possible for Sophie to have said something true, even though at the time of utterance she knows (and so do we) that there is no ice cream in the freezer.

(6.64) indicates that, at least in certain past-tense sentences, the information state that is relevant to the evaluation of a sentence does not correspond to the information that any part of the conversation possesses at the moment of utterance. Instead, in this case the relevant information state is Sophie's, at the time of her decision to open the freezer door.

Similarly, I suggest, past-tense and factive deontic sentences can be linked to an information state that no one in the conversation holds at the time of utterance, but someone (say, the speaker) did hold at some relevant previous point in time. In the case of (6.63), the relevant information state for evaluating *It is good that you spilled WHITE wine on the carpet* is one in which no wine has yet been spilled. Roughly, we can imagine going back to most recent point in time at which each of the focus alternatives was a realistic possibility, and asking whether $\mathbb{E}($*you spill white wine*$)$ is significantly greater than $\mathbb{E}($*you spill wine*$)$ relative to the information that we had at that time.

Similarly, we can account for insurance-type scenarios in past-tense sentences with *should* and *ought*:

(6.65) You should have paid for there to be a doctor at the game. Someone could have gotten hurt.

If the relevant information state were necessarily linked to the time of utterance, (6.65) would be a rather silly thing to say according to the expectation-based semantics of *should* that I will propose in a moment. This is because, since we know that no one did get hurt, having a doctor at the game would have been an unnecessary expense relative to our **current** information state. The state which did in fact occur is optimal relative to our present information, since we saved money and no one got hurt.

Nevertheless, (6.65) is a reasonable reprimand because, at the time when the decision was made not to pay for a doctor to attend, the addressee did not know that no one would be hurt. If this information state supplies our probability measure, then (6.65) will not come out as trivially false: it may well be that in such a state of ignorance, the expectation of paying for a doctor (and having someone to treat possible injuries) is greater than the expectation of not paying (and running the risk of an injury going untreated).

With this caveat in mind, we turn to *should* and *ought*, which, I claim, have meanings very close to those of *good* and *want*.

6.5.3.4 Intermediate-Strength Modals: *Should* and *Ought*

There are several reasons for assigning a semantics to *ought* and *should* closely related to the one that we did to *want* and *good*. The first is that all four items are neg-raisers, as we saw in chapter 5, a fact which Horn (1989) takes to be diagnostic for occupying an intermediate position on the scale. The second point, also from Horn (1972, 1989), is that they are clearly weaker than the high scalar counterparts *must*, *have to*, and *obligatory*:

(6.66) a. You should/ought to wash the dishes — indeed, you must.
b. # You must wash the dishes — indeed, you ought to.

Facts like these motivated von Fintel & Iatridou (2008) to add a second ordering source to Kratzer's semantics in order to obtain the result that *should* and *ought to* are weaker than *must*. I have already rejected the Kratzerian framework along with other quantificational theories due to various logical problems discussed in chapters 3 and 5, and presumably von Fintel & Iatridou's (2008) analysis must be reworked in order to fit in with my theory as well.

Fortunately this is not too far out of reach: Goble (1996) gives a semantics along these lines, treating *should* and *ought* as alternative-sensitive mid-scalar items just as I have argued for *want*. On the strongest theory of this type that we could adopt, the sole difference between *want* and *should/ought* is that the latter rely on a preference order which is deontic. In addition to making it possible to avoid the problems involving monotonicity and information-sensitivity which we have already discussed in the some detail, this analysis allows us to account for the fact that *should* and *ought* allow degree modification and form comparatives in some cases, which will be discussed in later sections.

The minimal modification of (6.55) for *should* and *ought* is (6.67):

(6.67) *should/ought* ϕ is true iff $\mathbb{E}(\phi) \geq \theta_{should/ought}$, where $\theta_{should/ought}$ is a value significantly greater than $\mathbb{E}(\bigcup \mathbf{ALT}(\phi))$.

This is close to Sloman's (1970) analysis and to the second proposal of Goble (1996) mentioned above, with two subtle differences. Goble defines *ought*(ϕ) as meaning that $\mathbb{E}(\phi) \geq \mathbb{E}(\psi)$ for all $\psi \in \mathbf{ALT}(\phi)$. The first difference is that my proposal requires that there be a **significant** difference between $\mathbb{E}(\phi)$ and θ_{ought}, while Goble does not. Second, my account treats *ought* and *should* as being closely related to relative-standard adjectives, while Goble's treats them essentially as superlatives. Neither of these differences is absolutely crucial for the overall project of this chapter; the initial motivation for (6.67) was just to maintain a maximally simple typology of scalar modals, and it could well turn out that the truth is more complicated. However, it can be shown that (6.67) does a better job of capturing the truth-conditions of sentences with *should* and *ought*.

On the question of a significance parameter, it seems clear that *ought*(ϕ) should not come out true if ϕ is just barely better than indifferent. Modifying an example which Yalcin (2010) uses to make a similar point for *probable*, suppose that it is absolutely morally indifferent whether I go to Paris on vacation or not, except for one factor: I know that my long-lost brother is somewhere in Paris, and if I go there there is a one in 11,000 chance that I will see him (it is, after all, a city of 11 million, and I will see roughly 1,000 of them). If I do see him, it will reunite us and he will be very happy, which is presumably a moral good. Despite all of this, it is hard to see how a reasonable person could accuse me of failing to do what I ought when I choose to go to somewhere else; the chance that I would see my brother was extremely small, and so the moral expectation $\mathbb{E}(\textit{I go to Paris})$ is just barely greater than $\mathbb{E}(\textit{I do not go to Paris})$. This is not enough, it appears, to make it the case that I ought to go to Paris. So we need to add some notion of significant difference to the definition of *ought*, on pain of making $\textit{ought}(\phi)$ come out true too easily.

The second difference between Goble's proposal and mine is that Goble's *ought* is effectively a superlative, so that $\textit{ought}(\phi)$ is true iff $\mathbb{E}(\phi)$ is at least as high as $\mathbb{E}(\psi)$ for each $\psi \in \textbf{ALT}(\phi)$. It follows that $\mathbb{E}(\phi)$ and $\mathbb{E}(\psi)$ must be **equal** if $\textit{ought}(\phi)$ and $\textit{ought}(\psi)$ are both to be true relative to the same alternative set. In contrast, my proposal allows that there may be ϕ and ψ such that $\textit{ought}(\phi)$ and $\textit{ought}(\psi)$ are both true, but one has slightly greater expectation than the other.

It is rather subtle to tease the predictions of the two theories apart, but the following naturally-occurring example suggests that my approach is correct:

(6.68) You should at least consume fatty foods and foods with a high trans fat and bad fat content in moderation or better YET not at all.[18]

The relevant alternative set here seems to be something like $\{$*consume fatty foods (etc.) in moderation, consume them not in moderation, don't consume them at all*$\}$. The first clause clearly entails that $\mathbb{E}(\textit{consume fatty foods in moderation}) \geq \theta_{should}$, but the continuation also indicates that $\mathbb{E}(\textit{don't consume fatty foods at all}) > \mathbb{E}(\textit{consume fatty foods in moderation})$. These conditions cannot be simultaneously true according to Goble's definition of *ought*, but they are compatible on my approach. The only caveat is that, with this alternative set, there is predicted to be an entailment that eating fatty foods not in moderation must be extremely bad; and in fact this inference is clearly intended here.

In either form, the alternative-based analysis is also able to explain the focus-sensitivity of *should* and *ought*, which is also a mystery on quantificational analyses. For example, we can modify the dialogue from the previous section so that it is one friend giving the other advice, and the results are the same:

(6.69) a. **Son's friend**: You should go to the grocery store for your mother.
b. **Son**: I don't want to go to the grocery store. I want to go to a movie.
c. **Son's friend**: Well, you should go to a CLEAN movie, then.

We can analyze this example just like the similar case with *want*. (6.69a) is true if and only if going to the grocery store is deontically above par relative to the alternative set $\{$*Son goes to the grocery store*, *Son does not go to the grocery store*$\}$, i.e., its expectation is significantly greater than the point of indifference $\mathbb{E}(\top)$. It follows, as usual, that the expectation of not going to the grocery must be below the point of indifference.

(6.69c), on the other hand, is true if and only if the expectation of going to a clean movie is significantly greater than the expectation of going to a movie. The latter can be true even if the

18 http://www.lowfatdietplan.org/low-fat-diet-plan/three-foods-you-should-avoid-eating

expectation of going to a movie is less than $\mathbb{E}(\top)$, as (6.69a) entails. As a result (6.69c) is not predicted to entail (6.70) —

(6.70) You should go to a movie.

— as indeed it does not. Here again, focus affects the calculation of θ_{should} in the same way that comparison classes affect relative adjectives, with the result that (6.69a) and (6.69c) are consistent in the present semantics.

Another example showing that focus affects the calculation of the threshold for *should* is the following. Imagine that old Mrs. Marple desperately needs to get out of the house, and she loves going out for lunch. Moreover, nothing would make her happier than to go out for lunch with her feckless daughter Sue. Her neighbor, Mary, is kindly but not at all dear to her. Here (6.71a) seems to be true, but (6.71b) is plausibly false:

(6.71) a. Mary should take Mrs. Marple out for lunch.

b. MARY should take Mrs. Marple out for lunch.

