Underspecificity in Modal Contexts

Richard Jefferson Booth

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
under the Executive Committee
of the Graduate School of Arts and Sciences

COLUMBIA UNIVERSITY

2022

Abstract

Underspecificity in Modal Contexts

Richard Jefferson Booth

This dissertation is on the semantics of modal expressions and attitude verbs like English 'must,' 'may,' 'ought,' and 'want' — expressions that allow us to discuss socially important modal facts like obligations, permissions, and desires. When discussing them, we rarely spell out the modal facts with complete specificity. For example, I may say, 'I want ice cream,' without specifying that what I desire, in particular, is unmelted, non-toxic, chocolate ice cream. If, in response, someone were to give me melted, toxic, vanilla ice cream, it would be fair for me to reply that what I have been given is not what I wanted. Similarly, I can truly say 'you must wash the dishes,' or 'you may have some wine,' without specifying that there are exceptions, that is, specific ways of washing the dishes, or of having wine, that are not ways of doing what you must or are allowed. I call modal claims that have such exceptions *underspecific*. Since avoiding underspecificity may require explicitly ruling out an infinite number of exceptions, most of our ordinary modal claims are underspecific. Our reliance on them is thus crucial to our very capacity to communicate about the modal facts.

I argue that modal claims can be true and underspecific by virtue of the meanings of the modal expressions that figure in them, and I defend a general semantic framework for modals that explains how. In the simplest cases, my framework predicts that speakers may describe the contents of desires, permissions, and obligations using conditions that are merely necessary for their fulfillment. Since getting unmelted, non-toxic chocolate ice cream necessitates getting ice cream *simpliciter*, this explains how I may truly state 'I want ice cream' even if the content of my desire is more specific.

Generalizing this account, however, poses significant challenges. Indeed, many philosophers and linguists have rejected semantic theories of modals that allow for true, underspecific modal claims for two main reasons. One is that previous accounts give rise

to several well-known logical puzzles, including the puzzle of free choice permission, Ross's puzzle, and the Samaritan paradox. A second reason is that existing accounts generate the wrong truth value judgments in examples involving complex preferences and forms of uncertainty. In response, theorists have offered semantics for modals like 'ought' and 'want' that rely on the theory of rational choice and the notion of expected utility. Straightforward expected utility analyses succeed in generating the desired truth value judgments, but at the cost of giving up the possibility of underspecifying the modal facts in the ways that we manifestly do.

I argue that we can meet both of these challenges, while still allowing for true, underspecific modal claims. In order to do so, I develop new solutions to the puzzle of free choice permission, Ross's puzzle, and the Samaritan paradox that improve on the empirical predictions of existing solutions. My framework also generates a methodologically desirable degree of independence between the proper semantics of modal expressions, on the one hand, and the proper analyses of desire or obligation in terms of rational choice, on the other. As a result, the semantic framework I develop in this dissertation can succeed where theorists who have turned to decision-theoretic semantics for 'want' and 'ought' have thought that existing theories fail. Besides the advantages that my semantic framework gains with respect to these problems in particular, it more generally provides new insights into the rich linguistic capacities and conventions that speakers leverage in order to communicate about the modal facts.

Table of Contents

List of Figures

Acknowledgements

This dissertation has two distinct roots that happened to grow together at the right time. The first was work I did for a seminar at NYU co-taught by Jessica Moss and David Velleman on contemporary Aristotelian ethics. In attempting to defend Donald Davidson's theory of 'pure' intention as a species of desire against some of Velleman's criticisms, I became particularly obsessed with some questions Davidson considers in the process of developing his view: what factors, exactly, determine the contents of an agent's desires? And how do our descriptions of our desires relate to their contents? These were the sorts of questions that seemed like they must have an obvious answer, but I quickly became convinced that they did not; perhaps, I thought, they even had *no* general answer. A major step along the way to realizing the problem was much more difficult and interesting than I had expected was reading Delia Graff Fara's paper, "Specifying Desires" (2013, *Noûs*). I was very saddened to learn that she passed soon after I discovered her work, and regret not having the chance to thank her in person for inspiring me.

For me, becoming convinced of the difficulty of answering the question about desires was both troubling and philosophically exhilarating; I felt the same sort of gleeful terror that I had felt when I first encountered Humean skepticism about induction — only this time, there did not already exist several proposed solutions and dissolutions of the problem. It felt like this was new ground to explore, and perhaps a place to develop a dissertation.

The second root of this dissertation was a seminar on nominalism with Achille Varzi. In that seminar, I became increasingly passionate about distinguishing the 'metaphysics' required for an adequate semantics of natural language from genuine metaphysical questions about the nature of reality. Achille was supportive of the general theme, but always careful to point out that these things were not as easy to distinguish as I sometimes pretended. In the process of forming my views, I became drawn to varieties of fictionalism, and in particular, Stephen Yablo's work, which had recently moved from metaphysical figuralism into a more general theory of aboutness and partial content. Yablo's work on aboutness was a trove of inspiration for me. First, it helped me realize that many of the problems I had identified in the attempt to systematically relate our descriptions of desires to their contents were much more general: they seemed to apply to most if not all modal states. Second, Yablo's work taught me that there was a vast variety of technical machinery beyond standard possible world semantics that could be put to productive use in addressing some of the problems with which I was concerned. Finally, I am also deeply indebted to Yablo for providing me with an example of a kind of philosophical optimism — about the possibility of actually shedding light on seemingly intractable problems — that reached my previously pessimistic temperament well enough that I was inspired to try, in my own way, to emulate it.

I would like to thank my dissertation committee for their boundless support, unwavering encouragement, and generous advice. Achille was always open to and supportive of venturing off the beaten path in search of solutions to the problems I set for myself, and was able to identify virtues and vices of doing so well before I could. He was a constant source of optimism, and a much-needed antidote to the sort of cold professionalism I had sometimes encountered. As one of the philosophical writers I most admire, his feedback on my own writing was invaluable, and I took his generous compliments on my drafts to heart.

I would like to thank Karen Lewis for her incredibly helpful feedback on my chapters. Her deep knowledge of the field helped me identify and stay focused on the issues properly belonging to my subject, when I sometimes began to drift into other territory. She was a talented advocate for views I was opposing, helping me develop my arguments against them much further than if I had focused only on what was in print. I would also like to thank her for her patience and care. I still often think about the first meeting I had with her about the beginnings of my dissertation proposal. After I anxiously tried to quickly introduce her to the questions I was interested in, she prompted me to slow down and be precise: what words, exactly, was I trying to give a semantics for, and what linguistic data, exactly, was I trying to explain?

I would like to thank Jessica Collins for her generous discussions with me about my chapters. I found her always both enthusiastic and difficult to fully persuade. This distinctive combination usefully pushed me to develop my views more fully, and to work to justify premises I had been taking for granted without noticing.

Outside my committee proper, I also want to give special thanks to Melissa Fusco, who gave me as much attention as if it had been her duty. Melissa's extensive and insightful feedback was invariably thorough, sometimes consisting in long email threads delving into foundational questions that no one had asked me about my account. My discussions with her have helped me develop my views in crucial ways, and I'm extremely grateful to her.

I am very grateful to Paolo Santorio for agreeing to serve as an external member of my defense committee, and for his probing questions about each of my chapters.

For providing crucial feedback, suggestions, and advice on some of these chapters as they were written, I would like to thank Kyle Blumberg, Ezra Keshet, Angelika Kratzer, Ben Holguin, Ben Lennertz, Wolfgang Mann, Christopher Peacocke, and Una Stojnic.

For providing this near-high-school-dropout the opportunity to eventually pursue a Ph.D., I would also like to thank the California public higher-education system, and Prof. Julian DelGaudio and Prof. Lisa Orr, whom I met at Long Beach City College. They gave

me encouragement, patience, and advice at the beginning of my academic career.

I would also like to thank my mother, Melody House, for her sacrifices and support. Reuniting with her in 2011 after our separation by adoption was one of the most meaningful experiences of my life. The love and support she has given me since then have been a ballast while navigating the tumult of a Ph.D.

Finally, my most profound thanks are to my partner, Janum Sethi. Her patience, love, and philosophical support have been so important for my work that it is impossible for me to imagine what it would have looked like without her. I have been frequently amazed by her ability to jump into a discussion of a technical problem I was stuck on, and to offer helpful insights and suggestions. Working on a Ph.D. has sometimes detracted from my ability to be the partner I would like to be, and among the various legitimate responses to these failures, Janum fortunately chose patience and understanding. She has helped me recover hopefulness in the face of discouragement, and (just as difficult) to celebrate encouraging news. Her presence in my life throughout my time in graduate school is the primary reason that I was able to pursue a core value that was sometimes at odds with the demands of the Ph.D. program: to lead a rich, interesting, and beautiful life that consists in more than just philosophy.

Dedication

For Janum, and for Melody.

Chapter 1: Introduction and Background

Like our nearest relatives, humans are social creatures. We coordinate with one another to achieve things that no individual could accomplish alone. The possibility of successful coordination is not trivial: groups of agents usually have many distinct ways that they may accomplish a goal, but success requires that each of their actions align in the right way. Going to my house or to your house for dinner might be equally good options; what is necessary for us to have dinner together is that we both go to the same place. Solving coordination problems like these requires that the agents involved share an understanding of the structure of the problems they face, each others' abilities, preferences, rationality, and numerous other factors.

Unlike our nearest extant relatives, humanity has a powerful tool for reaching the sort of common understanding requisite for successful coordination: language. Just by discussing them, we can quickly and concisely reach a common understanding of our options, each others' desires, beliefs, plans, abilities, obligations, permissions, and so on. These attitudes and states mentioned are often called intensional or *modal* (this is the term I will use) states, since they are relations to propositions that may not be true of the world as it actually is, but of the world as it could, should, or would be. Given the key role that a shared understanding of modal states plays in solving coordination problems, it is hardly surprising that despite the diverse types of linguistic constructions they use to do so, all known human languages appear to have some way of describing them (Matthewson 2016). In English, we can use auxiliaries like 'must', verbs like 'want' or 'plan', adverbs like 'necessarily', suffixes like '-able' to talk about modal states.

Normally, the modal states we are interested in are highly complex and detailed. The contents of my obligations, for example, might depend on a vast number of factors including moral duties, my and others' preferences, the environment, cultural norms, and what my options are. If I am obliged to prepare a cortado for a customer at the café, for example, it is usually not the case that any way of preparing it will fulfill my obligation. Perhaps I know that given the customer's preferences, it can only be satisfied by a cortado made with espresso of a certain quality, with plant-based milk, and within a certain temperature range. Perhaps prevailing cultural norms mean that my obligation can only be fulfilled if I prepare the cortado in a certain glass. Perhaps the law requires that the milk I use not be past its stated expiration date. Further reflection quickly reveals that even if the customer is not especially picky, the list of necessary properties the cortado I prepare must have in order to satisfy my obligation may be infinite: the drink must contain no discernable quantity of sharp glass, no anti-freeze, none of my phlegm, and so on, *ad infinitum*. The same sorts of considerations apply to all of the other modal states I have mentioned. The contents of mundane desires, beliefs, plans, abilities, obligations, and permissions are often infinitely more specific than they may appear.

There are some limits on how detailed we can be in using natural languages to refer to the propositions that form the contents of modal states. Because natural languages are recursive, sentences can in principle be arbitrarily long. But because they must be uttered in finite time, they cannot be *infinitely* long. In practice, they tend to be fairly short relative to the vast universe of in principle possible lengths. Furthermore, the lexicons of natural languages are finite. If, as is usually assumed, the space of possible worlds is infinite, then the set of all propositions, normally assumed to be at least as numerous as the set of all sets of worlds, is also infinite. Thus, the lexicons of natural languages do not include words uniquely identifying each proposition. To refer to most propositions, we must use finite sentences built compositionally from a finite vocabulary. If the propositional contents of the modal states we need to refer to are often infinitely specific, as I am suggesting, then a language in which we could only truly state what the modal facts are if we specified them in complete detail would be one in which true modal claims would be nearly impossible to make.

Fortunately, then, the semantics of modal expressions often allows us to state truths about the modal facts *without* being fully explicit about every detail making up their contents. A sentence like "I ought to make coffee," seems to require for its truth that it is *necessary* to make coffee in order to do what I ought to, but making coffee need not also be *sufficient* for fulfilling my obligation. As just discussed, there may be an infinite number of other additional necessary conditions that circumscribe the content of my obligation. The ability to make true claims about the modal facts that *underspecify* their contents means that we can make such claims in a compact (and in particular, finite) form, even if the contents of the modal states we talk about are infinitely specific.

The possibility of true, underspecific modal claims simplifies the task of coming to a shared understanding of the nature of a coordination problem and agreeing upon a solution. In order to share accurate information about my abilities, beliefs, preferences, and obligations, I need not specify every detail of their contents. Rather, I need only specify those details about the contents of these modal states that are *relevant* to the problem at hand. If it is common knowledge that no one typically wants to consume coffee mixed with a dash of antifreeze, then a customer ordering a cortado at the café need not mention that they require their coffee not to include antifreeze, and may feel assured that they have not committed themselves to accepting a drink that includes that ingredient by leaving out that detail.

Of course, the parsimony and flexibility that the semantic possibility of true, underspecific modal claims allows places a greater weight on the *pragmatics* of modal claims in successful coordination. Consider again a simple café problem: a café-goer and a barista need to coordinate their actions so that the barista prepares a drink that the café-goer likes. If the café-goer says "I want a cortado," the barista will need to rely on broadly Gricean reasoning about why the café-goer said what they did in the context in order to answer any questions that might arise about what unstated further conditions might attach to the fulfillment of their request. For example, if it is commonly understood that all espresso drinks may be prepared hot or iced, and that by default, drinks are prepared hot, then the barista has a good reason to infer from the fact that ice was not mentioned that the café-goer desires, more specifically, a *hot* cortado. Similarly, the café-goer must rely on broadly Gricean reasoning about the context in order to choose which sentence to use to state her request in order to coordinate with the barista on a successful outcome. In the context just described, they need to recognize that not mentioning whether they would like their cortado prepared hot or on ice will lead the barista to believe that their desire is for a hot drink in particular.