We can account for this example in the same way, though this time the threshold θ_{should} shifts upward rather than downward when there is focus. (6.71a) gets its truth-conditions relative to the alternative set {*Mary takes Mrs. Marple out, Mary doesn't take Mrs. Marple out*}, according to which θ_{should} is significantly greater than the point of indifference $\mathbb{E}(\top)$. This will be true, for example, if it would make Mrs. Marple happy to be taken out to lunch by just about anyone, and so it is better for Mary to do it than not to do it.

(6.71b), on the other hand, requires that θ_{should} be significantly greater than $\mathbb{E}(\bigcup\{$*x takes Mrs. Marple out* $\mid x \in \{$*Mary, Sue, ...*$\}\})$. But this means that the expectation of *Mary takes Mrs. Marple out for lunch* is being compared to $\mathbb{E}($*someone takes Mrs. Marple out*$)$. Since we already know that the latter proposition has expectation significantly greater than the point of indifference, (6.71b) will be true only if it would be especially good for Mrs. Marple's happiness if Mary were the person to take her out, as opposed to someone else. But this is not the case in this story; the only person (as far as we know) who could have this effect is unreliable Sue. As a result (6.71a) is true, but (6.71b) is false.

6.5.4 Deontic Conflicts and Inconsistent Requirements

Another interesting feature of the alternative-based scalar semantics for *ought* and *should* is that it predicts the possibility of genuine conflicts of obligation in some cases, i.e. cases in which *ought* ϕ and *ought* ψ are both true even though ϕ and ψ are inconsistent. I'll briefly describe how these cases go, and also use this issue to segue into discussion of the high scalar modals *must, have to, require*, and *need*, which do not appear to support conflicts in the same way. This fact gives us a first clue about the semantics of high scalar modals.

Recall from chapter 5 that standard deontic logic renders moral conflicts logically impossible, while Kratzer's theory incorrectly assigns falsity to the relevant *ought*-statements. As Goble (1996) points out, though, the scalar semantics with alternatives for *ought* and *should* that we have sketched makes it possible to have genuine moral conflicts in some cases: *ought* ϕ and *ought* ψ can both be true as long as both are substantial (moral) improvements on the status quo.

There are actually two somewhat different types of situation in which real or apparent moral conflicts are predicted possible, on the present theory. First, there are cases in which *ought* ϕ and *ought* ψ are evaluated with respect to different alternative sets. These are not true deontic conflicts, but they arise frequently and look like conflicts on the surface. Second, there are cases in which both *ought*-statements are true with respect to the same alternative set, because two inconsistent alternatives both surpass θ_{ought}. The latter are true deontic conflicts, and can be modeled in exactly the same way that incompatible desires were (cf. (6.58)).

For an example of the first type, suppose that my friend Ted calls me and tells me that he is driving 20 miles per hour on the highway while drunk. Either of the sentences in (6.72) constitutes sound advice on its own:

(6.72) a. Ted, you should speed up, it's dangerous to go so slow on the highway.
b. Ted, you should pull over, drunk driving is illegal and dangerous.

Obviously Ted cannot do both. Nevertheless, both of these utterances can be true if the set of alternatives with respect to which each is evaluated is different: speeding up is contrasted with continuing to drive *slowly*, while pulling over is contrasted with continuing to drive. Apparent (context-shifting) deontic conflicts receive a natural interpretation in the semantics argued for here. I suspect that they could be treated in quantificational theories as well, though, e.g. by an extension of Villalta's (2008) proposal for *want* to mid-scalar deontics.

However, such a treatment would not help with genuine deontic conflicts of the type that has interested moral philosophers, and that we illustrated in chapter 5 (§5.5) with the case of logically independent promises which cannot, due to circumstance, both be fulfilled. Genuine deontic conflicts are predicted to be possible in the present theory due to a special feature of the expectation-based semantics: when there are three or more alternatives, two (or more) can have expectation greater than θ_{should} as long as the other(s) have expectation well below this threshold. One such case is discussed by Harman (1993: 184). I have promised to give C a banana, and I have promised to give D a banana, but then I discover that I have only one banana. According to Harman, "most speakers ... [find] it quite acceptable to say the following:"

(6.73) You ought to give C a banana and you ought to give D a banana, but you can't give both of them a banana, so you have to decide.

Note that this example is very similar to the example of incompatible desires that we saw in (6.58); the analysis that I will offer is essentially the same as the one I gave there.

Specifically, the alternatives in (6.73) seem to be $\{$*Give A a banana, give B a banana, give both a banana, give no one a banana*$\}$. Since *give both a banana* is impossible, it has probability zero and drops out of the calculation of expectation. We are left with three options, and the predicted truth-conditions:

(6.74) a. *You ought to give C a banana* is true iff $\mathbb{E}($*give C a banana*$)$ is significantly greater than $\mathbb{E}(\top)$.
b. *You ought to give D a banana* is true iff $\mathbb{E}($*give D a banana*$)$ is significantly greater than $\mathbb{E}(\top)$.

The present theory predicts that *ought* can be true of two mutually exclusive alternatives can both be true just in case (a) there is no alternative which has much greater expectation, and (b) the expectation of (the disjunction of) the remaining alternatives is less than $\mathbb{E}(\top)$ by at least as much as either of these propositions are greater than $\mathbb{E}(\top)$. In other words, χ must be either quite bad or quite probable in order for this both (6.74a) and (6.74b) to be true simultaneously; in the case at hand, this is plausibly fulfilled, since fulfilling one promise or the other is a great deal better than breaking them both.

In contrast to both standard deontic logic and Kratzer's theory, expectation-based semantics for *ought* and *should* has no difficulty in allowing for that *ought* ϕ and *ought* ψ can be simultaneously true even when ϕ and ψ are incompatible. This is possible in either of two ways: either because the alternatives with respect to which the two statements are evaluated are different, or because the alternative set is rich enough that either would be much better than indifferent. Although more work is needed to check that all convincing examples of deontic conflicts fall into one of these categories, the fact that such conflicts are predicted to be possible seems to be a strong point in favor of the present theory over quantificational approaches (including Kratzer's) which are unable to make sense of the simultaneous truth of two conflicting *ought*-sentences.

6.5.5 High Scalar $\mathcal{D}$-Modals

There is, however, an interesting wrinkle here which leads us to consider the semantics of high scalar $\mathcal{D}$-modals. Lemmon (1962) points out that, as far as intuitions regarding deontic conflicts go, there is an asymmetry between *ought* on the one hand and *have to* and *must* on the other: *ought*(ϕ) and *ought*(ψ) can be simultaneously true for incompatible ϕ and ψ in some cases, but *must*(ϕ) and *must*(ψ) cannot. Harman (1993) also discusses this asymmetry with respect to the cases that we have considered, pointing out that while (6.73) is intuitively acceptable, the minimal modification replacing *ought* with the stronger item *have to* is not.

(6.75) # You have to give C a banana and you have to give D a banana, but you can't give both of them a banana, so you have to decide.

Similarly,

(6.76) # You must give C a banana and you must give D a banana, but you can't give both of them a banana, so you have to decide.

Recall that *must* and *have to* differ from *ought* and *should* in various other ways as well. Among other differences, *must* and *have to* are logically stronger and they do not participate in neg-raising; in addition there is not, as far as I am aware, any evidence for focus-sensitivity in their semantics. I will argue that these facts can be given a unified explanation if we treat the high scalar modals as placing restrictions on the expectation of both their propositional argument and its negation.

6.5.5.1 *Must* and *Have To*: The Basic Account

Recall the type of example which was used to show that *have to* is logically stronger than *should* and *ought to*:

(6.77) You ought to wash the dishes; in fact, you have to.

This example brings out the intuitive difference between *ought to* and *have to* nicely. The first clause, with *ought*, indicates that washing the dishes is a better option (morally or otherwise) than not washing the dishes. The information added by the second clause, essentially, is that you don't have a choice. This is the same gloss that Sloman (1970) and von Fintel & Iatridou (2008) give for these items: *ought* "picks out the best means without excluding the possibility of others", while *have to* "implies that no other means exists" (Sloman 1970: 390-1).

Another illustration of this difference is the contrast in (6.78). (6.78a) (from a set of instructions for testing water quality) is quite reasonable, but (6.78b) is not a coherent text.

(6.78) a. The solubility of the substance should be at least 2 mg/l, though in principle less soluble compounds could be tested ...[19]

b. # The solubility of the substance must be at least 2 mg/l, though in principle less soluble compounds could be tested.

This example illustrates again the essential difference between mid- and high-scalar modals: the former leave open the possibility of other alternatives, while the latter do not.

We can make this a bit more formal along the following lines. As I argued above, with mid-scalar $\mathcal{D}$-modals $\mathcal{D}(\phi)$ is true, roughly, when ϕ is a good option. The high scalar $\mathcal{D}$-modals are different in several ways. First, they have a higher threshold than *should* and *ought*, essentially as high degree adjectives (*huge, ecstatic*) relate to their relative-standard counterparts (*big, happy*). The second difference a direct consequence of the first: unlike, say, the domain of heights (where one individuals' height is logically independent of another's), there is a logical relationship between the expectation of mutually exclusive propositions. If one proposition ϕ has very high expectation relative to $\mathbb{E}(\top)$, then others must compensate by having an expectation which is lower than $\mathbb{E}(\top)$ to a degree proportional to the odds of ϕ (cf. §6.4.1.3).

The simplest way to account for high scalar modals would be to require that the threshold must be, not just significantly higher than average, but *extremely* high: $must(\phi)$ means that $\mathbb{E}(\phi)$ is close to the upper bound of the expectations of any proposition in the model which has significant probability. The result that $\mathbb{E}(\neg\phi)$ must be low follows as well; this is clear if we consider the unique admissible $\mu^P_{\mathcal{D}}$ such that $\mu^P_{\mathcal{D}}(\top) = 0$, so that better-than-indifferent propositions get positive expectation and worse-than-indifferent propositions get negative expectation. $\mathbb{E}(\top)$ is equal to $\mathbb{E}(\phi \vee \neg\phi) = \mathbb{E}(\phi) \times prob(\phi) + \mathbb{E}(\neg\phi) \times prob(\neg\phi)$, and so if (for the $\mu^P_{\mathcal{D}}$ in question) $\mathbb{E}(\phi) = +\eta$ for some large η, then $\mathbb{E}(\neg\phi) = -\eta \times \frac{prob(\phi)}{prob(\neg\phi)}$. In this way we get the desired result that $must(\phi)$ entails both that ϕ is very good and $\neg\phi$ is very bad (or moderately bad and very likely).

This will not quite work, however. The problem is that there can be incompatible ψ and χ such that both are extremely good, as long as these do not exhaust the possibilities. In this circumstance, it seems, $must(\psi)$ and $must(\chi)$ would both seem to be false, although $must(\psi \vee \chi)$ may well be true as long as there are no other comparably good options. For example, suppose that there are compelling reasons for spending my vacation in either of two incompatible ways: I could go see my parents who I have not seen in some time, or I could visit my ailing grandparents. Even though the sentences in (6.79) are plausibly both true –

19 http://fedbbs.access.gpo.gov/library/epa_835/835-3160.htm

(6.79) a. I ought to go to my parents'.

b. I ought to go to my grandparents'.

– neither of the sentences in (6.80) is true here:

(6.80) a. I must go to my parents'.

b. I must go to my grandparents'.

But since the expectation of each of these propositions could (let's say) be near the maximum of any proposition with significant probability in the model, simply requiring an extremely high expectation is not enough to derive the desired result; in that case both of the sentences in (6.80) would be true.

What we need is a semantics where $must(\phi)$ and $must(\psi)$ cannot both be true if ϕ and ψ are incompatible and both are sufficiently salient and/or plausible options. One possibility is:

(6.81) $must(\phi)$ is true iff

a. $\mathbb{E}(\phi) \geq \theta_{must}$, where θ_{must} is some very high threshold;

b. $prob(\phi)$ is significantly greater than 0; and

c. For all $\psi \subseteq \neg\phi$: if $prob(\psi)$ is significantly greater than 0, then $\mathbb{E}(\psi) < \mathbb{E}(\top)$.

The idea here is to require that ϕ must be, not only an extremely good option, but the **only** option with significant probability which is better than indifferent.

It can be shown that, on this semantics, the inference in (6.82) is valid:

(6.82) a. $must(\phi)$

b. $must(\psi)$

c. $\phi \wedge \psi$ is possible ($prob(\phi \wedge \psi) > 0$).[20]

So the truth-conditions in (6.81) account for Harman's (1993) observation that $must(\phi)$ and $must(\psi)$ cannot both be true when $\phi \wedge \psi$ is not possible.

6.5.5.2 Ross' Paradox Again

The proposed denotation for *must* also allows us to tie up a loose end from the discussion of Ross' Paradox in §6.4.1.1. Recall that Ross' Paradox is that, intuitively, $must(\phi)$ does not entail $must(\phi \vee \psi)$, but quantificational theories predict that it does. The non-monotonicity of expectation allowed us to resolve the problem, since $\mathbb{E}(\phi) \geq \theta_{must}$ does not entail that $\mathbb{E}(\phi \vee \psi) \geq \theta_{must}$.

20 **Proof of (6.82)**: Suppose (6.82c) is false. Then $prob(\phi \wedge \psi) = 0$, and so $prob(\psi) = prob(\neg\phi \wedge \psi)$. Either $prob(\neg\phi \wedge \psi)$ is significantly greater than zero, or it is not.

If it is, then $\mathbb{E}(\neg\phi \wedge \psi)$, being the expectation of a proposition inconsistent with ϕ, is less than $\mathbb{E}(\top)$ (by (6.81c)). Since $prob(\phi \wedge \psi) = 0$, $\mathbb{E}(\psi)$ — the probability-weighted average of $\mathbb{E}(\phi \wedge \psi)$ and $\mathbb{E}(\neg\phi \wedge \psi)$ — is equal to $\mathbb{E}(\neg\phi \wedge \psi)$. The latter quantity is less than $\mathbb{E}(\top)$, and so $\mathbb{E}(\psi) < \mathbb{E}(\top)$. But since $must(\psi)$ is true (6.82b), $\mathbb{E}(\psi) \geq \theta_{must}$, which is greater than $\mathbb{E}(\top)$; so we have a contradiction.

On the other hand, if $prob(\neg\phi \wedge \psi)$ is not significantly greater than zero, then $prob(\psi)$ is not significantly greater than zero since (we are assuming) $prob(\phi \wedge \psi) = 0$. But $must(\psi)$ is true by (6.82b), which entails that $prob(\psi)$ is significantly greater than zero (by (6.81b)); so we have a contradiction.

Since these two options exhaust the possibilities, we conclude that (6.82c) is true.

However, if we add the premise that $\mathbb{E}(\psi) > \mathbb{E}(\phi)$, it does follow that $\mathbb{E}(\phi \vee \psi) \geq \theta_{must}$, and so that $must(\phi \vee \psi)$ is true even if ϕ and ψ are incompatible. This seems strange, though: intuitively someone could not coherently tell me that I must spend my vacation with my parents, and that it would be even better to spend it with my grandparents. This suggests that there is a logical conflict between the premises (6.83b) and (6.83c):

(6.83) a. $must(\phi)$
b. $\phi \cap \psi = \varnothing$
c. ψ *is at least as good as* ϕ
d. $\therefore \ \bot$

The proposal in (6.81) explains this conflict, with one caveat: in order to make the argument valid we have to add premise (6.84d):

(6.84) a. $must(\phi)$
b. $\phi \cap \psi = \varnothing$
c. ψ *is at least as good as* ϕ
d. ψ is a reasonable possibility
e. $\therefore \ \bot$

With this addition, the argument in (6.84) is valid according to the semantics for *must* in (6.81). If ψ is a reasonable possibility and at least as good as ϕ, then $must(\phi)$ will be false. At best, in this situation, we can say that ϕ and ψ both ought to be the case, as in (6.79).

In order to be sure that this solution works, we must check that it is reasonable to assume that the additional premise (6.84d) is implicit in the relevant scenarios. For example, in the imagined conversation above, the fact that the possibility of my spending the vacation with my grandparents has even been mentioned strongly suggests that my interlocutor thinks that it is a reasonable possibility. In general, any normal situation in which $must(\phi)$ is true will also be a situation where $\mathbb{E}(\psi) > \mathbb{E}(\phi)$ is false for any incompatible ψ which is not rather far-fetched. As a result the conclusion $must(\phi \vee \psi)$, though not necessarily false, is misleading and inappropriate in these cases.

6.5.5.3 *Require* and *Need*

Based on tests from neg-raising and entailments, we have classified the modal verbs *require* and *need* as high-scalar modals. For instance:

(6.85) a. Mary shouldn't leave. $\rightsquigarrow$ Mary should stay.
b. Mary doesn't want to leave. $\rightsquigarrow$ Mary wants to stay.

(6.86) a. Mary doesn't need to leave. $\not\rightsquigarrow$ Mary needs to stay.
b. Mary isn't required to leave. $\not\rightsquigarrow$ Mary is required to stay.

Similarly, Horn (1989) points out that *require* and *need* are related to *want* essentially as *must* and *have to* are to *ought*.

(6.87) a. I want you to leave; in fact, I need/require you to.

b. # I need/require you to leave; in fact, I want you to.

There are some subtle difference between the two verbs involving raising vs. control structures; however, to a first approximation, *need* and *require* seem to be the bouletic counterparts of *must* and *have to*.

There is one subtle difference, however: while *require* patterns with *must* and *have to* in examples like (6.76), incompatible needs appear to be possible in some cases.

(6.88) a. # You are required to spend your vacation with your parents, and you are required to spend it with your grandparents. Unfortunately you can't do both.

b. You need to spend your vacation with your parents, and you need to spend it with your grandparents. Unfortunately you can't do both.

Someone who gives you a command using (6.88a) is clearly confused: it doesn't make sense to require something of a person if you know that the requirement cannot be fulfilled. (6.88b), in contrast, is a completely coherent way to state that someone is in a serious bind.[21] I take this contrast to indicate that

- ϕ *is required* and ψ *is required* together entail $\phi \wedge \psi$ *is possible*;
- *x needs* ϕ and *x needs* ψ together entail that ϕ *is possible* and that ψ *is possible*, but not that $\phi \wedge \psi$ is possible.