Although we navigate these situations fluently most of the time, failures of these forms of pragmatic reasoning are common, and are a major source of literary drama. In W.W. Jacobs's classic horror tale, "The Monkey's Paw," Mr. White makes a wish of the enchanted appendage of the title: "I wish for two hundred pounds." The next day, Mr. White is informed that his son has been mutilated and killed while at work, and that while the company denies any liability, they have offered a good will payment of two hundred pounds. Certainly one way to understand the tale is to hold that Mr. White, unused to stating a wish to a supernatural and malevolent interpreter, failed to recognize that he needed to be much more specific in stating its content: something more like, "I wish that I would get two hundred pounds *and that everything else I care about would go on as it would have if I hadn't made this wish*." Any dread one might naturally feel that no matter how specific the statement of the wish, there will always be some grotesque and undesirable way of realizing it, is precisely to my point. When it comes to modal facts like wishes, we are almost always left having to underspecify their contents.

1.1 The Source of Underspecificity

1.1.1 Basic Theories of Modal Semantics

The modal expressions that will be the focus of this dissertation all take a sentential argument at the level of syntax called the *prejacent* or *complement clause* that denotes a proposition. To illustrate (while glossing over some syntactic complications), in the following examples, the prejacent of the underlined modal expression in the (a) sentence is something like the (b) sentence:[1]

(1.1) a. They must be home.

b. They are home.

(1.2) a. They might be hungry.

b. They are hungry.

(1.3) a. She said they were coming.

b. They were coming.

(1.4) a. He wants ice cream.

b. He gets ice cream.

On *relational* semantics for modal logics due to Kripke (1963), modal operators like 'it is necessary that' and 'it is possible that' denote a relation between a set of *accessible worlds* and the set of worlds that make their prejacent clauses true. On the Kripkean theory, these relations are the same as those denoted by nominal quantifiers like 'every' and 'some'. A sentence like 'It is necessary that they are home' is true if and only if *every* accessible world is one that makes the complement clause 'they are home' true. 'It is

1. The complications I gloss over are foremost syntactic changes in mood and tense. The case of 'want' in the last example is even more complex, since current syntactic theory suggests that rather than 'he', the prejacent contains an unlexicalized reflexive pronoun often denoted PRO indexed to the subject of 'want', and rather than ordinary 'gets' the verb involved is unlexicalized and is associated with a weaker semantics. See Harley (2012) for discussion.

possible that they are home' is true if and only if there is *some* possible world that makes 'they are home' true.

Hintikka (1969) argued that we should understand the semantics of attitude verbs in a similar way. Consider:

(1.5) a. A believes Pim is home.

b. A wants Pim to be home.

c. A said Pim is home.

According to Hintikka, we should think of people's attitudes as determining sets of accessible worlds: given an agent A, there is the set of worlds compatible with her beliefs, $\mathsf{Bel}(A)$; the set of worlds compatible with her desires, $\mathsf{Des}(A)$; the set of worlds compatible with what she has said $\mathsf{Say}(A)$, and so on. (1.5a-1.5c) are true if and only if every world in $\mathsf{Bel}(A)$, $\mathsf{Des}(A)$, or $\mathsf{Say}(A)$, respectively, is one where Pim is home.

Contemporary research on the semantics of natural language modals builds on these ideas of Kripke and Hintikka. The dominant theory is due to Kratzer (1977, 1981, 1991). Kratzer noticed that meaningful uses of modal expressions in general could be categorized along two dimensions: their *flavor* and their *force*. In English, modals like 'might' and 'may' have the force of *possibility*: they express a weak relation like compatibility or existential quantification. Modals like 'must' and 'ought', on the other hand, have the force of *necessity*: they express a stronger relation like entailment or universal quantification.

In English and many other languages, the same modal words, for example, 'must,' and 'may,' can express different *flavors* of modality on different occasions of use. Consider:

(1.6) He must be home.

(1.7) The dishes must be washed tomorrow.

Sentence (1.6) says something like: the proposition *that he is home* is entailed by the contents of an epistemic modal state like our *knowledge*, giving the modal 'must' in that sentence an epistemic flavor. (1.7), on the other hand, says that the proposition *that the dishes are washed tomorrow* is entailed by the contents of a 'deontic' modal state like the relevant *norms*, giving 'must' in that sentence a deontic flavor. To make sense of flexibility in which flavors of interpretation modals like 'must' can receive in different contexts, Kratzer argues that this is an aspect of their meaning that is determined by the context of use.

1.1.2 Specificity and Monotonicity

The basic semantic frameworks for analyzing modal expressions just discussed allow that modal claims can be true even if they are not entirely specific about the modal facts. In general, this is because they predict that modals with the force of necessity or possibility are *upward monotonic*:

> (Upward Monotonicity). A propositional operator O is upward monotonic if and only if, when ϕ entails ψ, $O(\phi)$ entails $O(\psi)$.

Given this definition, what upward monotonicity amounts to will depend on the notion of entailment involved. For the purposes of this introduction, I will primarily focus on its most basic form, which relies on simple, truth-conditional entailment:[2]

> (Truth-Conditional Entailment). A sentence ϕ truth-conditionally entails ψ if and only if, whenever ϕ is true, so is ψ.

This allows us to further specify the most basic form of upward monotonicity, which I will call *Naïve Upward Monotonicity*:

> (Naïve Upward Monotonicity). A propositional operator O is naïvely upward monotonic if and only if, when ϕ truth-conditionally entails ψ, $O(\phi)$ truth-conditionally entails $O(\psi)$.

2. See von Fintel (1999) and Sharvit (2017) for refinements motivated by the influence of presuppositional aspects of meaning.

We can think of entailment as the relation that grounds the comparative notion of specificity in the sense I am interested in in this dissertation. If ϕ entails ψ, then ϕ is *at least as specific as* ψ. If ψ does not also entail ϕ, then ϕ is strictly *more specific* than ψ. If neither ϕ nor ψ entails the other, then neither is as specific as the other.

Thus, if the content of your obligation(s) is p, and 'ought' has an upward monotonic necessity semantics based on the contents of the relevant modal states, then for every sentence ψ that is entailed by p (including p itself), $\ulcorner$It ought to be the case that $\psi\urcorner$ is true. This means that for all ψ that are *less specific* than p (i.e. those that are entailed by p but which do not themselves entail p), $\ulcorner$It ought to be the case that $\psi\urcorner$ will be true and *underspecify* the contents of your obligation(s).

Some form of upward monotonicity is well supported by linguistic data. Let me illustrate with some examples. First, consider the following valid pattern of inference, which trades on the removal of a subsective modifier (*chocolate* modifies *ice cream* to denote one of its *subsets*):

(1.8) a. Coco has chocolate ice cream.

b. So, Coco has ice cream.

Embedding these sentences (adjusted for tense and mood) under modal expressions preserves the intuitive validity of the pattern:

(1.9) a. Coco $\left\{\begin{array}{c}\text{must}\\ \text{has to}\\ \text{ought to}\\ \text{wants to}\\ \text{said they}\\ \text{might}\\ \text{can}\end{array}\right.$ have chocolate ice cream.

b. So, Coco $\left\{\begin{array}{c}\text{must}\\ \text{has to}\\ \text{ought to}\\ \text{wants to}\\ \text{said they}\\ \text{might}\\ \text{can}\end{array}\right.$ have ice cream.

Second, consider cases involving determinable verbs (*sprinting* is a determinate of *running*). Clearly, the following inference is intuitively valid:

(1.10) a. Bolt is sprinting.

b. So, Bolt is running.

Again, embedding similar sentences under modal expressions preserves the validity of the inferences:

(1.11) a. Bolt $\left\{\begin{array}{c}\text{must}\\ \text{has to}\\ \text{ought to}\\ \text{wants to}\\ \text{might}\\ \text{can}\end{array}\right.$ sprint.

b. So, Bolt $\left\{\begin{array}{c}\text{must}\\ \text{has to}\\ \text{ought to}\\ \text{wants to}\\ \text{might}\\ \text{can}\end{array}\right.$ run.

Third, consider cases involving conjunction elimination. Clearly, the following inference is valid:

(1.12) a. Bolt sprints and laughs.

b. So, Bolt sprints.

Again, embedding these sentences under modal expressions appears to preserve validity:[3]

(1.13) a. Bolt {must / has to / ought to / wants to / might / can} sprint and laugh.

b. So, Bolt {must / has to / ought to / wants to / might / can} sprint.

Fourth, consider cases of plural generalization and their embeddings:

(1.14) a. Coco has chocolate ice cream.

b. Vanessa has vanilla ice cream.

c. So, Coco and Vanessa both have ice cream.

(1.15) a. Coco wants chocolate ice cream.

b. Vanessa wants vanilla ice cream.

c. So, Coco and Vanessa both want ice cream.

3. The validity of this inference pattern has been questioned, but I will not discuss this data here. See Jackson and Pargetter (1986), Cariani (2013), and Blumberg and Hawthorne (forthcoming[b]) for discussion.

Both inferences appear to be valid. Besides the intuitive validity of these patterns of reasoning, a battery of other tests for entailment reinforces the idea that necessity modals are upward monotonic. I invite the reader to perform some themselves: take any of the arguments in (1.9), (1.11), or (1.13), of the form $\Delta(\phi)$, therefore $\Delta(\psi)$ where Δ is one of our modal expressions. Then the sentence ⌜if $\Delta(\phi)$, then $\Delta(\psi)$⌝ will appear trivial, the sentence ⌜$\Delta(\phi)$ and $\neg\Delta(\psi)$⌝ will appear contradictory, and the contrapositive argument ⌜it is not the case that $\Delta(\psi)$, therefore it is not the case that $\Delta(\phi)$⌝ will also appear valid.

We can even go outside of our intuitions about semantic properties like validity, contradictoriness, and triviality, and find corroboration in our syntactic competence. Negative polarity items (NPIs) are linguistic expressions that are almost only licensed in linguistic environments that are *downward monotonic*.[4] In English, these include expressions like 'at all', 'any', and 'ever'. A propositional operator $\Delta(\phi)$ is downward monotonic if and only if whenever ϕ entails ψ, $\Delta(\psi)$ entails $\Delta(\phi)$. Given standard semantics for negation, this leads to the following prediction: necessity modals are upward monotonic if and only if their negations are downward monotonic. Thus, arguments like the following should be valid, as they seem to be:

(1.16) a. Bolt doesn't have to run.

b. So, Bolt doesn't have to sprint.

Clearly, NPIs are licensed within the prejacents of negated necessity modals, but not within the prejacents of bare (un-negated) necessity modals. Compare:

(1.17) a. # I have to ever bake any cookies.

b. I do not have to ever bake any cookies.

(1.18) a. # I ought to ever bake any cookies.

b. It is not the case that I ought to ever bake any cookies.

4. This is the analysis of Ladusaw (1980), building on work by Fauconnier (1975). Of course, there are complications with this account, such as the felicity of NPIs in the context of questions: *Did you ever cook?* See Zwarts (1995) and Giannakidou (1999) for futher discussion.

Thus, even our syntactic intuitions seem to support some form of upward monotonicity for modal expressions with the force of possibility or necessity.

1.2 Doubts about Monotonicity

Theorists have recently offered semantics for modal expressions that give up the upward monotonicity property in any form, primarily for two reasons. The first is that the dominant theories of modals are committed to naïve upward monotonicity, and this gives rise to several very undesirable predictions that I outline in §1.2.1. The second is that several theorists, dissatisfied with some of the truth-conditional predictions of the dominant Kratzerian semantics for modal expressions, have turned decision-theoretical semantics for modals expressions like 'ought' and 'want'. These semantics succeed in delivering the desired predictions about the examples motivating them, but typically at the cost of giving up upward monotonicity. I outline some of these examples, theories, and their consequences in §1.2.2.

1.2.1 Puzzles and Paradoxes of Monotonicity

Despite everything in favor of monotonicity, there are some famous and puzzling counterexamples to its naïve form. One pattern of counterexample trades on the validity of disjunction introduction inferences like the following:

(1.19) a. It rained yesterday.

b. So, either it rained or it snowed yesterday.

While pragmatically odd, such inferences are generally regarded as valid. But compare the validity of such inferences when embedded under modal expressions:

(1.20) a. You must mail the letter.

b. # So, you must either mail or burn the letter.

(1.21) a. You may make a single phone call.

b. # So, you may either make a single phone call or make twenty phone calls.

These inferences seem to be much worse than pragmatically odd. Indeed, many theorists have taken them to be *invalid*. The task of reconciling the upward monotonic semantics of modals with the intuition that arguments like (1.20) and (1.21) are invalid is known as *Ross's Puzzle*.

A related problem for the semantics of the modal-disjunction interaction is the puzzle of *free choice permission*. Although disjunctions do not entail their disjuncts, sentences of the form $\Diamond(\phi \vee \psi)$ appear to entail $\Diamond\phi$ and $\Diamond\psi$. Consider:

(1.22) a. You may either go play soccer or work on your novel.

b. So, you may go play soccer.

c. So, you may work on your novel.

If we give a semantics for modals and disjunctions that accords with our intuitions about free choice inferences like this one, this presents problems for upward monotonicity. If a modal expression is both naïvely upward monotonic (validating arguments like (1.21) *and* validates arguments like (1.22), then it becomes functionally trivial: we can prove that if $\Delta(\phi)$ is true for such a modal Δ, then for any sentence at all ψ, $\Delta(\psi)$ is true. To see why, suppose $\Delta(\phi)$ is true. By the validity of disjunction introduction and upward monotonicity, $\Delta(\phi \vee \psi)$ is also true. But, by the validity of free choice inferences like (1.22), it follows that $\Delta(\psi)$ is true. This would mean that absurd inferences like the following would be valid:

(1.23) a. You may go play soccer.

b. So, you may decapitate me.

Another major puzzle for upward monotonic semantics of modals trades on presuppositional features of natural language sentences. Generally, the phrase 'the person who was robbed' is thought to introduce the presupposition that there is a unique, salient person who was robbed, and without the intervention of other linguistic phrases or pragmatic mechanisms, sentences that contain this phrase will inherit this presupposition. Among other things, this is usually taken to mean that such sentences will also *entail* that there is a person who was robbed. Thus, the following inference seems valid:

(1.24) a. Someone helped the person who was robbed.

b. So, there was a person who was robbed. ✓

But when these sentences are embedded under modals, they often seem definitely *invalid*:

(1.25) a. It ought to be that someone helps the person who was robbed.

b. So, it ought to be that there was a person who was robbed. ✗

(1.26) a. Pim wants someone to help the person who was robbed.

b. So, Pim wants there to be a person who was robbed. ✗

(1.27) a. Pim said that someone helped the person who was robbed.

b. So, Pim said that there was a person who was robbed. ✗

The challenge of reconciling upward monotonic semantics for these modals with the sense that such inferences are invalid is known as the *Samaritan Paradox*.