We can capture this contrast is by treating *require* as logically stronger than *need*: the former has the same basic semantics as *must* and *have to* given above, but *need* does not place the extra condition on its negation, so that it has a semantics is more closely related to that of ordinary high degree adjectives.[22]

(6.89) ϕ *is required* is true iff

a. $\mathbb{E}(\phi) \geq \theta_{required}$, where $\theta_{required}$ is some very high threshold; and

b. For all $\psi \subseteq \neg\phi$: if $prob(\psi)$ is significantly greater than 0, then $\mathbb{E}(\psi) < \mathbb{E}(\top)$.

(6.90) *x needs* ϕ is true iff $\mathbb{E}(\phi) \geq \theta_{need}$, where θ_{need} is some very high threshold.

---

21 A naturally-occurring example which makes this point is:

> Just like [Sharon] Angle's looming teabagger rally speech, the repubs in general are in trouble because they have two very different audiences they very much need to appeal to: the wild eyed teabaggers and the moderate independents.
>
> But they can't appeal to both — the spiels are mutually exclusive. And when immigration comes back to the fore, the gulf will just grow even wider...

(From discussion on washingtonpost.com, June 18 2010, http://voices.washingtonpost.com/plum-line/2010/06/new_dnc_ad_calls_on_republican.html)

22 I assume that the relevant comparison here is with high degree adjectives, rather than maximum-standard adjectives, simply because the scale of expectation does not have a maximum element, and imposing one is undesirable for various logical reasons.

Very roughly: ϕ is *needed* if it would be really great if ϕ were the case; it is *required* in addition if there is no other way to get an even middling result. This proposal accounts for the subtle difference in *require* and *need* shown in (6.88): two incompatible propositions can be needed, but they cannot both be required, since (as we saw in (6.82)) this would entail that their conjunction has non-zero probability.[23]

6.5.6 Weak Scalar $\mathcal{D}$-Modals

Finally we come to the weak scalar $\mathcal{D}$-modals *may, allowed*, and *permissible*. These have not played a major role in the puzzles addressed in this and the previous chapter, and I will be brief and speculative about their lexical semantics. It is usual in modal logic to define the strong modals *necessarily, must, obligatory*, etc. and the weak modals *may, allowed, possible*, etc. in terms of each other:

(6.91) a. ϕ is necessary (etc.) if and only if $\neg\phi$ is not possible (etc.).
b. ϕ is possible (etc.) if and only if $\neg\phi$ is not necessary (etc.).

Within a scalar theory built on probability-weighted preference, we want to retain the intuition of a strong connection between possibility and necessity while also treating these items as establishing a relatively low threshold value $\theta_{\mathcal{D}}$. One option might be to set $\theta_{\mathcal{D}} = \mathbb{E}(\top)$. However, this would seem to give the wrong result in some cases: for example, since $\mathbb{E}(\phi) \geq \mathbb{E}(\top)$ if and only if $\mathbb{E}(\phi) \geq \mathbb{E}(\neg\phi)$, this would predict that (6.92a) and (6.92b) should both be contradictions.

(6.92) a. You may eat that candy, but it's better if you don't.
b. You may eat that candy, but you shouldn't.

Both of these sentences are perfectly coherent, though; they express permissions, albeit begrudging ones. So it seems that $may(\phi)$ can be true even if $\mathbb{E}(\neg\phi) > \mathbb{E}(\phi)$, as both of these sentences entail.

The best option that I can see is to adopt a close variant of (6.91b) as our definition of the low scalar $\mathcal{D}$-modals:

(6.93) a. $may(\phi)$ is true iff $must(\neg\phi)$ is false.
b. $allowed(\phi)$ is true iff $required(\neg\phi)$ is false.

Now according to (6.81), $\neg\phi$ *is required* is false if and only if either of two conditions hold: either

(6.94) a. $\mathbb{E}(\neg\phi) < \theta_{required}$; or
b. $\mathbb{E}(\psi) \geq \mathbb{E}(\top)$ for some reasonably likely subset ψ of ϕ.

May/allowed ϕ will be true if either of these conditions holds: either if $\neg\phi$ is not extremely good, or if there is some subset of ϕ which is as good as indifferent and has non-zero probability.

Now, if *may* ϕ is true because condition (6.94a) holds, it may still be the case that ϕ is worse than indifferent; all that is required is that it is not extremely bad. This condition is even compatible

23 Note that incompatible ϕ and ψ can be coherently required as long as the requirements are imposed by different sources. In such cases we are presumably dealing with two different preference orders: ϕ is required (by *x*) and ψ is required (by *y*), but it is not possible to do both. Thanks to Chris Barker for pointing this out.

with the truth of *should*$(\neg\phi)$, as in (6.92b), as long as $\mathbb{E}(\neg\phi)$ falls somewhere between θ_{should} and θ_{must}.

This proposal is still tentative, but it gives a sense of how a reasonable scalar analysis of weak modals would work and how it captures what is correct in the traditional semantics. More work is needed to ascertain, for example, whether this analysis makes new predictions with respect to the classic puzzle of Free Choice Permission; the non-monotonicity of *may* in this theory may well lead to different predictions. For now, however, I will leave this issue to the side.

6.6 Chapter 6 Summary and Conclusions

We have covered a lot of ground in this chapter; let me try to summarize briefly the high points. In the first section I argued that the fatal flaw of quantificational theories of deontic and bouletic modals, including Kratzer's, is that they attempt to assign a degree of obligation or desirability to a proposition while only taking into account the relative positions of the highest-ranked worlds in the propositions. I then suggested an alternative which makes better use of the information in the preference order by treating the desirability of a proposition as the weighted average of the desirabilities of the worlds in it ("probability-weighted preference" or just "expectation"). This move required almost no new semantic machinery beyond the preference orders that are used in standard deontic logic and the probability measures that, as I argued in chapter 3, are needed to account for epistemic modals.

Nevertheless, the logical properties of obligation and desire are strikingly different in this theory. In §6.4 I showed that, because probability-weighted preference is non-monotonic, several classic puzzles involving compound propositions in interaction with $\mathcal{D}$-modals can be resolved in a straightforward way. In addition, several puzzles from chapter 5 which demonstrated that quantificational theories of modality are not sufficiently sensitive to probabilistic information were shown to have a simple resolution. In particular, the semantics given here makes it possible to treat the Miner's Paradox in the same terms in which the medicine and insurance cases were dealt with.

Finally, the proposed semantics was shown to offer a good account of gradability and comparison of $\mathcal{D}$-modals, which are problematic in various ways for the standard theory due to Kratzer (1991). I also argued for a three-way typology of $\mathcal{D}$-modals related to the various types of gradable adjectives, and then showed that this theory yields new predictions about moral conflicts, the interaction of mid-scalar modals with focus, and a more natural semantic account of "weak necessity" modals, among other points.

CHAPTER 7

Overview and Future Directions

7.1 Summary of Proposal and Results

This dissertation has proposed a substantially new approach to the semantics of modality according to which modals are **measure functions**: expressions which map propositions to scales and compare them to a threshold value. Nearly all approaches to modality in formal semantics have treated modals instead as quantifiers over possible worlds, including the theory of Kratzer (1981, 1991) which has dominated discussion in linguistic semantics. Nevertheless, I argued that epistemic and deontic modals are much more similar to gradable adjectives than they are to quantifiers, both grammatically and logically. I also showed that a number of classic puzzles involving modal semantics, as well as several that are new here, create problems for a quantificational analysis of modals but have a natural resolution in the scalar semantics given here.

In chapter 3, I showed that Kratzer's theory has a number of empirical problems involving epistemic modals: it makes clearly incorrect predictions about the interaction of epistemic comparatives and equatives with disjunction, fails to give consistent truth-conditions to epistemic modals with degree modifiers, and forcibly declares too many epistemic comparatives undefined. By analyzing the degree modifiers which adjectival epistemic modals accept, I demonstrated that the scales underlying these expressions are equivalent to finitely additive probability measures. With this conclusion in hand, it turned out that the modal auxiliaries could not be treated as quantifiers either, on pain of making absurd predictions about the logical relationship between the adjectival and auxiliary modals; I also proposed an information-theoretic account of question-embedding *certain*. Chapter 4 discussed experimental evidence showing that the relative adjectives *likely* and *probable* are sensitive to contextual alternatives, showing that — contrary to the received interpretation of these results — this is not evidence of irrationality, but the expected semantic behavior of these items and perfectly consistent with the assumption that subjects are making coherent probability judgments.

Chapter 5 on deontic modals and desire verbs posed a number of problems for quantificational theories, including Kratzer's. I argued on a variety of grounds that these items are non-monotonic, sensitive to probabilistic information, come in a greater variety of grades than quantificational semantics can account for, do not behave as expected in comparatives, and allow for robust deontic conflicts in some cases. Although quantificational theories have piecemeal accounts of a few of the issues discussed, no unified theory has ever been proposed, and in many cases the data are utterly mysterious for these theories.

In chapter 6, I gave a simple scalar semantics for deontic modals and desire verbs and showed that it is able to account for the puzzles in chapter 5. The key is to treat these expressions as functions which map propositions to a scale of probability-weighted preference, a construct which is familiar from decision theory and economics. In addition to avoiding the logical problems for quantificational semantics and incorporating probability in the appropriate way, this approach accounts naturally for the three-way distinction of modal strength identified by Horn (1989) and von Fintel & Iatridou (2008), accounts for the facts involving gradability, comparison, and degree modification, and leaves room for genuine deontic conflicts.

Overall, the scalar account developed here is an attractive alternative to standard quantificational theories of modal semantics, and improves on the predictions of these theories in a variety of ways.

7.2 Is a Unified Modal Semantics Possible?

I have argued that Kratzer's semantics for modality is incorrect in some fairly deep ways, along with other proposals which assume that modals are quantifiers over possible worlds. However, Kratzer's theory has a desirable feature which my theory cannot obviously match, since it associates epistemic and deontic/bouletic modals with domains that have quite different structures: a firm commitment to giving a unified semantics for modals such as *must* and *can*, rather than treating, for instance, epistemic and deontic *must* as being distinct but homophonous lexical items. As Kratzer (1981: 340) argues, it is desirable to have an analysis of these items in which "there is something in the meaning they have ... which stays invariable". As Kratzer presents it, this is based on a theoretical intuition. I share the belief that a theory with this feature is desirable, but to my mind this is a methodological desideratum following from Grice's (1989) Modified Occam's Razor: "Senses are not to be multiplied beyond necessity". In general, it's better to avoid positing lexical ambiguity if you can avoid it.

It seems clear, though, that if the theory which best accounts for the data does not validate Kratzer's theoretical intuition and Grice's methodological principle, we have no choice but to abandon the intuition and sideline the principle. As Swanson (2008: 1204-5) points out,

> The substantial differences between epistemic and so-called 'root' (non-epistemic) modals make it unclear precisely what one is aiming for in giving a 'relatively unified' semantics of different modal expressions. At the same time, there are enough similarities between the flavors of modality ... that it is fruitful to look for interesting phenomena involving one flavor of modality where we've already found such phenomena involving another. A plausible explanation of these similarities is that at least some of them are due to shared history. But it's consistent with this that some ways of expressing modal information have come to have quite different features, and that they now demand quite different semantics.

While I don't have a firm answer to this question, I want to bring up two further relevant considerations which seem to go in opposite directions. First, most of the modal expressions that we have considered in this dissertation are *not* as promiscuous in their modal flavor as *must*, *can*, *should*, and *ought*. For example, *want* seems to be restricted to bouletic and teleological uses; *probable, likely*, and *certain* have only an epistemic meaning; *good, permissible, obligatory*, etc. are primarily deontic; and so on. The auxiliaries are not typical representatives of the domain of modal expressions, as Kratzer's discussion might lead us to think; they are really quite extreme in the number of distinct uses that they have. Even if we were to abandon the project of a unified modal semantics, it is just not clear how much theoretical duplication this would make for: in many cases, it seems, the range of uses of modal expressions is so limited that there would be very little effect.

Still, I'm not completely pessimistic about the prospects for modal unification. We may still be able to unify the meanings of these expressions in terms not too different from Kratzer's. In her theory, *must* and *can* have stable meanings in the face of changing conversational backgrounds by maintaining a stable quantificational force: *must* universally quantifies over the set of $\succcurlyeq_{\mathbf{g}(w)}$-maximal worlds, however context and semantics work under the hood to determine what set this is. A scalar semantics can similarly maintain a stable core meaning for *must* and *can* in the face of changing orderings over worlds and propositions by keeping the threshold value associated with *must* stable. For example, the proposals in chapters 3 and 6 both require that the propositional argument of *must* exceed a very high threshold. The details of these proposals differed in enough ways, as do the underlying structures of the scales in question, that I can't offer a rule for moving from a high value on an upper-closed scale like probability to a high value on an open scale like expectation. Still, it's fairly clear in a pre-theoretical sense that there is something stable here. Empirical work on the semantics of scalar adjectives might also turn out to be relevant here: if there are high-degree adjectives that can be associated with either open or closed scales, as I suggested for *obligatory* in chapter 6, we may find that maximum-standard and high-degree adjectives reflect a single underlying category which surfaces differently depending on the boundedness of the scale.

The semantics that I proposed for *should* and *ought* in their epistemic and deontic uses are even more obviously related than the semantics for *must*: in each case, the propositional argument is compared to an average which is calculated across a set of contextual alternatives. In sum, to what extent a unified semantics for different modal flavors is an obligatory feature of a good theory of modality is an open question. However, scalar semantics does hold out the possibility of a relatively unified semantics for different modal flavors. Further development will be needed both in scalar semantics for modality and in the semantics of gradability more generally before we can be sure how this issue will play out.

7.3 Future Directions

There are many issues in the semantics of modality, gradability, and related issues that I have dealt with quickly or ignored altogether in this dissertation. In some cases, this is because the issue was felt to be not crucial given the focus on the structure of modality. For instance, I have sidelined many details of the precise compositional implementation of the account. In other cases, issues were avoided because they would have led to a much longer dissertation. For example, the scalar semantics proposed here also differs from quantificational semantics in many cases in its predictions about the interactions between modals and other operators. As I have hinted, initial indications are that the predictions of the scalar account are markedly better than those of quantificational theories in at least some cases: see Lassiter (2010b) for a brief discussion relating to weak islands and Heim's (2001) puzzle about degree operator scope, and Lassiter (2011a) for a discussion of the puzzle about minimum and maximum requirements due to Nouwen (2010b,a). In both of these cases standard accounts of modal semantics make markedly incorrect predictions about the interaction with other degree expressions, due specifically to the fact that modals are treated as quantifiers over possible worlds. The scalar account proposed here does better in these cases and quite possibly in others as well.

Future work in this vein should also expand the range of modal flavors covered. In the case of teleological modals, I expect that this will be straightforward: the expectation-based semantics for deontic and bouletic modals proposed in chapter 6 will probably extend directly to teleological modality as well. (We may well want to consider whether deontic and teleological modals are really different in any deep sense.) I am much less sure what to say about circumstantial modals such as *able* and *can* (in the relevant use). These modals are peculiar in a number of ways. For one, they seem to lack duals; in addition their intuitive gradability is fairly limited. However, there are some convincing examples of *more able*: for instance, the stated goal of the UK Department of Education and Employment's publication *Mathematical challenges for able pupils* is "to help primary teachers cater for pupils who are more able in mathematics".[1] In addition, the closely related item *capable* is quite readily graded. However, I have not looked into circumstantial modals in enough detail to offer a concrete theory of their behavior as scalar items at the moment.

Another important class of expressions frequently discussed in connection with modal semantics are counterfactual conditionals. Although I have said nothing about counterfactuals here, there is reason to think they these too can be given a semantics closely related to the proposals made here. For one, the most influential semantics for counterfactuals to date, due to Lewis (1973), is very closely related to — and served as a direct inspiration for — Kratzer's semantics for modality. This connection hints that some of the methods employed here may also be fruitfully extended to the study of counterfactuals. More importantly, though, there is a well-developed scalar semantics for counterfactuals already on offer: the probabilistic account in *Causality* (Pearl 2000), whose influence has been enormous in computer science, psychology, and some areas of philosophy, but curiously absent in formal semantics.[2] Perhaps this dissertation may inspire semanticists to look seriously at Pearl's work in search of a good semantics for counterfactuals.

Finally, it will be important to consider how the semantics for $\mathcal{D}$-modals developed in chapter 6 relates to the semantics of imperatives. It seems plausible that expectation will play a role in this area as well, but the details remain to be seen. One useful point of comparison is Starr (2010), who develops a semantics for imperatives based on preference orders. This proposal is different in the details but similar in spirit to the current approach.

7.4 Interdisciplinary Connections

Many of the core formal notions developed in this dissertation fall beyond the standard toolkit of formal semantics. Probabilistic analyses have played an important role in philosophy and are increasingly dominant in psychology and artificial intelligence, for example. However, formal semantics has for the most part remained anchored firmly to a symbolic tradition which is often thought (wrongly) to be in competition with this approach. Expected utility and related ideas were among the most influential conceptual developments of the 20^{th} century for economics, psychology, and computer science; but most formal linguists are only dimly aware of this framework, with the happy exception of those involved in the recent trend of game-theoretic pragmatics. The unifying

1 http://www.bgfl.org/bgfl/custom/files_uploaded/uploaded_resources/12212/mathspuzzlesall.pdf

2 An important exception is Schulz (2007, 2011) — who, however, abandons the crucial probabilistic component of Pearl's theory.

framework of Measurement Theory, which makes it possible to connect previously unrelated areas of formal semantics, has exerted only limited influence in the formal semantics of gradation and has not previously been connected with the semantics of modality. Measurement Theory, too, seems to have suffered from a lack of cross-disciplinary communication.

The connections between formal semantics and these related academic fields that have been drawn here are exciting, for two reasons. First, I think it will be fruitful for formal semantics to expand the traditional toolkit by looking into the large and sophisticated literature on the representation of uncertainty, uncertain reasoning and decision-making, learning, and beyond. Many different approaches have been studied in great detail and from a variety of perspectives by psychologists, economists, philosophers, logicians, mathematicians, computer scientists, and statisticians, among others, and we have only scratched the surface of this rich literature here. I believe that there is a great deal to be gained in formal semantics by careful attention to the formal tools developed in related disciplines.