In response to puzzles like these, many theorists have given up upward monotonicity completely, offering semantics for modal expressions that not only prevent the puzzles from arising, but also fail to make sense of the data in favor of some form of upward monotonicity adduced above, and fail to explain how modal claims can be true even if they underspecify the relevant modal facts.

1.2.2 Decision-Theoretic Semantics

Many theorists working on the semantics of modals have recently given up upward monotonicity as a result of trying to improve on some of the predictions of standard upward monotonic theories in cases of complex uncertainty and preference. For example, take Kai von Fintel's semantics for 'want' (von Fintel 1999), which is the only tenable theory in the literature that makes 'want' upward monotonic. His semantics roughly amounts to the following: ⌜A wants ϕ⌝ is true if and only if, among all the worlds compatible with A's beliefs, the most preferred ones make ϕ true. Levinson (2003) introduces an example that he takes to show the error of these truth conditions. The structure of the example is as follows. A always prefers to get more money than less, and knows that they face the following situation. A fair coin will be flipped. If the coin lands heads, A gets \$200. If the coin lands tails, then a second event will occur (say, casting lots), with two possible outcomes associated with different chances: (Tails 1) A gets \$300 (10% probability); or (Tails 2) A gets nothing (90% probability). According to Levinson, the following sentence is true, since if the coin lands tails, it is overwhelmingly likely that A will receive nothing, and if it lands heads, A will receive \$200:

(1.28) A wants the coin to land heads.

von Fintel's truth conditions disagree; they deem (1.28) false. Given that A prefers to get the most money possible, the maximally preferred worlds compatible with A's beliefs are those in which Tails 1 obtains. Clearly, in Tails 1, the flipped coin has landed *tails*, and *not* heads. Thus, it is not true that in all of A's most preferred doxastically-possible worlds, the coin lands heads. On von Fintel's truth conditions, then, (1.28) comes out false. Since von Fintel's semantics for 'want' so closely resembles Kratzer's semantics for necessity modals, examples with exactly this form can easily be constructed for other modal expressions like 'ought' and 'must.'

Truth-value judgments about examples like these have driven many theorists to semantics that draw on theories of rational choice. For example, the semantics for 'want' that Levinson offers assumes that for each agent A, there is a cardinal utility function g_A measuring A's preferences over worlds ($g_A(w)$), and relative to a world w, a probability function $P_{A,w}$ measuring A's credences at w over the space of worlds W. Relative to these parameters, Levinson gives the following truth functions for 'want' in terms of standard subjective expected utility (with conditional probability defined in the standard way):

(Levinson's 'want'). ⌜A wants ϕ⌝ is true with respect to g, w if and only if:

$$\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\phi\urcorner \text{ is true}) > \sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\phi\urcorner \text{ is false})$$

To see what this semantics predicts about the above example of Levinson's, we may assume that the agent A's credence and utility functions have the following structure at the actual world w before the game is played (where $W = \{\textsf{Heads}, \textsf{Tails1}, \textsf{Tails2}\}$ and Heads (and similarly for the other names) does double duty as a name for the world where the coin lands heads and the sentence saying that the coin lands heads):

Heads	$g_A(\textsf{Heads}) = 200$	$P_{A,w}(\textsf{Heads} \mid \text{'Heads' is true}) = 1$
		$P_{A,w}(\textsf{Heads} \mid \text{'Heads' is false}) = 0$
Tails1	$g_A(\textsf{Tails1}) = 300$	$P_{A,w}(\textsf{Tails1} \mid \text{'Heads' is true}) = 0$
		$P_{A,w}(\textsf{Tails1} \mid \text{'Heads' is false}) = .1$
Tails2	$g_A(\textsf{Tails2}) = 0$	$P_{A,w}(\textsf{Tails2} \mid \text{'Heads' is true}) = 0$
		$P_{A,w}(\textsf{Tails2} \mid \text{'Heads' is false}) = .9$

This leads to the following calculations relevant to the evaluation of (1.28). The expected value of the coin landing heads is calculated as follows:

$$\begin{aligned}\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\textsf{Heads}\urcorner \text{ is true}) &= g_A(\textsf{Heads}) \times P_{A,w}(\textsf{Heads} \mid \text{'Heads' is true}) \\ &\quad + g_A(\textsf{Tails1}) \times P_{A,w}(\textsf{Tails1} \mid \text{'Heads' is true}) \\ &\quad + g_A(\textsf{Tails2}) \times P_{A,w}(\textsf{Tails2} \mid \text{'Heads' is true}) \\ &= 200 \times 1 \\ &\quad + 300 \times 0 \\ &\quad + 0 \times 0 \\ &= 200\end{aligned}$$

The expected value of the coin *not* landing heads is calculated as follows:

$$\begin{aligned}\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\textsf{Heads}\urcorner \text{ is false}) &= g_A(\textsf{Heads}) \times P_{A,w}(\textsf{Heads} \mid \text{`}\textsf{Heads}\text{' is false })\\ &\quad + g_A(\textsf{Tails1}) \times P_{A,w}(\textsf{Tails1} \mid \text{`}\textsf{Heads}\text{' is false })\\ &\quad + g_A(\textsf{Tails2}) \times P_{A,w}(\textsf{Tails2} \mid \text{`}\textsf{Heads}\text{' is false })\\ &= 200 \times 0\\ &\quad + 300 \times .1\\ &\quad + 0 \times .9\\ &= 30\end{aligned}$$

Thus, Levinson's truth conditions for (1.28) are satisfied:

$$\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\textsf{Heads}\urcorner \text{ is true}) = 200 > 30 = \sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner\textsf{Heads}\urcorner \text{ is false})$$

So, on Levinson's semantics, (1.28) comes out true. By adopting a different way of combining beliefs and preferences from von Fintel, Levinson predicts a different truth value for (1.28) in the described example. Decision-theoretic semantics for necessity modals like 'ought' and 'must' likewise offer different truth value predictions than the standard Kratzerian semantics in cases structurally similar to this one.

Assuming the data they hope to account for is correct (such as the judgment that (1.28) is true), these decision-theoretic semantics clearly gain some coverage over the standard Kratzerian theories. They do so, however, at an enormous cost: they predict that these modals are not upward monotonic. For even if $\ulcorner\phi\urcorner$ asymmetrically entails $\ulcorner\psi\urcorner$ and $\ulcorner A$ wants $\phi\urcorner$ is true because A's expected utility for ϕ is greater than that for not-ϕ, it is always possible that A's expected utility for ψ is not greater than A's expected utility for not-ψ. Indeed, Levinson's example witnesses this possibility. Quick calculation shows that (where Tails is the proposition that the coin did not land heads):

$$\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner \mathsf{Tails1} \urcorner \text{ is true}) > \sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner \mathsf{Tails1} \urcorner \text{ is false})$$

$$\sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner \mathsf{Tails} \urcorner \text{ is true}) \not> \sum_{w' \in W} g_A(w') \times P_{A,w}(w' \mid \ulcorner \mathsf{Tails} \urcorner \text{ is false})$$

Thus, on Levison's semantics, 'A wants the coin to land tails and to win the lot' is true, but 'A wants the coin to land tails' is not.

Many proponents of decision-theoretic semantics have seen the failure of upward monotonicity as a virtue: by giving semantics for modal expressions that are not upward monotonic, they *also* avoid generating the logical puzzles discussed in §1.2.1. The cost, however, is great: these theories thereby give up the ability to explain all of the inference patterns surveyed above and the licensing of NPIs in negated modal constructions.

1.3 Aims and Overview of the Chapters

In this dissertation, I will argue that we can give semantic theories that allow for true, underspecific modal claims, while meeting both of the challenges discussed in the last section. In order to do so, I develop new solutions to the puzzle of free choice permission, Ross's puzzle, and the Samaritan paradox that improve on the empirical predictions of existing solutions but still allow for underspecificity. The framework I develop in response to these puzzles also generates a methodologically desirable degree of independence between the proper semantics of modal expressions, on the one hand, and the proper analyses of modal states like desire or obligation in terms of rational choice, on the other. As a result, my framework can succeed where theorists who have turned to decision-theoretic semantics for 'want' and 'ought' have thought that existing theories fail.

In chapter 2, "Underspecifying Desires," I focus on underspecificity in the context of 'want' ascriptions (sentences of the form $\ulcorner A$ wants $\phi \urcorner$) in particular. I argue in detail that 'want' ascriptions are true even when they underspecify the contents of the subject's desires since (i) 'want' is upward monotonic and (ii) the set of propositions that characterize the contents of a subject's desires is not closed under entailment. I then argue that existing upward monotonic accounts of 'want' make incorrect empirical predictions and do not have the resources to account for the relationship between a true 'want' ascription and the contents of an agent's desire(s). I argue that both problems stem from tying the semantics of 'want' too directly to a function of the subject's beliefs and preferences. In response, I outline what I call a 'desire-based' semantics that gives truth conditions for 'want' ascriptions directly in terms of the contents of the subject's desires. This semantics, I argue, makes better empirical predictions than existing theories, and offers a systematic account of the relationship between true 'want' ascriptions and the contents of agents' desires that makes room for underspecificity.

In chapter 3, "Independent Alternatives: Ross's Puzzle and Free Choice," I address two well-known and much discussed puzzles for the modal-disjunction interaction that arise for semantics on which modals are upward monotonic: the puzzle of free choice permission (the challenge of explaining why $\Diamond(\phi \text{ or } \psi)$ conveys $\Diamond(\phi)$) and Ross's puzzle (the challenge of explaining why $\Box(\phi)$ does not seem to entail $\Box(\phi \text{ or } \psi)$).[5] I argue that existing solutions to these puzzles are insufficiently general one or another of two respects. First, they tend to rely on features specific to *deontic* modality where the puzzles are sharpest, but, I argue, the puzzles arise for many other non-deontic flavors of modality. Second, while they can account for the simplest versions of the two puzzles, they cannot account for minor variations of these puzzles or the interactions between modals with embedded disjuncts and conditionals. The key to a better solution, I argue, is to see that modals with embedded disjunctions (e.g. "I ought to mop the floor or wipe the tables") are used to describe the status of alternatives denoted by the disjuncts that are *independently* relevant (e.g. mopping-without-wiping, and wiping-without-mopping). I

5. Two papers based on chapter 3 are forthcoming in *Philosophical Studies* and *Proceedings of Sinn und Bedeutung: 26*.

relate the relevant notion of independence to a topological relation between sets (the *minimal covering* relation), and use this relation to outline a bilateral, inquisitve semantics for modals and disjunctions that improves on existing solutions to the puzzles in several important respects: it allows for true and underspecific modal claims; it sides with intuitive judgments in the puzzle cases; it is flexible enough to work for any flavor of modality; it retains duality between necessity and possibility modals; and it generates the correct truth conditional predictions for *negated* modals with embedded disjunctions. I conclude by arguing that the minimal covering relation at work in the semantics of the modal-disjunction interaction reveals a deep unity between the non-standard logic of modals with embedded disjunctions, on the one hand, and the more familiar logic of collective predicates and plural terms in the nominal domain, on the other.

In chapter 4, "The Samaritan Paradox," I address another well-known puzzle for the interaction between modals and *presuppositions* that arises for semantics on which modals are naïvely upward monotonic. Since 'the burglar gets caught' presupposes and therefore *entails* 'there is a burglar,' standard upward monotonic semantics for necessity modals like 'ought' predict that 'The burglar ought to be caught' entails 'There ought to be a burglar.' Intuitively, this inference (an instance of a pattern I call the *Samaritan inference*) is a complete *non sequitur*. I argue that the best candidate for a 'received' solution to the paradox, together with some standard assumptions about the lexicon, predicts that Samaritan inferences should also be non-sequiturs when modals are interpreted with *epistemic* flavors. I then present data to demonstrate that this prediction is incorrect. Furthermore, as with the puzzles that are the focus of chapter 2, existing solutions to the Samaritan paradox usually rely on properties specific to deontic flavors of modality. Since the paradox also arises for non-deontic modals like verbs of *saying* (e.g., 'Anna said the burglar was caught', does not entail 'Anna said there is a burglar'), I argue that these solutions are insufficiently generalizable. In response to these problems, I argue that the parameters that set the domains of quantification for modals and attitude verbs can make a three-fold distinction between worlds. First, they distinguish between *relevant* and *irrelevant* worlds. Second, among relevant worlds, they make a further distinction between those that are correct and those that are incorrect. For example, a requirement to catch burglars deems worlds where there are no burglars irrelevant; it counts worlds where there are burglars and they are caught *correct*; and it counts worlds where there are burglars and they are uncaught *incorrect*. I assume that *deontic* modal states like this requirement often deem worlds irrelevant. Epistemic states, by contrast, count every world relevant as a rule. I develop a formal semantics for necessity modals based on this theory of the structure of the parameters that modals are sensitive to. Together with the assumption that necessity modals presuppose there are relevant worlds where their prejacent is true, and relevant worlds where its negation is true, my theory generates the desired prediction that Samaritan inferences are invalid for deontic readings of necessity modals, but *valid* for their epistemic readings. My theory also easily extends to verbs of *saying*: by treating 'say' as a necessity modal with the same semantics, and making the plausible assumption that the relevant worlds are those compatible with what the speaker presupposes, the correct are those compatible with what she asserted, and the incorrect are those compatible with what she denied, my semantics also blocks Samaritan inferences in this case. I close by outlining an extension of the theory that delivers similar results for a version of the Samaritan paradox that arises for the interaction between necessity modals and *conditional* constructions. Just as with the original Samaritan inference pattern, the pattern ⌜If ϕ, then $\Box(\phi)$⌝ seems to be valid for epistemic readings of the embedded necessity modal, but invalid for deontic readings.

Chapter 5, "Supposition, Presupposition, and Trivalent Conditionals," explores the prospects of some ideas suggested by the solutions I develop to the Samaritan paradoxes in chapter 4. The theory of chapter 4 suggests that the parameters with respect to which modals are evaluated can make a three-fold distinction between worlds, and furthermore, that presuppositions and conditional antecedents are the linguistic resources we have for tracking one dimension of this distinction: which worlds are relevant or not. This suggests a unity between how we might model the semantic functions of presuppositions, on the one hand, and conditional antecedents (which we might think of as *suppositions*), on the other. In this chapter, I defend the coherence of a particularly strong way of modeling this unity: presuppositions and conditionals both denote trivalent truth functions which are neither-true-nor-false when their (pre-)suppositions fail. Because the idea of modeling presuppositions using a trivalent logic has been much discussed in the literature, I begin by outlining some less-well-known motivations for modeling indicative conditionals with the same tools. I then introduce and develop arguments for two serious objections to giving a trivalent semantics for both presuppositions and conditionals within the same theory: (i), that doing so predicts absurdly that conditionals presuppose their antecedents; and (ii), that doing so leaves us unable to explain why indicative conditionals appear to presuppose that their antecedents are compatible with the conversational common ground. I then argue that it is possible to defuse both of these worries by distinguishing between the different characteristic effects of two different types of assertion we may make using conditional and unconditional sentences. I develop a formal dynamic pragmatics that distinguishes between conditional and unconditional assertions, and show how the resulting theory addresses the two worries. I note a third worry about accounting for the different embedding behavior of conditionals and presupposition-carrying sentences, and extend my dynamic pragmatics in order to meet two of the most frequently discussed embeddings: within the scope of negation, and within the scope of modal operators.

I have written the chapters so that they may be read independently. Thus, my solution to each problem for underspecificity in modal contexts can be assessed largely on its own, without the interference of any theoretical assumptions made for other problems.

Chapter 2: Underspecifying Desires

Coco wants ice cream. In particular, she wants *chocolate* ice cream — she hates all other flavors. If, knowing this, I were to give her a scoop of *vanilla* ice cream, it would be entirely fair for Coco to complain that I have not satisfied her desire. Although Coco wants ice cream, not just any ice cream will do.

I want to draw attention to two features of this mundane story. First, it does not seem to follow from the fact that Coco wants ice cream that Coco has a desire that would be satisfied by *any* way of getting ice cream. In general, when A wants p, the truth of p may not suffice to satisfy any of A's desires.[1] Call any true 'want' ascription of the form $\ulcorner A$ wants $p\urcorner$, where the truth of p is necessary for satisfying some desire of A's, but insufficient for satisfying any desire of A's, *underspecific*.

The second feature was implicit: Coco's wanting *chocolate ice cream* appeared to suffice for the truth of 'Coco wants ice cream'. In other words, the inference from (2.1) to (2.2) seemed valid:

(2.1) Coco wants chocolate ice cream.

(2.2) ∴ Coco wants ice cream.

In general, I will argue, more specific 'want' ascriptions like (2.1) entail less specific ones like (2.2). Let's call the inference pattern, of which the inference from (2.1) to (2.2) is an instance, the *underspecific inference pattern*, without yet saying what it is, exactly.[2]

1. Throughout this chapter, I assume that 'want' always takes, at the level of logical form, a proposition-denoting complement: so, e.g., (2.2) has a logical form like 'Coco$_i$ wants (PRO$_i$ GETS ice cream)'. This is an assumption commonly made by both semanticists (see von Fintel (1999), Heim (1992), Levinson (2003), Villalta (2008)), and syntacticians (Harley (2012)). However, the arguments in this chapter are largely independent of this assumption: doing without it may require complicating them, but would not require altering their substance.

2. I say what the pattern is in §1.

As Fara (2013) points out, the first feature — the possibility of true, underspecific 'want' ascriptions — conflicts with what I will call the *Simple View* of the relationship between 'want' ascriptions and the desires we use them to talk about.[3] As I will understand it here, the Simple View comes in two parts. First, there is a principle governing how the contents of the complements of true 'want' ascriptions relate to the contents of desires:[4]

> *Identity*. If ⌜A wants p⌝ is true at a context c, then the agent denoted by A at c has a desire the content of which is *identical* to the proposition expressed by p at c.

Second, there is a principle saying what it is for a desire to have the content it does:

> *Content*. For a desire to have the content p is for it to be satisfied just in case p is true.[5]

The Simple View entails that no true 'want' ascription is underspecific. For suppose (2.2) is true. Then, by *Identity*, Coco has a desire with content identical to what is expressed by '[Coco gets] ice cream', in other words, the proposition *that Coco gets ice cream*.[6] By *Content*, it follows she has a desire satisfied just in case *that Coco gets ice cream* is true. In other words, there are no situations in which the complement of (2.2) is true but Coco's desire for ice cream is not satisfied. So, the Simple View ensures that (2.2) cannot be

3. Lycan (2012) and Grant and Phillips-Brown (2020) also defend underspecific 'want' ascriptions. See Stampe (1986) and Braun (2015) for defenses of the Simple View. When it is not argued for overtly, it is often implicit in discussions of propositional attitude reports in general (Fodor (1978) Searle (1983)) and desire in particular (e.g., McDaniel and Bradley (2008)).

4. *Identity* places a necessary condition on the truth of a 'want' ascription. Among the few philosophers who discuss principles like *Identity*, it is not always clear whether it is meant to be a *semantic* truth or some other kind (analytical, metaphysical) of necessary condition.

5. This principle is compatible with many ways of spelling out what it is for a desire to be *satisfied*: e.g. that an agent is no longer motivated to act so as to promote the truth of p, that an agent no longer has a reason to act so as to promote the truth of p, that an agent experiences pleasure, and so on.

6. Abstracting from issues involved in assigning propositions to tenseless clauses. If this complication bothers the reader, it might help to append, to the end of each 'want' ascription, a reference to a future time, e.g., the reader may take (2.2) as elliptical for:

(2.2′) Coco wants ice cream after dinner.

true and underspecific. The argument generalizes: the Simple View entails that *no* true 'want' ascription is underspecific. Thus, we can either accept the Simple View, or we can accept the possibility of true, underspecific 'want' ascriptions, but we cannot accept both.

The Simple View is attractive for many reasons. For one, it offers a straightforward account of how we do what we do with 'want' ascriptions. We often like to please, help, or benefit others, and satisfying a desire they have is typically a good way to do so. When someone tells me, $\ulcorner A$ wants $p\urcorner$, according to the Simple View, I understand what is necessary and sufficient for satisfying A's desire just by virtue of computing the proposition denoted by p in the context.

Unfortunately, I think things are more complicated than the Simple View pretends. I will argue in this chapter that we must accept the possibility of true, underspecific 'want' ascriptions, and thus reject the Simple View. As an alternative, I will propose a semantics for 'want' that allows for underspecificity.

The plan is as follows. In the first part of this chapter (§§1-2), I argue against *Identity* by appealing to the second feature of Coco's story: the validity of the underspecific inference pattern. I argue that the bulk of the linguistic evidence suggests the pattern is valid (§1), and that if the pattern is valid, we should think that true 'want' ascriptions can be underspecific (§2). Accepting the possibility of true, underspecific 'want' ascriptions means we must reject the Simple View, and in particular, *Identity*.[7] Doing so raises a question: what systematic relation *do* true 'want' ascriptions bear to the desires we use them to talk about? When I learn that A wants p, what can I conclude about what it would take to satisfy A's desire for p?

In the second part of this chapter (§3), I turn to the lexical semantics of 'want' for an answer to this question. I argue that existing theories do not have the resources to provide an adequate alternative to the Simple View. In §4, I propose an original account

7. In this chapter, I set aside the possibility of rejecting *Content* instead.

that does. My *Desire-Based Semantics* for 'want' departs from most others by giving truth conditions for 'want' ascriptions directly in terms of an agent's *desires*, rather than some function of her beliefs, credences, comparative preferences, or cardinal utility function. I defend this choice on both empirical and methodological grounds. Then, I show that the account offers a straightforward alternative to *Identity*, namely, a principle I call *Entailment*. I conclude in §5 by discussing the ramifications of *Entailment* for understanding our communicative practices involving 'want' ascriptions and clarifying some aspects of the account.

2.1 The Underspecific Inference Pattern

Consider the following examples of the underspecific inference pattern, drawn from Fara 2013:

(2.3) a. Fiona wants to catch a fish big enough to make a meal.

b. ∴ Fiona wants to catch a fish.

(2.4) a. Charlotte wants to have enough champagne to feel it go to her head.

b. ∴ Charlotte wants to have some champagne.

In both (2.3) and (2.4), the truth of the (a)-sentence seems sufficient for the truth of the (b)-sentence. Furthermore, it is easy to imagine contexts in which one might utter the underspecific (b)-sentence even though the only ways of satisfying the subject's relevant desire for fish or champagne, respectively, are more completely specified by the complement of the (a)-sentence.

One who denies the possibility of true, underspecific 'want' ascriptions might deny that these are genuine instances of the underspecific inference pattern. As Fara discusses, they might appeal to the mechanisms of quantifier domain restriction Stanley and Szabó (2000), or nominal domain restriction Stanley (2002), in order to suggest that a hearer of the argument in (2.3), say, would assign a logical form to (2.3b) whereby it is synonymous with the more specific (2.3a); roughly, something like, 'Fiona wants to catch a fish [*among those that are big enough to make a meal*]'. If that were correct, these examples would not show that more specific 'want' ascriptions entail less specific ones; rather, they would simply illustrate the mundane fact that if sufficient information is available in a context, a speaker need not be as *explicit* as she might otherwise.

For this reason, Fara (2013) also provides examples that do not involve underspecifying nominal or quantifier expressions. Removing *adverbial* modifiers is a source of equally compelling examples of the inference pattern:

(2.5) a. Heidi wants to wear high heels on her feet.

b. ∴ Heidi wants to wear high heels.

(2.6) a. Winona wants to win the race fairly.

b. ∴ Winona wants to win the race.

Here, again, the (a)-sentences appear to entail the (b)-sentences. And in these cases, the information left unspecified by the (b)-sentence could not attach implicitly to a nominal or quantifier expression.

Of course, it is not hard to think of further moves one might make to deny to deny that these adverbially underspecified examples are genuine. Perhaps like nouns or quantifiers, verbs also come with domains that can take implicit restrictions, or perhaps there is some other mechanism whereby 'to wear high heels' can just mean, in a context, more specifically *to wear high heels on one's feet*.[8]

Evidence against any version of this strategy comes from 'want' ascriptions to pluralities. For example, the following arguments appear valid:

(2.7) a. Coco wants chocolate ice cream.

b. Vanessa wants vanilla ice cream.

8. For example, as a result of *modulation*, forms of whicha re discussed in Travis (1994), Recanati (2010), and Del Pinal (2018).

c. ∴ Coco and Vanessa both want ice cream.

(2.8) a. Winona wants to win fairly.

b. Chelsea wants to win by cheating.

c. ∴ Winona and Chelsea both want to win.

(2.9) a. Democrats want to raise taxes on the rich.

b. Republicans want to raise taxes on the poor.

c. ∴ Both parties want to raise taxes.

Each of these arguments seems valid. But this cannot be because the less specific ascription in the conclusion is somehow interpreted more specifically. Take, as an example, (2.7). The fact that the conclusion appears to follow from the premises cannot be because we interpret it as implicitly ascribing a desire for *chocolate* or for *vanilla* ice cream. For then it would clearly not follow. Similarly for (2.8) and (2.9). If these arguments are valid, it can only be because the (c) sentence ascribes a less specific 'want' than the (a) and (b) sentences.

All of this data suggests that the underspecific inference pattern trades on 'want' being *upward monotonic*, or *upward entailing* (UE) in its complement position:[9]

An operator O is *Upward Entailing* iff:

If p entails q, then $\ulcorner O(p) \urcorner$ entails $\ulcorner O(q) \urcorner$.

As applied to 'want', this gives us:

If p entails q, then $\ulcorner A$ wants $p \urcorner$ entails $\ulcorner A$ wants $q \urcorner$.

9. Throughout the chapter, I use the usual linguistic semanticist's notion of entailment, rather than the logician's. One key difference is that the former notion allows the set of admissible models to be constrained, for example, by a set of meaning postulates. These added constraints make it so that, e.g., 'Bolt sprints' entails 'Bolt runs'. Of course, this leaves much else unsettled with respect to the notion of entailment in question. For now, let us think of it classically as truth-preservation (see von Fintel (1999) and Sharvit (2017) for alternatives). Also, I will abstract in this chapter from the much-discussed issues surrounding hyperintensionality in attitude contexts, since the issue I am concerned with is orthogonal to those questions (see Fara (2013) for some discussion). I touch on this and related issues in §5.

The underspecific inference pattern may thus be characterized as follows:

> *Underspecific Inference Pattern*. When p entails q, one may infer $\ulcorner A$ wants $q\urcorner$ from $\ulcorner A$ wants $p\urcorner$.

The hypothesis that the underspecific inference pattern is underwritten by 'want' being UE leads straightaway to another prediction. 'Want' is UE if and only if negated 'want' ascriptions are *downward entailing* (DE):

> An operator O is *Downward Entailing* (DE) iff:
>
> If p entails q, then $\ulcorner O(q)\urcorner$ entails $\ulcorner O(p)\urcorner$.

Applying this to negated 'want' ascriptions gives us:

> If p entails q, then $\ulcorner A$ does not want $q\urcorner$ entails $\ulcorner A$ does not want $p\urcorner$.

Indeed, this prediction seems to be borne out. The negation of the underspecific (2.2) seems to entail the negation of the more specific (2.1):

(2.10) a. Coco does not want ice cream.

b. ∴ Coco does not want chocolate ice cream.

Further examples can easily be multiplied. Besides the intuitive status of such examples, the DE property of negated 'want' ascriptions leads to a more systematic prediction. If negations of 'want' ascriptions are DE, then we should expect them to license *negative polarity items* (NPIs) within their scope. NPIs are expressions like 'any', 'ever', or 'at all' which are only felicitous in certain linguistic contexts. The paradigmatic context in which they are felicitous is in the scope of DE operators like negation — hence the name, '*negative* polarity items'. For example, compare the acceptability of the following (a) and (b) pairs:

(2.11) a. # I think I have any gourds.

b. I do not have any gourds.

(2.12) a. # I think I ever cook.

b. I do not ever cook.