Second, greater engagement between the formal study of natural language meaning will likely lead to unexpected theoretical engagement with other fields which already make use of these tools. An especially promising connection, in my mind, is with the psychological study of reasoning, decision-making, learning, and related areas. These fields make considerable use of formal tools closely related to those studied in this dissertation, and — as the small example of such engagement studied in chapter 4 suggests — the empirical and theoretical insights that a careful analysis of linguistic meaning can offer for these fields is considerable. In addition to theoretical contributions to psychology from linguistics, there are new and exciting theoretical movements in psychology which linguists would do well to pay attention to. See Chater et al. (2006); Griffiths et al. (2008) for overviews and references to the rich and growing literature developing empirical insights and computational models which indicate that probability is of fundamental importance in human psychology. These trends lend indirect but important support to the style of semantics for gradability and modality developed here, and the scalar approach will no doubt support further engagement between psychology and formal semantics.

References

Bach, Emmon. 1986. The algebra of events. *Linguistics and Philosophy* 9(1). 5–16.
Bale, Alan C. 2006. *The universal scale and the semantics of comparison*: McGill University dissertation.
Bale, Alan C. 2008. A universal scale of comparison. *Linguistics and Philosophy* 31(1). 1–55.
Bale, Alan C. 2011. Scales and comparison classes. *Natural Language Semantics (*to appear*)* .
Barker, Chris. 2002. The dynamics of vagueness. *Linguistics and Philosophy* 25(1). 1–36.
Barner, David & Jesse Snedeker. 2008. Compositionality and statistics in adjective acquisition: 4-year-olds interpret *tall* and *short* based on the size distributions of novel noun referents. *Child development* 79(3). 594–608.
Bartsch, Renate & Theo Vennemann. 1973. *Semantic structures: a study in the relation between semantics and syntax*. Athenäum.
Bastiaanse, Harald. 2011. The rationality of round interpretation. In U. Sauerland R. Nouwen, R. van Rooij & H.-C. Schmitz (eds.), *Vagueness in communication*, 37–50. Springer.
Beaver, David & Brady Clark. 2003. *Always* and *only*: Why not all focus-sensitive operators are alike. *Natural Language Semantics* 11(4). 323–362.
Beaver, David & Brady Clark. 2008. *Sense and sensitivity: How focus determines meaning*. Wiley-Blackwell.
Bennett, Jonathan F. 2003. *A Philosophical Guide to Conditionals*. Oxford University Press.
Bierwisch, Manfred. 1989. The semantics of gradation. In M. Bierwisch & E. Lang (eds.), *Dimensional adjectives: Grammatical structure and conceptual interpretation*, 71–261. Springer-Verlag.
Bresnan, Joan. 1973. Syntax of the comparative clause construction in English. *Linguistic inquiry* 4(3). 275–343.
Broome, John. 1995. *Weighing Goods: Equality, Uncertainty and Time*. Wiley-Blackwell.
Broome, John. 1999. *Ethics out of Economics*. Cambridge University Press.
Cariani, Fabirzio. 2011. *Ought* and resolution semantics. *Noûs* (forthcoming) .
Carroll, Lewis. 1866. *Alice's Adventures in Wonderland*. Macmillan.
Charlow, Nate. 2011. What We Know and What to Do. *Synthese* doi:10.1007/s11229-011-9974-9.
Chater, Nick, Joshua B. Tenenbaum & Alan Yuille. 2006. Probabilistic models of cognition: Conceptual foundations. *Trends in Cognitive Sciences* 10(7). 287–291. doi:10.1016/j.tics.2006.05.007.
Chisholm, Roderick M. 1964. The ethics of requirement. *American Philosophical Quarterly* 1(2). 147–153.
Chomsky, Noam. 1977. On *wh*-movement. In Peter Culicover, Thomas Wasow & A. Akmajian (eds.), *Formal syntax*, vol. 132, Academic Press.
Copley, Bridget. 2006. What should *should* mean? Ms., CNRS/University of Paris 8 .
Costa, Horacio Arlo & William Taysom. 2005. Contextual Modals. In David Leake A. Dey, Biocho Kokinov & Roy Turner (eds.), *Modeling and using context*, 175–198. Springer.
Cover, Thomas M. & Joy A. Thomas. 1991. *Elements of information theory*. Wiley.
Cresswell, M.J. 1976. The semantics of degree. In B. Partee (ed.), *Montague grammar*, 261–292. Academic Press.
Davis, Christopher, Christopher Potts & Margaret Speas. 2007. The pragmatic values of evidential

sentences. In *Proceedings of salt*, vol. 17, .
Driver, Julia. 1997. The Ethics of Intervention. *Philosophy and Phenomenological Research* 57(4). 851–870.
Edwards, Ward. 1968. Conservatism in human information processing. In B. Kleinmuntz (ed.), *Formal representation of human judgment*, 17–52. New York: John Wiley and Sons.
Egré, Paul & Mikaël Cozic. 2011. If-clauses and probability operators. *Topoi* 1–13. http://www.springerlink.com/content/t61r3u06227811q8/.
Ellis, Brian. 1966. *Basic Concepts of Measurement*. Cambridge University Press.
Evans, Jonathan St. B.T. 2008. Dual-processing accounts of reasoning, judgment, and social cognition. *Psychology* 59(1). 255.
Fagin, Ronald, Joseph Y. Halpern, Yoram Moses & Moshe Y. Vardi. 2003. *Reasoning About Knowledge*. MIT Press.
Fara, Delia Graff. 2000. Shifting sands: An interest-relative theory of vagueness. *Philosophical Topics* 20. 45–81.
Fine, Terrence L. 1973. *Theories of Probability: An Examination of Foundations*. Academic Press.
von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97.
von Fintel, Kai & Anthony Gillies. 2008. CIA leaks. *The Philosophical Review* 117(1). 77.
von Fintel, Kai & Anthony Gillies. 2010. Must... stay... strong! *Natural Language Semantics* 18(4). 351–383.
von Fintel, Kai & Sabine Iatridou. 2008. How to say *ought* in foreign: The composition of weak necessity modals. In Jacqueline Guéron (ed.), *Time and modality*, 115–141. Springer.
Fishburn, Peter C. 1986. The axioms of subjective probability. *Statistical Science* 1(3). 335–345.
Fox, Danny & Martin Hackl. 2006. The universal density of measurement. *Linguistics and Philosophy* 29(5). 537–586.
van Fraassen, Bas C. 1966. Singular terms, truth-value gaps, and free logic. *Journal of Philosophy* 63(17). 481–495.
van Fraassen, Bas C. 1968. Presupposition, implication, and self-reference. *Journal of Philosophy* 65(5). 136–152.
van Fraassen, Bas C. 1973. Values and the heart's command. *Journal of Philosophy* 70(1). 5–19.
Frankish, Keith. 2010. Dual-process and dual-system theories of reasoning. *Philosophy Compass* .
Frazee, Joey & David Beaver. 2010. Vagueness is rational under uncertainty. *Proceedings of the 17th Amsterdam Colloquium* .
Geurts, Bart. 2005. Entertaining alternatives: disjunctions as modals. *Natural Language Semantics* 13(4). 383–410.
Gibbard, Allan & William L. Harper. 1978. Counterfactuals and two kinds of expected utility. In Harper, Stalnaker & Pearce (eds.), *Ifs: Conditionals, belief, decision, chance, and time*, D. Reidel.
Gigerenzer, Gerd. 1991. How to make cognitive illusions disappear: Beyond "heuristics and biases". *European Review of Social Psychology* 2(1). 83–115. doi:10.1080/14792779143000033.
Gigerenzer, Gerd. 2000. *Adaptive thinking: Rationality in the real world*. Oxford University Press, USA.

Goble, Lou. 1996. Utilitarian deontic logic. *Philosophical Studies* 82(3). 317–357.

Goldblatt, Robert. 1987. *Logics of Time and Computation*. CSLI Publications.

Goldblatt, Robert. 2003. Mathematical modal logic: a view of its evolution. *Journal of Applied Logic* 1(5-6). 309–392.

Gould, Stephen Jay. 1992. *Bully for Brontosaurus: Reflections in Natural History*. WW Norton & Company.

Grice, H. Paul. 1989. *Studies in the Way of Words*. Harvard University Press.

Griffiths, Thomas L., Charles Kemp & Joshua B. Tenenbaum. 2008. Bayesian models of cognition. In R. Sun (ed.), *Cambridge handbook of computational psychology*, 59–100. Cambridge University Press.

Groenendijk, Jeroen & Martin Stokhof. 1984. *Studies in the Semantics of Questions and the Pragmatics of Answers*: University of Amsterdam dissertation.

Hacking, Ian. 2001. *An introduction to probability and inductive logic*. Cambridge University Press.

Halpern, Joseph Y. 1997. Defining relative likelihood in partially-ordered preferential structures. *Journal of Artificial Intelligence Research* 7. 1–24.

Halpern, Joseph Y. 2003. *Reasoning about Uncertainty*. MIT Press.

Hamblin, Charles L. 1973. Questions in Montague English. *Foundations of Language* 41–53.

Hare, R.M. 1967. Some alleged differences between imperatives and indicatives. *Mind* 76(303). 309.

Harman, Gilbert. 1993. Stringency of Rights and "Ought". *Philosophy and Phenomenological Research* 53(1). 181–185.

Harsanyi, John C. 1978. Bayesian decision theory and utilitarian ethics. *The American Economic Review* 68(2). 223–228.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183.