In each case, the (a) sentence sounds extremely odd, but the (b) sentence is perfectly natural. The standard theory about the difference between the (a) and the (b) sentence in each pair is that the NPIs 'any', 'ever', and 'at all' are licensed within the scope of DE operators like negation, but not in UE contexts like 'think'.[10] Thus within the UE context of the (a) sentence, they are infelicitous, and within the DE context of the (b) sentence, they are felicitous.

We observe exactly this pattern within the complement of 'want':

(2.13) a. # I want to have any gourds.

b. I do not want to have any gourds.

(2.14) a. # I want to ever cook.

b. I do not want to ever cook.

The standard theory of NPIs thus confirms that negated 'want' ascriptions are DE. That can only be so if un-negated 'want' ascriptions are UE. Thus, the felicity of NPIs in negated 'want' ascriptions provides linguistic evidence, independent of our intuitions about particular cases, that 'want' is UE.

Given these considerations, it is perhaps surprising that most existing semantic accounts of 'want' (e.g. Heim (1992), Levinson (2003), Villalta (2008), Lassiter (2011), Jerzak (2019a), Phillips-Brown (2021), Blumberg and Hawthorne (forthcoming)) do not predict it to be UE, and thus invalidate the underspecific inference pattern. Many of these theorists have been impressed by a putative counterexample to the pattern due to Asher (1987).[11] The example goes as follows. Suppose the thrill of supersonic travel on the

10. This is the analysis of Ladusaw (1980), building on work by Fauconnier (1975). Of course, there are complications with this account, such as the felicity of NPIs in the context of questions: *Did you ever cook?* See Zwarts (1995) and Giannakidou (1999).

11. There are other proposed counterexamples in the literature (see, e.g., Levinson (2003)), but Asher's is by far the most commonly discussed.

Concorde is incredibly appealing to Nicholas, but the cost of a normal ticket would bankrupt him. According to Asher (1987), (2.15a) is true but (2.15b) is false in this situation:

(2.15) a. Nicholas wants a free trip on the Concorde.

b. ∴ Nicholas wants a trip on the Concorde.

If this assessment were correct, then 'want' would not be UE, and the underspecific inference pattern would be invalid. I do not think it is correct, however. First, if (2.15a) is true but (2.15b) is false, then *ceteris paribus*, we would expect the conjunction of (2.15a) with the negation of (2.15b) to be felicitous and, indeed, true. But it is infelicitous, no matter the order of the conjuncts:

(2.16) a. # Nicholas wants a free trip on the Concorde but he doesn't want a trip on the Concorde.

b. # Nicholas doesn't want a trip on the Concorde but he wants a free trip on the Concorde.

To my ear, these sentences sound like outright contradictions. If 'want' is UE, then there is a simple explanation for why: they just *are* contradictions. Those who agree with Asher's assessment of Nicholas' case incur the burden of explaining their infelicity in other ways.

Of course, I do not mean to deny that acceptable versions of the sentences in (2.16) may be derived by adding stress, for example:

(2.17) Nicholas doesn't want a *trip* on the Concorde, he wants a *free trip* on the Concorde.

But the acceptability of (2.17) does not show that (2.16) is not a contradiction. (2.17) patterns with cases of 'scalar' or 'meta-linguistic' negation (Horn (1989), Carston (1996)). Compare, for example:

(2.18) a. It's not *red*, it's *crimson*.

b. It's not *good*, it's *extremely good*.

The stress in these sentences seems to perform the same function as in (2.17): it makes acceptable a sentence that would otherwise be a contradiction. I do not want to take a stand in this chapter on what the correct account of this phenomenon is. But it is widely assumed that these examples do not refute the theory that *being crimson* entails *being red* or that *being extremely good* entails *being good*. Usually, examples like these can be adequately paraphrased by explicitly adding exclusivity constructions like 'just', 'only', or 'merely' Coppock and Beaver (2013):

(2.18′) a. It's not *just red*, it's *crimson*.

b. It's not *merely good*, it's *extremely good*.

Similarly, the acceptable reading of (2.17) can be paraphrased by:

(2.17′) Nicholas doesn't want *just a trip* on the Concorde, he wants a *free* trip on the Concorde.

Clearly (2.17′) is not equivalent to (2.16). If (2.17′) is equivalent to (2.17), as I am suggesting, then the acceptability of the stressed (2.17) does not show that the non-equivalent (2.16) is not a contradiction.

Interestingly, this analysis of conjunctions like (2.17) also explains why Asher feels the need to add the modifier 'simpliciter' to state his conclusion about Nicholas' case. "If I want to ride on the Concorde and not pay for it," Asher says, "it doesn't necessarily follow that I also want to ride on the Concorde *simpliciter*" (Asher 1987, 171). 'Simpliciter' clearly functions as an exclusive (like 'simply'), and does some substantive work in making Asher's statement of the conclusion coherent. If that is right, then it is obvious why, if I want to ride on the Concorde and not pay for it, it does not follow that I want to ride on the Concorde *simpliciter* — for clearly I want something else, too: to not pay for my trip on the Concorde.

Besides conjunctions like (2.16), the theory that 'want' is not UE faces related problems in other embedded contexts. If (2.15a) and (2.15b) are logically independent, then their disjunction, *ceteris paribus*, should sound perfectly ordinary. If on the other hand, 'want' is UE and (2.15a) entails (2.15b), then their disjunction should sound redundant, since one disjunct entails the other.[12] Indeed, the latter is what we find (no matter the order):

(2.19) a. # Either Nicholas wants a free trip on the Concorde, or he wants a trip on the Concorde.

b. # Either Nicholas wants a trip on the Concorde, or he wants a free trip on the Concorde.

Again, if 'want' is UE, there is a simple explanation: they sound redundant because they *are* redundant. Each is truth-conditionally equivalent to the weakest disjunct, (2.15b).[13]

Similarly, if 'want' is not UE and (2.15a) does not entail (2.15b), one would expect the following conditionals to sound on a par with respect to their potential informativeness:

(2.20) a. If Nicholas wants a free trip on the Concorde, he wants a trip on the Concorde.

b. If Nicholas wants a trip on the Concorde, he wants a free trip on the Concorde.

But only (2.20b) sounds informative; (2.20a) sounds trivial. If 'want' is UE, then this is because (2.20a) *is* trivial (since the antecedent semantically entails the consequent); and (2.20b) *is* informative (since it is possible for the antecedent to be true and the consequent false).

Of course, further problems for the non-UE view can be derived wherever an inference pattern trades on an upward or downward entailing context. Agreeing with Asher's analysis of (2.15) and the conclusion that 'want' is not UE means giving up the

12. The norm against redundant disjuncts is known as *Hurford's Constraint*. See Hurford (1974).

13. Again, this is not to deny that, as with (2.16), there may exist repair strategies available for interpreters of these disjunctions. See, for example, Ciardelli and Roelofsen (2017a).

obvious explanation for all of this data. We should only do so if there is no other natural way to explain the intuition that some people have that (2.15a) may be true and (2.15b) may be false.

In fact, as von Fintel (1999) argues, the phenomenon of implicit quantifier domain restriction offers a plausible alternative explanation for this intuition. One thing that is distinctive about Asher's example is that *free* trips on the Concorde are extremely rare — perhaps unheard of. Thus, a typical utterance of (2.15b) will be made in a context in which it is presumed that the only trips worth talking about are the expensive ones. In such a context, the domain of the indefinite 'a trip' in (2.15b) is plausibly restricted to quantify over those, none of which is a trip Nicholas would like to take. Thus, holding fixed the facts about Nicholas' desires, (2.15b) would be false in a normal context.[14] When considering (2.15a), however, we identify a domain for the indefinite quantifier that includes possible free trips and, knowing Nicholas, know it to be true. In each case, the default context we are tempted to interpret the sentence in delivers Asher's intuitions. But this does not mean that 'want' is not UE. The theory that 'want' is UE says that *in the context where* (2.15a) is true, (2.15b) must also be true. It does not require that if (2.15a) is true in *some* context, (2.15b) is true in *every* context.

If the intuitions about Asher's Concorde example can be explained in this or some other way, then the main motivation for denying that the underspecific inference pattern is valid is rather weak. I think the bulk of the linguistic evidence — from plural ascription (2.7-2.8), NPI distribution (2.13), and judgments about semantic properties like contradictoriness, redundancy, and tautologousness (2.16-2.20b) — suggests that 'want' is UE, and that the underspecific inference pattern is therefore valid.

14. Likewise, even if Coco wants chocolate ice cream and hates all other flavors, in a context where it is presumed that there is no chocolate ice cream available, 'Coco wants ice cream' may be false.

2.2 The Openness of Desire

Despite the suggestive name I've given to the inference pattern, its validity alone does not entail that true 'want' ascriptions can be underspecific. To see why, suppose (2.1) is true and exactly specifies the content of one of Coco's desires. If (2.1) entails (2.2), then (2.2) is true as well. Even so, (2.2) would underspecify Coco's desire for chocolate ice cream only if Coco did not *also* have a desire with the less specific content *that she gets ice cream*. For if she did have such a desire, then (2.2) would *exactly* specify it. This suggests the following way of holding on to the Simple View: we concede that 'want' is UE, but we also accept the following principle:[15]

> *Desire Closure*. If p entails q and A has a desire with the content p, then necessarily, A has a desire with the content q.

Notice that if we accept *Content*, the principle identifying the contents of desires with their satisfaction conditions, we can rephrase *Desire Closure* as follows:

> *Desire Closure* (equivalent). If p entails q and A has a desire satisfied just in case p is true, then necessarily, A has a desire satisfied just in case q is true.

Accepting *Desire Closure* would mean that, for every less specific ascription made true by virtue of the underspecific inference pattern, there is also a desire with a less specific content for the ascription to *exactly* specify. Thus, if we accept *Desire Closure*, we may accept the validity of the inference pattern and still deny that 'want' ascriptions can be true and underspecific.

Although it is not usually discussed directly, most contemporary analyses of desire do not entail *Desire Closure*. For some (harmlessly) simplified examples:[16]

15. Note that this thesis is distinct from the thesis I endorse (but will not argue for here) that the ordinary English verb 'desire' is upward entailing with respect to its propositional complement, just like 'want'. So, for example, 'Coco desires that she get chocolate ice cream' entails 'Coco desires that she get ice cream'. *Desire Closure* uses a slightly technical sense of the ordinary English 'desire' that is common in philosophy of mind and ethics. This is I hope clear enough from the locution ⌜desire with the content ϕ⌝ (a non-UE operator, as I argue in this section), which I take to be philosophical jargon rather than ordinary English.

16. See Schroeder (2017) for a helpful survey.

Metaphysics of Desire. S has a desire with the content p just in case:

- *Disposition to Action*: S is disposed to take whatever actions are likely to bring about p, given S's beliefs.[17]
- *Pleasure / Displeasure*: S experiences pleasure when it seems to S that p / displeasure when it seems to S that not-p.[18]
- *Guise of the Good*: p appears good to S.[19]
- *Expected Utility*: S assigns a greater expected utility to p than to not-p.[20]

While probably few theorists actually subscribe to these analyses in simple form, more precise versions of these views are accepted and they usually share the same relevant feature: they do not entail *Desire Closure*.

Let me briefly illustrate, in each case, why. Suppose again that Coco has a desire with the content *that she gets chocolate ice cream*, and she hates vanilla. Suppose furthermore that vanilla ice cream is widely available in her immediate environment, but chocolate is hard to get, and Coco knows all of this. Then there are many actions that are likely to bring about *that Coco gets ice cream* which Coco is not disposed to take (say, reaching for the nearby vanilla), so on the *Disposition to Action* view, she does not have a desire with the content *that she gets ice cream*.

Alternatively, imagine in the described scenario that Coco is told that she will get ice cream. Since she thinks it is likely to be vanilla (without assuming it definitely *is*),

17. Stalnaker (1984) and Smith (1987, 1994) defend dispositional views along these lines. See Ashwell (2009, 2014) for discussion of and response to some common challenges for a dispositional analysis. See also Schroeder (2007), which defends the following analysis of S's having a desire with the content p: S is in a psychological state grounding the following disposition: when for some action a and proposition r believed by S, given S's beliefs r obviously helps to explain why S's doing a promotes p, S finds r salient, and this tends to prompt S to do a and S's attention is directed toward considerations like r.

18. The connection between desire and pleasure, displeasure, and the reward system is emphasized in, for example, the accounts of Schroeder (2004) and Sinhababu (2009).

19. See Tenenbaum (2013) for some discussion.

20. Arguably, something like this analysis of desire is implicit in some recent semantic work on 'want' (see Levinson (2003) and Jerzak (2019a), for example). Phillips-Brown (2021) defends a more complex version of an expected utility account, more explicitly directed at analyzing the folk psychological concept of desire.

she experiences no pleasure in its seeming to her merely *that she gets ice cream*. Thus, on the *Pleasure* analysis, she does not have a desire with the content *that she gets ice cream*. In fact, she may see getting ice cream as an ill: now she will have to find a way to get dispense with her ice cream. So getting ice cream does not seem good to her. Thus, on the *Guise of the Good* analysis, Coco does not have a desire with the content *that she gets ice cream*.

Finally, let's turn to the *Expected Utility Analysis*. For simplicity, imagine that the space of possible outcomes is partitioned by the following three options, by Coco's lights: (i) the set of worlds where she gets no ice cream ($\neg\textsf{ice}$); (ii) the set where she gets chocolate ice cream ($\textsf{choc}$); and (iii) the set where she gets vanilla ice cream ($\textsf{van}$). Then, imagine her credence function P and utility function U have the following structure over these possibilities:

$$P(\neg\textsf{ice}) = .5 = P(\textsf{ice}) \qquad\qquad U(\neg\textsf{ice}) = 0$$

$$P(\textsf{van}) = .45 \quad \text{and} \quad P(\textsf{van} \mid \textsf{ice}) = .9 \qquad\qquad U(\textsf{van}) = -1$$

$$P(\textsf{choc}) = .05 \quad \text{and} \quad P(\textsf{choc} \mid \textsf{ice}) = .1 \qquad\qquad U(\textsf{choc}) = 1$$

Then, Coco's expected utility function (EU), computed in the standard way, delivers the following values:

$$EU(\textsf{ice}) = P(\textsf{van} \mid \textsf{ice}))U(\textsf{van}) + P(\textsf{choc} \mid \textsf{ice})U(\textsf{choc}) \qquad EU(\neg\textsf{ice}) = U(\neg\textsf{ice})$$

$$= (.9 \times -1) + (.1 \times 1) \qquad = 0$$

$$= -.9 + .1$$

$$= -.8$$

$$EU(\neg\textsf{choc}) = P(\textsf{van} \mid \neg\textsf{choc})U(van) + P(\neg\textsf{ice} \mid \neg\textsf{choc})U(\neg\textsf{ice}) \qquad EU(\textsf{choc}) = U(\textsf{choc})$$

$$= (.47 \times -1) + (.53 \times 0) \qquad = 1$$

$$= -.47$$

This means Coco's expected utilities for the relevant propositions bear the following relations:

$$EU(\textsf{choc}) = 1 > -.47 = EU(\neg\textsf{choc})$$

$$EU(\textsf{ice}) = -.8 \not> 0 = EU(\neg\textsf{ice})$$

On the *Expected Utility* analysis of desire, then, Coco has a desire with the content *that she gets only chocolate ice cream* (choc) (since $1 > -.47$), but not a desire with the content *that she gets ice cream* (since $-.8 \not> 0$). In fact, it predicts the opposite: that Coco has a desire with the content *that she does not get ice cream* (since $0 > -.8$).

In sum, most contemporary theories of the metaphysics of desire do not entail *Desire Closure*. In light of these considerations, I will take *Desire Closure* to be false. Thus, I will endorse the following thesis, which I take to be a widespread commitment among metaphysical analyses of desire:

> *Desire Openness*. If p entails q and A has a desire with the content p, then it is possible for A to not have a desire with the content q.

Or, again, given *Content*:

> *Desire Openness* (equivalent). If p entails q and A has a desire satisfied just in case p is true, then it is possible for A to not have a desire satisfied just in case q is true.

Desire Openness entails, for example, that it is possible for Coco to have a desire with the content *that she gets chocolate ice cream and no other flavors* without having a desire with the content *that she gets ice cream*. In contexts like that, the sentence (2.1), 'Coco wants chocolate ice cream', will presumably be true. But, if the underspecific inference pattern is valid, (2.1) entails (2.2), 'Coco wants ice cream'. So, (2.2) is also true in this situation, even though Coco has no desire with the content *that she gets ice cream*. Thus, (2.2) is true but does not exactly specify any of Coco's desires: while the truth of the complement clause is necessary, it is not also sufficient for the satisfaction of her desire with the content *that she gets chocolate ice cream*. Thus, (2.2) is true and underspecific. In general, given *Desire Openness*, the validity of the underspecific inference pattern entails that 'want' ascriptions can be true and underspecific.

This concludes my argument for the possibility of true underspecific 'want' ascriptions. In the last section, I argued that the underspecific inference pattern is valid. But as I noted, the validity of the pattern does not by itself mean that true 'want' ascriptions can be underspecific. As I have just argued, a second assumption is required, namely, *Desire Openness*. Given *Desire Openness*, the validity of the underspecific inference pattern does entail that true 'want' ascriptions can be underspecific. If I am right about the validity of the pattern and *Desire Openness*, then we should reject the Simple View, and in particular *Identity*, since it requires 'want' ascriptions to *exactly* specify the contents of desires.

2.3 Desire and von Fintel's 'want'

If true 'want' ascriptions can be underspecific, this raises interesting questions about how we are able to do all of the important things we do with them. What can I conclude about how to satisfy Coco's desire for ice cream when I learn that she wants ice cream?

In trying to answer this question, a natural place to look is the truth-conditional semantics of 'want', for the question we are asking is, what conditions does the truth of ⌜*A* wants *p*⌝ place on the content of *A*'s desire for *p*? As mentioned in §1, most existing theories do not validate the underspecific inference pattern, so they are non-starters for our purposes. But there is one viable proposal that validates it, namely the semantics of von Fintel (1999). Let us consider whether von Fintel's semantics can help provide an alternative to *Identity* that illuminates our practices surrounding 'want' ascriptions.

Von Fintel's account of 'want' is inspired by the well-known semantics for non-attitude modals developed in Kratzer (1977, 1981, 1991).[21] In particular, on von Fintel's account, the truth of a 'want' ascription like ⌜*A* wants *p*⌝ depends on two factors (simplifying slightly):[22]

- The set of worlds compatible with everything *A* believes, which I write as Bel_A.[23]
- A set of propositions (sets of worlds) that "[specifies] the preferences" of *A*, which I will write as Pref_A (von Fintel 1999, 118).

With some (harmless) simplification, von Fintel's semantics for 'want' is roughly as follows:[24]

> ⌜*A* wants *p*⌝ is true at *c* iff the proposition expressed by *p* at *c* is true at every world in Bel_A that is maximally preferred given Pref_A.

A world w_1 is strictly preferred to w_2 ($w_1 <_{\mathsf{Pref}_A} w_2$), just in case every preferred proposition (element of Pref_A) that is true of w_2 is also true of w_1, and at least one preferred

21. To briefly summarize: modals are sensitive to two contextual parameters that supply sets of propositions. One parameter, the *modal base*, is meant to supply a domain of relevant worlds, while the other, the *ordering source*, determines a *ranking* of those worlds. Modals act like quantifiers over the undominated (relative to the ranking) elements of the modal base. As I explain, in the case of 'want', the modal base is a set of the subject's beliefs, and the ordering source is a set of propositions expressing her preferences.

22. There are a lot of questions one might ask about the natures of these two entities. Since von Fintel's essay is not focused on 'want', in particular, he does not pursue them in much detail. His primary goal in that paper, when it comes to 'want', is to outline a plausible semantics that validates the underspecific inference pattern, and to defend it against cases like Asher's Concorde example.

23. Strictly speaking, it is a larger set of worlds, consisting of those worlds compatible with everything *A* believes to be the case no matter how *A* chooses to act.

24. I ignore here the definedness conditions meant to explain the presupposition behavior of 'want'.

proposition true of w_1 is not also true of w_2. More precisely, von Fintel defines the strict partial order determined by Pref_A as follows:

$$w_1 <_{\mathsf{Pref}_A} w_2 := \forall p \in \mathsf{Pref}_A : \ w_2 \in p \Rightarrow w_1 \in p \text{ and}$$
$$\exists q \in \mathsf{Pref}_A : \ w_1 \in q \ \& \ w_2 \notin q$$

Given an arbitrary set of worlds S, the *maximally* preferred ones ($\mathsf{max}_{\mathsf{Pref}_A}(S)$) are those in S for which there is no strictly more preferable world also in S:

$$\mathsf{max}_{\mathsf{Pref}_A}(S) = \{w \mid \neg\exists v : \ v \in S \ \& \ v <_{\mathsf{Pref}_A} w\}$$

Thus, von Fintel's truth conditions are, more precisely:

$$[\![\ulcorner A \text{ wants } p\urcorner]\!]^{c,w} = \textsc{true} \text{ iff } \mathsf{max}_{\mathsf{Pref}_{[\![A]\!]^{c,w}}}(\mathsf{Bel}_{[\![A]\!]^{c,w}}) \subseteq \{v \mid [\![p]\!]^{c,v} = \textsc{true}\}$$

This semantics has many virtues. Foremost from our perspective, it makes 'want' UE and validates the underspecific inference pattern. If p is true in every maximally preferred world, and p entails q, q must also be true in every maximally preferred world.

Despite this, von Fintel's semantics does not obviously help much with our current concern: identifying what one can, as a rule, conclude about the content of A's desire for p, given that $\ulcorner A$ wants $p\urcorner$ is true. To begin with, von Fintel's semantics makes no mention of A's desires. Furthermore, it is not clear how we could interpret or extend the framework in order to systematically relate the content of A's desire for p to the truth conditions of $\ulcorner A$ wants $p\urcorner$. Consider what is likely the most obvious way of trying to do so. If $\ulcorner A$ wants $p\urcorner$ is true, on von Fintel's semantics, that is because there is a certain specific proposition that entails p and is especially preferred by A, namely, the proposition determined by the set of A's maximally preferred belief-worlds ($\mathsf{max}_{\mathsf{Pref}_A}(\mathsf{Bel}_A)$). One might therefore suggest that we think of *this* proposition as forming the specific content of A's desire for p. Doing so would deliver the following alternative to *Identity*:

Maximal Preference. If ⌜A wants p⌝ is true at a context c, then A has a desire for p with the (potentially more specific) content, $\mathsf{max}_{\mathsf{Pref}_{[\![A]\!]^c}}(\mathsf{Bel}_{[\![A]\!]^c})$.

But there are some clear problems with this approach. For one, it makes the contents of an agent's desires too narrow. Suppose I prefer to get chocolate ice cream, I prefer that the U.S. Senate pass a minimum wage increase, and I think that both things are live possibilities. Then the worlds I most prefer are those where I get chocolate ice cream and the U.S. Senate passes a minimum wage increase. If we accept *Maximal Preference*, it follows that my desire for chocolate ice cream has as part of its content that the U.S. Senate passes a minimum wage increase. Given *Content*, it would follow that my desire for chocolate ice cream cannot be satisfied unless a minimum wage increase is passed.

This points to an even worse problem: *Maximal Preference* entails the absurd result that all of an agent's desires have the same content. In particular, for any agent A and any true 'want' ascription ⌜A wants p⌝, A's desire for p has the content $\mathsf{max}_{\mathsf{Pref}_A}(\mathsf{Bel}_A)$. This means that A's desires all stand to be satisfied or frustrated together. But ordinarily, if I want ice cream and I also want my political ideals to be realized, I will have two desires that can be satisfied or frustrated independently of each other.[25]

Perhaps there are other candidates, besides *Maximal Preference*, made available by the resources of von Fintel's semantics, but I will not explore them further here. Instead, I want to explore an alternative approach. What if, rather than trying to fit desires into a semantics for 'want' that has no clear place for them, we try to give a semantics for 'want' directly in terms of desires? As I will argue in the next section, doing so will provide a more plausible alternative to *Identity*, and comes with empirical and methodological advantages.

25. *Ordinarily*, since we must allow for the possibility that one's political ideals just are the getting of ice cream.

2.4 A Desire-Based Semantics for 'want'

Let us assume that for any agent A, there is a set of propositions (which I will model as sets of worlds) that exactly specify the satisfaction conditions of A's desires. Call this set Des_A. So, for example, if Coco has a desire that is satisfied just in case she gets chocolate and no other flavor of ice cream, and another desire that is satisfied just in case the U.S. minimum wage is \$15/hour, we have:

$$\mathsf{Des}_{\text{Coco}} = \{\ \{w \mid \textit{Coco gets chocolate and no other flavor of ice cream in } w\},$$
$$\{w \mid \textit{U.S. minimum wage is \$15/hour in } w\}\ \}$$

Of course, there are many interesting and reasonable questions one might ask about the nature of the set Des_A.[26] I take it that the existence of such a set is uncontroversial enough, at least for any theorist who accepts an ontology of desires and *Content*. In order to remain ecumenical about various theories of desire, I will try to make as few assumptions as possible. The only assumption I will definitely make is that, given *Desire Openness*, it is not as a rule closed under entailment. This means that my account is compatible with any view of the metaphysics of desires (for example, any of those discussed in §2) that accepts *Desire Openness*.

Given this resource, we can give a semantics for 'want' that validates the underspecific inference pattern and avoids many of the problems we found for von Fintel's account. We will say that $\ulcorner A$ wants $p\urcorner$ is true just in case the satisfaction of one of A's desires entails the truth of p. In other words, that there is an element d of Des_A such that d entails p:[27]

26. For example, which facts determine whether or not a proposition is in Des_A? Are there structurally interesting properties of the set such as closure under some function? Does Des_A bear inclusion or exclusion relations to certain obvious 'cousins', like the set of propositions giving the fulfillment conditions of A's intentions, fears, hopes, or wishes?

27. Of course, like other possible worlds semantics for modals and attitude verbs, this semantics will give rise to the well-known puzzle of 'logical omniscience' (see, e.g., Stalnaker (1988)). I abstract from this and related problems in this chapter.

Desire-Based Semantics:

$$[\![\ulcorner A \text{ wants } p\urcorner]\!]^{c,w} = \textsc{true} \text{ iff } \exists d \in \mathsf{Des}_{[\![A]\!]^{c,w}} : \; d \subseteq \{v \mid [\![p]\!]^{c,v} = \textsc{true}\}$$

It is straightforward to see that this semantics makes 'want' UE, and thus generates the validity of the underspecific inference pattern. If the satisfaction of one of Coco's desires entails that *Coco gets chocolate ice cream*, then it also entails that *Coco gets ice cream*. Similarly, negated 'want' ascriptions are DE. If Coco does not have a desire the satisfaction of which entails that *she gets ice cream*, it follows that she also does not have a desire the satisfaction of which entails that *she gets chocolate ice cream*. Thus, the felicity of NPIs in the scope of negated 'want' ascriptions, as well as the rest of the data from §1, can be explained in the usual way.

Furthermore, the present semantics vindicates the other feature of Coco's story: if she has a desire satisfied just in case *she gets chocolate ice cream and no other flavors*, then it follows that (2.1) and (2.2) are true by virtue of this desire. In cases where she does not also have a desire with the less specific content *that she gets ice cream*, then her desire for ice cream would not be satisfied by just *any* way of getting ice cream. Thus, the *Desire-Based Semantics* readily accounts for the phenomenon that this chapter is centrally concerned with: that 'want' ascriptions can be true and underspecific.

Clearly, the semantics also underwrites a straightforward alternative to *Identity*:

> *Entailment*. If $\ulcorner A$ wants $p\urcorner$ is true at a context c, then A's desire for p has a (potentially more specific) content that entails the proposition expressed by p at c.

Thus, when I learn that 'Coco wants ice cream' is true, I learn that there is *some way* for Coco to get ice cream that would satisfy her desire, but I cannot conclude that *any way* of getting ice cream would do so.

Clearly, *Entailment* is weaker than *Identity*. This means that the content of a true 'want' ascription is, on my view, less informative about the conditions of satisfaction of the agent's desire than it would be if the Simple View were correct. As I discuss below, this means that successful coordination in satisfying each other's desires essentially involves background reasoning over-and-above the reasoning involved in semantic interpretation.