Heim, Irene. 2001. Degree operators and scope. In Fery & Sternefeld (eds.), *Audiatur vox sapientiae: A festschrift for arnim von stechow*, Berlin: Akademie Verlag.

Heim, Irene. 2006. Remarks on comparative clauses as generalized quantifiers. Ms., MIT .

Helmholtz, Hermann von. 1887. Zahlen und Messen erkenntnistheoretisch betrachtet. *Philosophische Aufsätze Eduard Zeller gewidmet* 356–391.

Hintikka, Jaako. 1962. *Knowledge and Belief: An Introduction to the Logic of the Two Notions*. Cornell University Press.

Hölder, Otto. 1901. Die Axiome der Quantität und die Lehre vom Mass. *Ber. Verh. Kgl. Sächsis. Ges. Wiss. Leipzig, Math.-Phys. Klasse* 53. 1–64.

Horn, Laurence. 1989. *A Natural History of Negation*. University of Chicago Press.

Horn, Laurence R. 1972. *On the Semantic Properties of Logical Operators in English*: UCLA dissertation.

Horney, Karen. 1942. *The collected works of Karen Horney (volume II)*. W.W. Norton & Company.

Jackson, Frank. 1985. On the semantics and logic of obligation. *Mind* 94(374). 177.

Jackson, Frank. 1991. Decision-theoretic consequentialism and the nearest and dearest objection. *Ethics* 101(3). 461–482.

Jackson, Frank & Robert Pargetter. 1986. Oughts, options, and actualism. *The Philosophical Review*

95(2). 233–255.
Jeffrey, Richard C. 1965a. Ethics and the logic of decision. *Journal of Philosophy* 62(19). 528–539.
Jeffrey, Richard C. 1965b. *The logic of decision*. University of Chicago Press.
Joyce, James M. 1999. *The foundations of causal decision theory*. Cambridge University Press.
Kahneman, Daniel, Paul Slovic & Amos Tversky. 1982. *Judgment under Uncertainty: Heuristics and Biases*. Cambridge University Press.
Kant, Immanuel. 1797. Über ein vermeintes Recht aus Menschenlieve zu lügen. *Berliner Blätter* .
Keenan, Edward L. & Leonard M. Faltz. 1985. *Boolean semantics for natural language*. Kluwer Academic Publishers.
Kennedy, Chris. 1997. *Projecting the adjective: The syntax and semantics of gradability and comparison*: U.C., Santa Cruz dissertation.
Kennedy, Chris. 2001. Polar opposition and the ontology of 'degrees'. *Linguistics and Philosophy* 24(1). 33–70.
Kennedy, Chris. 2007. Vagueness and grammar: The semantics of relative and absolute gradable adjectives. *Linguistics and Philosophy* 30(1). 1–45.
Kennedy, Chris & Louise McNally. 2005. Scale structure, degree modification, and the semantics of gradable predicates. *Language* 81(2). 345–381.
Keynes, John Maynard. 1921. *A Treatise on Probability*. Macmillan.
Klecha, Peter. 2011. Positive and conditional semantics for gradable modals. In *Proceedings of sinn und bedeutung 16* (to appear), .
Klein, Ewan. 1980. A semantics for positive and comparative adjectives. *Linguistics and Philosophy* 4(1). 1–45.
Klein, Ewan. 1982. The interpretation of adjectival comparatives. *Journal of Linguistics* 18(1). 113–136.
Klein, Ewan. 1991. Comparatives. In A. von Stechow & D. Wunderlich (eds.), *Semantik: Ein internationales handbuch der zeitgenössischen forschung*, Walter de Gruyter.
Kolmogorov, Andrey. 1933. *Grundbegriffe der Wahrscheinlichkeitsrechnung*. Julius Springer.
Kolodny, N. & J. MacFarlane. 2010. Ifs and oughts. *The Journal of philosophy* 107(3). 115–143.
Kraft, Charles H., John W. Pratt & A. Seidenberg. 1959. Intuitive probability on finite sets. *The Annals of Mathematical Statistics* 30(2). 408–419.
Krantz, David H., R. Duncan Luce, Patrick Suppes & Amos Tversky. 1971. *Foundations of Measurement*. Academic Press.
Kratzer, Angelika. 1981. The notional category of modality. In Hans-Jürgen Eikmeyer & Hannes Rieser (eds.), *Words, worlds, and contexts: New approaches in word semantics*, 38–74. de Gruyter.
Kratzer, Angelika. 1986. Conditionals. In *Chicago linguistics society*, vol. 22 2, 1–15.
Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, de Gruyter.
Kratzer, Angelika. 2012. *Modality and Conditionals: New and Revised Perspectives*. Oxford University Press (To Appear).
Krifka, Manfred. 1989. Nominal reference, temporal constitution and quantification in event semantics. *Semantics and contextual expression* 75. 115.

Krifka, Manfred. 1990. Four thousand ships passed through the lock: Object-induced measure functions on events. *Linguistics and Philosophy* 13(5). 487–520.

Krifka, Manfred. 1998. The origins of telicity. *Events and grammar* 197. 235.

Krifka, Manfred. 2007a. Approximate interpretation of number words: A case for strategic communication. *Cognitive foundations of interpretation* 111–126.

Krifka, Manfred. 2007b. Basic notions of information structure. In Caroline F''ery, Gisbert Fanselow & Manfred Krifka (eds.), *Working papers of the sfb632, interdisciplinary studies on information structure*, vol. 6, 13–36. Universitätsverlag Potsdam.

Kripke, Saul. 1963. Semantical considerations on modal logic. *Acta philosophica fennica* 16(1963). 83–94.

Laplace, Pierre. 1829. Essai philosophique sur les probabilités.

Lassiter, Daniel. 2010a. Gradable epistemic modals, probability, and scale structure. In Nan Li & David Lutz (eds.), *Semantics and linguistic theory (SALT) 20*, 197–215. Ithaca, NY: CLC Publications. http://elanguage.net/journals/index.php/salt/article/view/20.197.

Lassiter, Daniel. 2010b. The Algebraic Structure of Amounts: Evidence from Comparatives. *Interfaces: Explorations in Logic, Language and Computation* 38–56.

Lassiter, Daniel. 2011a. Nouwen's puzzle and a scalar semantics for obligations, needs, and desires. In Neil Ashton, Anca Chereches, & David Lutz (eds.), *Semantics and linguistic theory 21*, eLanguage.

Lassiter, Daniel. 2011b. Vagueness as probabilistic linguistic knowledge. In U. Sauerland R. Nouwen, R. van Rooij & H.-C. Schmitz (eds.), *Vagueness in communication*, 127–150. Springer.

Lemmon, E.J. 1962. Moral dilemmas. *The Philosophical Review* 139–158.

Levinson, Dmitry. 2003. Probabilistic model-theoretic semantics for *want*. *Semantics and Linguistic Theory 13* .

Lewis, David. 1973. *Counterfactuals*. Harvard University Press.

Lewis, David. 1978. Reply to McMichael. *Analysis* 38(2). 85.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1). 339–359. doi:10.1007/BF00258436.

Lewis, David. 1981. Ordering semantics and premise semantics for counterfactuals. *Journal of Philosophical Logic* 10(2). 217–234.

Link, Godehard. 1983. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bäuerle, C. Schwarze & A. von Stechow (eds.), *Meaning, use and interpretation of language*, vol. 21, 302–323. Walter de Gruyter.

Link, Godehard. 1998. *Algebraic Semantics in Language and Philosophy*. CSLI Publications.

Luce, R. Duncan & Louis Narens. 1985. Classification of concatenation measurement structures according to scale type. *Journal of Mathematical Psychology* 29(1). 1–72.

Macchi, Laura, Daniel Osherson & David H. Krantz. 1999. A note on superadditive probability judgment. *Psychological Review* 106(1). 210.

MacFarlane, John. 2011. Epistemic modals are assessment-sensitive. In A. Egan & B. Weatherson (eds.), *Epistemic modality*, Oxford University Press.

MacKay, David J.C. 2003. *Information theory, inference, and learning algorithms*. Cambridge

University Press. http://www.inference.phy.cam.ac.uk/itprnn/book.pdf.

Mellor, D.H. 2005. *Probability: A philosophical introduction*. Routledge.

Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In J. Hintikka, J. Moravcsik & P. Suppes (eds.), *Approaches to natural language*, vol. 49, 221–242. Reidel.

Narens, Louis. 1985. *Abstract Measurement Theory*. MIT Press.

Narens, Louis. 2007. *Theories of probability: an examination of logical and qualitative foundations*. World Scientific Pub Co Inc.

Nerbonne, John. 1995. Nominal comparatives and generalized quantifiers. *Journal of Logic, Language and Information* 4(4). 273–300.

von Neumann, Jon & Oskar Morgenstern. 1944. *Theory of Games and Economic Behavior*. Princeton University Press.

Nouwen, Rick. 2008. Upper-bounded *no more*: the exhaustive interpretation of non-strict comparison. *Natural Language Semantics* 16. 271–295.

Nouwen, Rick. 2010a. Two kinds of modified numerals. *Semantics and Pragmatics* 3(3). 1–41.

Nouwen, Rick. 2010b. Two puzzles about requirements. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager & Katrin Schulz (eds.), *Logic, language and meaning: 17th amsterdam colloquium*, 345–354. Springer.