I suspect that a semanticist might, at this point, raise an objection to the *Desire-Based Semantics*. Both it and von Fintel's semantics make 'want' UE, and can therefore explain all of the linguistic data adduced in §1. The move from von Fintel's semantics to the *Desire-Based Semantics* might seem to unfairly inject extra-linguistic philosophical considerations about desires and their satisfaction conditions into our semantic theory of 'want'.

In response, I want to argue that the *Desire-Based Semantics* for 'want' comes with some empirical and methodological benefits for the semanticist. First, let me discuss the empirical benefits. In contrast to von Fintel's semantics, the *Desire-Based Semantics* makes 'want' a non-normal modal operator.[28] This frees 'want' from some of the baggage that comes with normal necessity modals.

For example, most metaphysical accounts of desire allow that the contents of our desires are not closed under conjunction: I may have desires with the contents p and q, respectively, without having a desire with the content $p \wedge q$. If we adopt this assumption, then our semantics for 'want' will not validate conjunction *agglomeration*, as illustrated in the following example:

(2.21) a. Alicia wants to go to Kerala this summer.

b. Alicia wants to go to Costa Rica this summer.

c. # Thus, Alicia wants to go to both Kerala and Costa Rica this summer.

Suppose it is the end of winter, and Alicia begins planning her summer vacation. She has two weeks off from a job she loves, so she plans to take one trip this year. As she considers her options, she finds herself with the desire to go to Kerala, and with the

28. In particular, it gives 'want' a *neighborhood semantics*. See Pacuit (2017) for an excellent overview. My semantics treats 'want' as the *diamond-box* modal ('$\langle\,]$') from (9).

desire to go to Costa Rica. She understands and is unbothered by the fact that she can only choose one; she simply has not yet decided which. This is the sort of ordinary situation in which (2.21a)-(2.21b) may be true while (2.21c) false. On a normal modal semantics like von Fintel's, by contrast, (2.21a) and (2.21b) entail (2.21c), so it would have to be true as well.

Second, using a non-normal modal semantics means we straightforwardly allow for the ascription of *necessarily conflicting* desires.[29] Suppose Alicia wants to stay at the party in order to keep talking with Coco, but she also wants to leave the party in order to get a good night's rest. She cannot decide which desire to act upon. It is natural enough to say something like the following:

(2.22) Alicia cannot decide what to do; she both wants to leave and wants to stay.

On my non-normal modal semantics, this situation can be modeled simply by supposing that Alicia has two separate and necessarily incompatible desires: one to leave and one to stay. On a normal modal semantics like von Fintel's, we could only account for this by further complicating the pragmatics associated with the theory, for example, by positing a contextual shift in which preferences or beliefs are relevant for determining the meanings of the two instances of 'want' in (2.22).[30]

Third, we have not required, as von Fintel's semantics does, that the worlds relevant for evaluating a 'want' ascription are a subset of the the worlds compatible with the agent's beliefs.[31] This generates a desirable measure of independence between the

29. See Marino (2009). Some of the simple theories of desire discussed in §2 do not allow for necessarily conflicting desires. But many more precise versions do. For example, see Ashwell (2009, 2014) for a version of the *Dispositional Analysis* that does so.

30. Indeed, Levinson (2003) appeals to context-shift (in the selection of which preferences are relevant) to explain such data. Although I will not argue this here, I think Levinson's context-shifting proposal needs to be constrained to prevent over-generation (being able to select *any* preferences at all is too permissible). My semantics, by contrast, is already tightly constrained: it generates the truth of conflicting 'want' ascriptions exactly when the agent has conflicting desires.

31. This does not mean I am assuming there is *no* relationship at all between the agent's belief state and her desires. If we suppose, for example, that desire's constitutive aim is at the *accessible* (by the agent's lights) Velleman (2014), then a natural way to spell this out would be to require that $p \in \mathsf{Des}_A$ only if $p \cap \mathsf{Bel}_A \neq \varnothing$. This relation is weaker than the one posited by von Fintel's semantics, and would not lead to the problem I discuss here. See also Drucker (2017) for reasons to think the contents of an agent's

'belief' and 'want' ascriptions true of an agent. For example, the *Desire-Based Semantics* allows for the ascription of *contingently conflicting* desires — cases where one's desires are in conflict *given what one believes*. On von Fintel's semantics, the following three sentences are jointly inconsistent:

(2.23) a. Alicia wants to get tenure.

b. Alicia believes that she will get tenure only if she stops caring about her personal life.

c. Alicia wants to continue caring about her personal life.

If (2.23a) is true, it follows on von Fintel's semantics that Alicia's maximally preferred belief-worlds are ones where she gets tenure. If (2.23b) is also true, these are all worlds where she stops caring about her personal life. Therefore, in all of Alicia's maximally preferred belief worlds, she stops caring about her personal life, so on von Fintel's semantics, (2.24) must also be true:

(2.24) Alicia wants to stop caring about her personal life.

Even though Alicia clearly faces some form of tension between her desires to get tenure and to care about her personal life (by her lights, she can only satisfy one of them), I think it is obvious that (2.24) is not a semantic consequence of (2.23a) and (2.23b) (Kant's doctrine about willing the necessary means to one's ends aside). On a normal modal semantics like von Fintel's, the truth of (2.24) entails the falsity of (2.23c), so the three sentences in (2.23) are predicted to be jointly inconsistent. On my semantics, by contrast, they are consistent.

Each of these cases suggests, I think, that the *Desire-Based Semantics* does a better job than its main UE-competitor of explaining the empirical data surrounding one of the central uses of 'want' ascriptions in ordinary discourse: to lay out the respects in which an agent's values are in conflict with one another.

---

desires are not as tightly constrained by their beliefs or credences as semantics like von Fintel's assumes.

Now let me turn to some reasons to think a *Desire-Based Semantics* is also methodologically preferable as a semantics for 'want'. Generally, the other existing semantics for 'want' of which I am aware (including not just von Fintel (1999) but also non-UE semantics like Heim (1992), Levinson (2003), Villalta (2008), Lassiter (2011), Jerzak (2019a), Phillips-Brown (2021), and Blumberg and Hawthorne (forthcoming[c])), make the truth conditions of a 'want' ascription some function of an information state or probability function, on the one hand, and the agent's preferences or utility function, on the other. Each of these semantic theories is therefore committed to a straightforward reduction of facts about what someone wants to some relation between these factors. In von Fintel's case, for example, wants are determined by preference among doxastic possibilities. An account like Levinson's, on the other hand, says that wanting p is just a matter of assigning a higher subjective expected utility to p than to $\neg p$.

I have some sympathy for the philosophical motivations behind such accounts. It is plausible enough that wanting is not metaphysically fundamental, and that it might be grounded in some combination of more basic psychological factors. But research in many fields on the nature of desire suggests that the psychology of desire is probably more complex than these semantic accounts suppose. For example, besides preference and credence, the literature in behavioral economics and psychology suggests that other psychological factors might influence what an agent wants, such as their attitude toward risk,[32] their antecedently formed plans,[33] the limited supply of cognitive resources,[34] salience of various alternatives to the agent, or the ways in which a decision problem is 'framed'.[35] On my view, the semanticist would be better off if she can account for the linguistic data surrounding 'want' while avoiding commitment to a view on these matters.

This is not to say that my view *precludes* any form of analysis along these lines. Rather,

32. Buchak (2013)
33. Bratman (1987)
34. See, for example, the literature on bounded rationality Wheeler (2020).
35. See, for example, the classic work of Kahneman and Tversky (1979).

my account distinguishes between two versions of the analytical question we might be interested in:

- Metaphysical Question: what is it for p to be in the set Des_A?
- Semantic Question: what is it for ⌜A wants p⌝ to be true?

Distinguishing between these two questions means that while they have the option, semanticists interested in 'want' do not *have to* commit themselves to any particular answer to the metaphysical question — they may leave that task to philosophers of mind, psychologists, and others working on the nature of desire. The *Desire-Based Semantics* answers the semantic question but not the metaphysical one. It is intended to be compatible with *any* answer to the metaphysical question that does not entail *Desire Closure*.[36]

2.4.1 Levinson's example

The Desire-Based semantics for 'want' and the clear distinction it draws between the semantic and metaphysical questions offers us an appealing response to the worry Levinson raises for von Fintel's semantics that I introduced in the introduction (§1.2.2). Recall the structure of the example:

> A always prefers to get more money than less, and knows that they face the following situation. A fair coin will be flipped. If the coin lands heads, A gets \$200. If the coin lands tails, then a second event will occur (say, casting lots), with two outcomes with different chances: (Tails 1) A gets \$300 (10% probability); or (Tails 2) A gets nothing (90% probability).

36. On my view, many of the questions about the relationships between wanting, believing, preferring, and so on that some other semantic accounts of 'want' try to answer are better posed when they target the metaphysical question, rather than the semantic one. Of course, different analyses of the metaphysics of desire will lead to some different semantic predictions, as discussed in the case of conjunction agglomeration. While I do think that conjunction agglomeration fails for 'want', and thus that the proper metaphysics of desire should not predict that the contents of an agent's desires are necessarily closed under conjunction, I take it to be a minor virtue of my semantics for 'want' that it is compatible with the either view on the matter.

Recall that according to Levinson, the sentence (1.28), repeated here, is true, while von Fintel's semantics for 'want' seems to predict that it is false:

(1.28) *A* wants the coin to land heads.

In my opinion, Levinson's description of the example does not contain enough information to settle things either way about the truth value of (1.28). For example, if *A* is risk-averse, perhaps Levinson is right that (1.28) is true. If *A* is risk-seeking, on the other hand, perhaps it is false, since the possibility of getting $300 if the coin lands tails is most attractive and the possibility of losing out on $200 is untroubling.

Still, suppose Levinson is right that (1.28) is true precisely because *A*'s expected value for the coin landing heads is greater than their expected value for its not landing heads. We can make sense of this within my *Desire-Based* semantics for 'want' by assuming that *A* has a desire for p just in case the expected utility they assign to p is greater than that which they assign to $\neg p$. If that were so, then *A* clearly has a desire with the content *that the coin land heads*. Clearly, then, (1.28) is true.

I want to note that an advantage of this alternative form of explanation of Levinson's truth value judgment about (1.28) is that by separating the metaphysics of desire from the semantics of 'want', it allows that different agents might settle on their desires in different ways. The contents of *A*'s desires might be determined by her comparative standard expected utility assignments as Levinson suggests, but perhaps *B*'s desires are also affected by their causal assumptions, their risk tolerance, the framing of the decision problem, or other factors. Any semantics for 'want' that gives truth conditions in terms of a unique function of an agent's doxastic and preferential states seems to preclude that our use of 'want' might be compatible with these kinds of differences among people. My Desire-Based semantics, by contrast, does not.

2.5 Conclusion

If I am right that 'want' ascriptions can be true and underspecific, and that we should reject *Identity* in favor of the weaker principle *Entailment*, this has some important ramifications for understanding how we do what we do with 'want' ascriptions. If someone says, 'Coco wants ice cream' and I take them to be truthful, I can conclude that Coco's getting ice cream is *necessary* for the satisfaction of one of her desires, but not that it is also sufficient. Identifying sufficient conditions for the satisfaction of Coco's desire requires non-semantic knowledge about, for example, Coco's psychology and abilities, the normal course of events, or the context in which the utterance was made. I might know that Coco hates all flavors besides chocolate, and conclude that only *chocolate* would be sufficient. Alternatively, I might know very little about Coco's likes and dislikes, and rely on other resources to draw further information. If the speaker is answering the question *who would like to have the mint ice cream on the table?*, I might conclude Coco has a desire that would be satisfied by getting the mint ice cream in front of us. If the speaker is answering the question *why did Coco leave class early?*, I might assume very little about what specific flavors of ice cream would satisfy her desire. More generally, I might draw on common-sensical knowledge about how people normally like their ice cream. For example, I would likely assume Coco's desire is for, in particular, ice cream that is *unmelted*, and that *contains no large shards of broken glass*. Having rejected *Identity* in favor of the weaker *Entailment* does not mean it is mysterious how we can rely on true 'want' ascriptions to coordinate on actions that satisfy each other's desires. Rather, we need only accept that successful coordination of this kind essentially involves background reasoning over-and-above the semantic interpretation of 'want' ascriptions, or, in case such reasoning leaves things unsettled, further inquiry: what *kind* of ice cream do you want, Coco?

I want to close with some clarifications about the account I have developed in this chapter. First, I do not take the *Desire-Based Semantics* to be the full story when it comes to the truth conditional semantics of 'want' ascriptions. There is obviously some semantic data it does not explain. For example, 'want' ascriptions are gradable: one may want p *more* than one wants q. Ultimately, we need to refine the *Desire-Based Semantics* in order to account for this behavior. This is an issue that is crucial to the proper semantic account of 'want', but I ignore it here because I think it is unlikely that resolving it will help us further refine our alternative to the Simple View. On the other hand, there is another category of semantic data that I think may. Like many semantics for necessity modals,[37] the *Desire-Based Semantics* inherits several paradoxes that are well known from the literature on deontic logic, such as Ross's Puzzle (⌜A wants p⌝ entails ⌜A wants p or q⌝) and the Samaritan Paradox (⌜A wants p_q⌝ entails ⌜A wants q⌝ where p_q presupposes q). Accounting for this data is deeply important and may help us further refine *Entailment*. Solving them properly requires a greater level of generality, since these paradoxes arise not just for 'want', but also for many other modals and attitude verbs. I adopt such generality and turn to solving these puzzles in chapters 3 and 4.

Here, I hope to have made the case for two main points. First, that the semantic and philosophical considerations I drew upon in §§1-2 suggest that 'want' ascriptions can be true and underspecific. Second, that the *Desire-Based Semantics* — or some refinement of it — can account for this feature in a philosophically, empirically, and methodologically attractive manner.