Nozick, Robert. 1969. Newcomb's problem and two principles of choice. In Nicholas Rescher (ed.), *Essays in honor of carl g. hempel: A tribute on the occasion of his sixty-fifth birthday*, 114–115. D. Reidel.

Pearl, Judea. 1988. *Probabilistic Reasoning in Intelligent Systems: Networks of Plausible Inference*. Morgan Kaufmann.

Pearl, Judea. 2000. *Causality: models, reasoning and inference*. Cambridge University Press.

Piatelli-Palmarini, Massimo. 1994. *Inevitable Illusions: How Mistakes of reason Rule Our Minds*. Wiley.

Portner, Paul. 2009. *Modality*. Oxford University Press.

Potts, Christopher. 2008. Interpretive Economy and Schelling points. Ms., University of Massachusetts at Amherst .

Regan, Donald. 1980. *Utilitarianism and Co-operation*. Oxford University Press.

Roberts, Craige. 1996. Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics* 91–136. http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.129.8440&rep=rep1&type=pdf.

Roberts, Fred S. 1979. *Measurement Theory with Applications to Decisionmaking, Utility, and the Social Sciences*. Addison Wesley Publishing Company.

Roberts, Fred S. & R. Duncan Luce. 1968. Axiomatic thermodynamics and extensive measurement. *Synthese* 18(4). 311–326.

van Rooij, Robert. 1999. Some analyses of pro-attitudes. In H. de Swart (ed.), *Logic, game theory, and social choice*, Tilburg University Press.

van Rooij, Robert. 2003. Questioning to resolve decision problems. *Linguistics and Philosophy* 26(6). 727–763. doi:10.1023/B:LING.0000004548.98658.8f.

van Rooij, Robert. 2004. Utility, informativity and protocols. *Journal of Philosophical Logic* 33(4). 389–419. doi:10.1023/B:LOGI.0000036830.62877.ee.

van Rooij, Robert. 2009. Up and down the scale: Adjectives, comparisons, and measurement. Ms., ILLC, Universteit van Amsterdam .
van Rooij, Robert. 2010. Measurement, and interadjective comparisons. *Journal of Semantics* .
Rooth, Mats. 1985. *Association with Focus*: University of Massachusetts dissertation.
Rooth, Mats. 1992. A theory of focus interpretation. *Natural language semantics* 1(1). 75–116.
Ross, Alf. 1944. Imperatives and logic. *Philosophy of Science* 30–46.
Ross, Jacob. 2010. The Irreducibility of Personal Obligation. *Journal of Philosophical Logic* 39(3). 307–323.
Rothschild, Daniel. 2011. A note on conditionals and restrictors. Ms., Oxford University .
Rotstein, Carmen & Yoad Winter. 2004. Total adjectives vs. partial adjectives: Scale structure and higher-order modifiers. *Natural Language Semantics* 12(3). 259–288.
Rullmann, Hotze. 1995. *Maximality in the Semantics of* wh-*constructions*: University of Massachusetts, Amherst dissertation.
Sapir, Edward. 1944. Grading: A study in semantics. *Philosophy of Science* 11(2). 93–116.
Sassoon, Galit. 2007. *Vagueness, Gradability and Typicality, A Comprehensive semantic analysis*: Tel Aviv University dissertation.
Sassoon, Galit. 2010. Measurement theory in linguistics. *Synthese* 174(1). 151–180.
Sauerland, Uli & Penka Stateva. 2007. Scalar vs. epistemic vagueness: Evidence from approximators. In M. Gibson & T. Friedman (eds.), *Proceedings of semantics and linguistic theory xvii*, CLC Publications, Cornell University.
Savage, Leonard J. 1954. *The Foundations of Statistics*. Wiley.
Schmidt, Laura A., Noah D. Goodman, David Barner & Joshua B. Tenenbaum. 2009. How tall is *Tall*? Compositionality, statistics, and gradable adjectives. In *Proceedings of the 31st annual conference of the cognitive science society*, .
Schulz, Katrin. 2007. *Minimal models in semantics and pragmatics: Free choice, exhaustivity, and conditionals*: Institute for Logic, Language and Computation, University of Amsterdam dissertation.
Schulz, Katrin. 2011. "If you'd wiggled A, then B would've changed": Causality and counterfactual conditionals. *Synthese* 179(2). 239–251. doi:10.1007/s11229-010-9780-9.
Schwarzschild, Roger. 2004. Scope-splitting in the comparative. Handout from MIT colloquium, available http://www.rci.rutgers.edu/~tapuz/MIT04.pdf .
Schwarzschild, Roger & Karina Wilkinson. 2002. Quantifiers in comparatives: A semantics of degree based on intervals. *Natural Language Semantics* 10(1). 1–41.
Scott, Dana & Patrick Suppes. 1958. Foundational aspects of theories of measurement. *Journal of Symbolic Logic* 23(2). 113–128.
Sen, Amartanya. 1969. Quasi-transitivity, rational choice and collective decisions. *The Review of Economic Studies* 36(3). 381–393.
Shannon, Claude E. 1948. A Mathematical Theory of Communication. *Bell System Technical Journal* 27. 379–423.
Shoham, Yoav & Kevin Leyton-Brown. 2009. *Multiagent systems: Algorithmic, game-theoretic, and logical foundations*. Cambridge University Press.
Simons, Mandy. 2005. Dividing things up: The semantics of or and the modal/or interaction.

Natural Language Semantics 13(3). 271–316.
Sloman, Aaron. 1970. 'Ought' and 'Better'. *Mind* 75(315). 385–394.
Slovic, Paul, Baruch Fischhoff & Sarah Lichtenstein. 1976. Cognitive processes and societal risk taking. In J. S. Carrol & J. W. Payne (eds.), *Cognition and social behavior*, Erlbaum.
Solt, Stephanie. 2011. Notes on the comparison class. In U. Sauerland R. Nouwen, R. van Rooij & H.-C. Schmitz (eds.), *Vagueness in communication*, 127–150. Springer.
Starr, William. 2010. *Conditionals, Meaning and Mood*: Rutgers dissertation.
von Stechow, Arnim. 1984. Comparing semantic theories of comparison. *Journal of Semantics* 3(1). 1–77.
Stevens, S.S. 1946. On the theory of scales of measurement. *Science* 103(2684). 677–680.
Suppes, Patrick. 1959. Measurement, empirical meaningfulness, and three-valued logic. *Measurement: definitions and theories* 129.
Suppes, Patrick & Joseph L. Zinnes. 1963. Basic measurement theory. In R. D. Luce, R. R. Bush & E. Galanter (eds.), *Handbook of mathematical psychology*, vol. 1, 1–63. Wiley.
Swanson, Eric. 2006. *Interactions With Context*: MIT dissertation. http://mit.dspace.org/bitstream/handle/1721.1/37356/123190034.pdf?sequence=1.
Swanson, Eric. 2008. Modality in language. *Philosophy Compass* 3(6). 1193–1207.
Swanson, Eric. 2011. On the treatment of incomparability in ordering semantics and premise semantics. *Journal of Philosophical Logic* 1–21.
Szabolcsi, Anna & Frans Zwarts. 1993. Weak islands and an algebraic semantics for scope taking. *Natural Language Semantics* 1(3). 235–284.
Teigen, Karl. 1988. When are low-probability events judged to be 'probable'? Effects of outcome-set characteristics on verbal probability estimates. *Acta Psychologica* 68. 157–174.
Tversky, Amos & Derek J. Koehler. 1994. Support theory: A nonextensional representation of subjective probability. *Psychological Review* 101(4). 547–566.
Unger, Peter. 1971. A defense of skepticism. *The Philosophical Review* 80(2). 198–219.
Villalta, Elisabeth. 2008. Mood and gradability: An investigation of the subjunctive mood in Spanish. *Linguistics and Philosophy* 31(4). 467–522.
Wedgwood, R. 2006. The Meaning of "Ought". *Oxford studies in metaethics* 1. 127–60.
Windschitl, P.D. & G.L. Wells. 1998. The alternative-outcomes effect. *Journal of Personality and Social Psychology* 75(6). 1411–1423.
Winter, Yoad. 2001. *Flexibility Principles in Boolean Semantics: The Interpretation of Coordination, Plurality, and Scope in Natural Language*. The MIT Press.
Yalcin, Seth. 2007. Epistemic modals. *Mind* 116(464). 983–1026. doi:10.1093/mind/fzm983.
Yalcin, Seth. 2009. The language of probability. Talk presented at the Department of Linguistics, U.C. Berkeley .
Yalcin, Seth. 2010. Probability Operators. *Philosophy Compass* 5(11). 916–937. doi:10.1111/j.1747-9991.2010.00360.x.
Yalcin, Seth. 2011. Nonfactualism about epistemic modality. In A. Egan & B. Weatherson (eds.), *Epistemic modality*, Oxford University Press.
Zadeh, L.A. 1965. Fuzzy sets. *Information and Control* 8(3). 338–353.
Zadeh, L.A. 1978. Fuzzy sets as a basis for a theory of possibility. *Fuzzy Sets and Systems* 1(1).

3–28.
Zimmermann, Thomas Ede. 2000. Free choice disjunction and epistemic possibility. *Natural Language Semantics* 8(4). 255–290.