37. For example, those of von Fintel (1999) and Kratzer (1977, 1991). See Portner (2018) for a survey.

Chapter 3: Independent Alternatives: Ross's Puzzle and Free Choice

3.1 Introduction

Orthodox semantic theories of natural language modals make the seemingly correct prediction that we can make true, underspecific statements about what is necessary or possible.[1] For example, if it is necessary that I mail the letter *overnight*, orthodox theories predict that my utterance of 'I ought to mail the letter' is true, even though this sentence is not fully specific about what I have to do. When combined with the standard theory of disjunction, however, this apparently correct prediction gives rise to two well-known puzzles.[2] First, there is *Ross's Puzzle* (Ross 1941), which is the puzzle of reconciling the fact that orthodox semantic theories predict the following inference to be valid with the appearance that it is *invalid*:[3]

(3.1) a. Alicia ought to mail the letter. $\Box M$

b. So, Alicia ought to either mail the letter or burn it. $\Box(M \vee B)$

I will call an inference that exhibits the pattern in (3.1) a *Ross inference*.

Second, there is the *Puzzle of Free Choice Permission* (von Wright 1968; Kamp 1973). Here, the challenge is to reconcile the apparent *validity* of the following inference with the fact that orthodox theories predict it to be invalid:

1. By 'orthodox', I mean any theory that makes necessity and possibility modals upward monotonic, e.g. the standard Kripke semantics that uses accessibility relations, and the context-sensitive semantics of Kratzer (2012c). 'Underspecific' has many different uses in the philosophy of language; the one I employ here has precedent in Zimmermann (2006) and Fara (2013).

2. By 'standard theory of disjunction', I mean especially theories like that of Partee and Rooth (1983), where disjunction is treated as the Boolean dual of conjunction.

3. In keeping with much of the recent literature, I use declarative 'ought' claims to illustrate Ross's puzzle. Ross's original example was presented using imperatives: where 'Post the letter!' entails 'Post the letter or burn it!' Although I do not discuss imperatives in this chapter, my account can be straightforwardly extended to them.

(3.2) a. Alicia may either mail the letter or burn it. $\Diamond(M \vee B)$

b. So, Alicia may burn the letter. $\Diamond B$

I will call an inference that exhibits the pattern in (3.2) a *free choice inference*.

There have been two kinds of solution to the puzzles. Pragmatic accounts defend the semantic predictions of the orthodox theories and give a pragmatic explanation of the apparent (in)validity in each case. Semantic accounts, by contrast, reject orthodoxy in favor of revisionary semantics for modals and/or disjunction in order to vindicate the appearance of (in)validity.

Both kinds of solution must explain *why* speakers' intuitions about the (in)validity of these arguments are in opposition to the predictions of the orthodox semantics. Since a speaker's judgment of the (in)validity of an argument depends on what meanings the speaker assigns to the premises and the conclusion, providing such an explanation involves taking a stand on what speakers intuitively take these sentences to mean. In fact, both semantic and pragmatic solutions to the puzzles have typically agreed on this point: over and above their orthodox truth conditions, sentences of the form $\Diamond(p \vee q)$ and $\Box(p \vee q)$ convey that p and q are each compatible with the set of relevant worlds. In this chapter, I argue that this analysis cannot explain the full range of data. Instead, I argue that claims of the form $\Diamond(p \vee q)$ and $\Box(p \vee q)$ convey something stronger: that p and q are each *independent* alternatives among the set of relevant worlds. I then develop an original revisionary semantics that validates these inferences, and thereby vindicates our unorthodox intuitions about both puzzles.

The plan is as follows. In the next section (§3.2), I explain why orthodox semantic theories make the counter-intuitive predictions that they do in our two puzzles, and in §3.3, I survey a range of different modals for which the two puzzles arise. In §3.4, I argue that the puzzles need to be explained in terms of what I call *Independence inferences*. In §3.5, I show that these inferences are validated when a simple topological relationship — a minimal covering relation — holds between the truth sets of the disjuncts and the set of relevant worlds over which a modal quantifies. In §3.6, I use this relation to outline an inquisitive semantics for modals that validates the Independence inferences, and I show how it vindicates our intuitions in the two puzzles. Then, in §3.7, I extend the model to a 'bilateral' version in order to generate better predictions for the interactions between modals and negation. Finally, in §3.8, I show that the bilateral theory generates some truth value gaps, and argue that these well-placed gaps shed new light on the original puzzles.

3.2 Orthodox Predictions

Orthodox semantic theories treat necessity modals like 'ought' and possibility modals like 'may' as universal and existential quantifiers, respectively, over sets of worlds.[4] Where $R(w)$ is the set of worlds relevant at the world w,[5] the orthodox theory of 'ought' ($\Box$) and 'may' ($\Diamond$) roughly amounts to the following, with $[\![\phi]\!]^w$ giving the truth value of ϕ at a world w (for simplicity, I ignore some syntactic complexity, the dependence of R on context, and semantic interactions with tense):

$$[\![\Box\phi]\!]^w = \textsc{true} \text{ iff for every } w' \in R(w) : [\![\phi]\!]^{w'} = \textsc{true}$$

$$[\![\Diamond\phi]\!]^w = \textsc{true} \text{ iff for some } w' \in R(w) : [\![\phi]\!]^{w'} = \textsc{true}$$

4. See Kripke (1963). In Kratzer (1977, 1991), the dominant orthodox theory of natural language modals, there are several worthwhile complications of the basic quantificational idea, but they do not matter for our purposes. See Portner (2009) for a textbook presentation of various semantic theories of natural language modals.

5. In the Kratzerian dialect (Kratzer 1977, 1991), where modals are relative to a *modal base* f and an *ordering source* g, we define $R(w)$ as follows:

$$R(w) = \mathsf{max}_{g(w)}(\bigwedge f(w))$$

where $\mathsf{max}_{g(w)}$ is a function that takes a proposition and returns the subset of worlds that are maximal with respect to the order determined by g.

In the Kripkean dialect, where modals are sensitive to an *accessibility relation* R between worlds in a set W, we define $R(w)$ as follows:

$$R(w) = \{v \in W \mid (w, v) \in \mathsf{R}\}$$

That is, $\Box(\phi)$ is true iff ϕ is true at every relevant world, while $\Diamond(\phi)$ is true iff ϕ is true at at least one relevant world. As a result, orthodox semantic theories make modals like 'ought' and 'may' *upward monotonic* operators:

> For any sentential operator Δ: Δ is upward monotonic iff whenever ϕ entails ψ, $\Delta(\phi)$ entails $\Delta(\psi)$.

The upward monotonicity property means that in order to speak truly about what is necessary or possible, one need not be completely specific. For example, consider 'ought'. By the orthodox semantics above, 'Alicia ought to mail the letter overnight' is true at w iff every world in $R(w)$ is one where 'Alicia mails the letter overnight' is true. Since 'Alicia mails the letter overnight' entails 'Alicia mails the letter', whenever that latter condition holds, every world in $R(w)$ will also be one where 'Alicia mails the letter' is true. Thus, the less specific sentence 'Alicia ought to mail the letter' is true whenever the more specific sentence 'Alicia ought to mail the letter overnight' is.

By the same token, however, orthodox theories make the counter-intuitive predictions they do in our two puzzles. Both are a direct result of the fact that a disjunction $\phi \vee \psi$ is asymmetrically entailed by its disjuncts, ϕ and ψ. This means that in general, $\Box\phi$ asymmetrically entails $\Box(\phi \vee \psi)$, and $\Diamond\phi$ asymmetrically entails $\Diamond(\phi \vee \psi)$:

$$\Box\phi \vDash \Box(\phi \vee \psi)$$

$$\Diamond\phi \vDash \Diamond(\phi \vee \psi)$$

The Ross inference in (3.1) is just an instance of this schema. Since 'Alicia mails the letter' entails 'Alicia either mails the letter or burns it', orthodox theories predict that 'Alicia ought to mail the letter' entails 'Alicia ought to either mail the letter or burn it' ($\Box M \vDash \Box(M \vee B)$).

The invalidity of the free choice inference has the same source. Suppose there is only one relevant world, w. Further, suppose M is true and B is false at w. Then $\Diamond M$ is true, since w is a relevant world where M is true. Since $\Diamond$ is an upward monotonic operator, it follows that $\Diamond(M \vee B)$ is automatically true. However, since every relevant world (i.e., w) makes B false, we also have that $\Diamond B$ is false. This presents a counterexample to the validity of the argument in (3.2): the premise $\Diamond(M \vee B)$ is true while the conclusion, $\Diamond B$ is false. So the free choice inference is invalid.

3.3 Range

Before moving on to my analysis of the two puzzles, I want to briefly note their generality. While they have sometimes been treated as arising only for a special class of *deontic* modals like 'ought' and 'may',[6] recent research has made clear that they arise for a much wider class of modals,[7] and are not tied to any particular 'flavors' of modality.[8] To see this, let us schematize the inferences as follows, where Δ is a modal operator:[9]

(Ross schema) $\Delta(p)$ therefore $\Delta(p \vee q)$

(Free choice schema) $\Delta(p \vee q)$ therefore Δp and Δq

Consider the following instances of the Ross schema with non-deontic modals:

6. The puzzles were discovered in early work on deontic logic and the logic of imperatives (von Wright 1968; Kamp 1973; Ross 1941). Some recent research has continued this focus on deontic flavors of modality, including Barker (2010), Cariani (2013), Fusco (2015), and Starr (2016). Of course, as some of these authors mention, there are likely ways to extend these accounts, tailored to the deontic case, to other flavors of modality.

7. See Zimmermann (2006), Yablo (2014), Abreu Zavaleta (2019) for some examples of invalid, non-deontic Ross inferences, and Zimmermann (2000), Nickel (2010), Romoli and Santorio (2017), and Willer (2021) for some discussion of valid, non-deontic free choice inferences.

8. It is standardly assumed that modal auxiliaries like 'must' are context sensitive, and in various contexts can express different *flavors* of necessity, such as metaphysical, epistemic, deontic, and so on.

9. Throughout, by 'modal' or 'modal operator', I mean a sentential operator that is standardly analyzed as shifting the world of evaluation for its complement proposition. 'Modal' as I am using it thus includes not just modal auxiliaries like 'must' and 'may', but also attitude verbs like 'want' and 'believe'.

(3.3) a. An object in motion {must / will / has to / is likely to / might / can} stay in motion.

b. # So, an object in motion {must / will / has to / is likely to / might / can} either stay in motion or come to absolute rest.

The subject matter in these examples encourages physical, metaphysical, or epistemic interpretations of the modals, and yet the inference appears just as invalid as the one in (3.1). Similarly, the pattern appears invalid for many attitude verbs, which are usually given a necessity modal semantics. For some examples:

(3.4) a. Alicia {seeks / intends / hopes / wants / expects} to mail the letter.

b. # So, Alicia {seeks / intends / hopes / wants / expects} to either mail the letter or burn it.

(3.5) a. Alicia $\left\{\begin{array}{c}\text{said}\\ \text{claims}\\ \text{believes}\\ \text{thinks}\end{array}\right.$ that Bulmaro mailed the letter.

b. # So, Alicia $\left\{\begin{array}{c}\text{said}\\ \text{claims}\\ \text{believes}\\ \text{thinks}\end{array}\right.$ that Bulmaro either mailed the letter or burned it.

Each seems to be a complete non sequitur. See Zimmermann (2006), Yablo (2014), Abreu Zavaleta (2019) for further discussions of Ross inferences with some of these verbs.

Similarly, as Zimmermann (2000), Nickel (2010), Romoli and Santorio (2017), and Willer (2021) argue, the free choice pattern appears to be valid for non-deontic possibility modals. For some examples:

(3.6) a. Given their skill, Alicia or Bulmaro $\left\{\begin{array}{c}\text{might}\\ \text{could}\end{array}\right.$ have won the game.

b. So, Alicia $\left\{\begin{array}{c}\text{might}\\ \text{could}\end{array}\right.$ have won and Bulmaro $\left\{\begin{array}{c}\text{might}\\ \text{could}\end{array}\right.$ have won.

(3.7) a. Hydrangeas can grow either pink or blue flowers.

b. So, hydrangeas can grow pink flowers and they can grow blue flowers.

Each inference seems as well-supported as the original in (3.2), despite the fact that the modals take on a non-deontic interpretation.

This data suggests that the intuitive status of the two inferences has nothing to do with a deontic interpretation of the modals involved.[10] An adequate account of the puzzles, therefore, should be general enough to accommodate modals of any flavor.

10. This is not to claim that there are no modals for which the Ross inference is valid — perhaps there are. My claim is a weaker one: given these examples, the solution we propose to the puzzles should be flexible enough to handle any flavor of modality.

3.4 Independence

As mentioned in the introduction, a speaker's judgment of the (in)validity of an argument depends on what meanings the speaker takes the premises and conclusion to have. In judging the Ross inference to be invalid, a speaker senses that $\Box(M \vee B)$, for example, intuitively means something that the truth of $\Box M$ fails to guarantee. And in judging a free choice inference to be valid, a speaker senses that $\Diamond(M \vee B)$ intuitively means something that should guarantee the truth of both $\Diamond M$ and $\Diamond B$. Thus, a crucial step toward providing a solution to the two puzzles is to identify what intuitive meanings speakers take sentences of the form $\Delta(p \vee q)$ to have (I will use Δ as an variable ranging over possibility and necessity modals that give rise to either of the puzzles), such that orthodox theories fail to correctly predict that these sentences have this intuitive meaning.

Both pragmatic and semantic accounts of the puzzles have typically agreed on this issue.[11] They say that over and above its orthodox truth conditions, a modal claim of the form $\Delta(p \vee q)$ conveys that each disjunct is *compatible* with the relevant set of worlds. That is, they say that the information that $\Delta(p \vee q)$ intuitively conveys, and which orthodox theories fail to predict, is underwritten by what I will call *Diversity* inferences (where $\Rightarrow$ is ambiguous between semantic entailment, implicature, or some other robust form of licensing):

11. See, for example, von Fintel (2012) for a pragmatic version of this thesis and Simons (2005a) for a semantic one. As I mention below, Menéndez-Benito (2005, 2010) has dissented in the analogous case of free choice 'any' under possibility modals, and building on this work, Aloni and Ciardelli (2013) dissent in the analogous case of imperatives.

Diversity Inferences:

$$\begin{aligned} \Box(p \vee q) &\Rightarrow \Diamond p \\ &\Rightarrow \Diamond q \\ \Diamond(p \vee q) &\Rightarrow \Diamond p \\ &\Rightarrow \Diamond q \end{aligned}$$

For modals without lexicalized duals, like 'want' and 'intend', we can give a meta-linguistic characterization of the Diversity inferences as follows:

Meta-Linguistic Diversity Inferences:

Where Δ is a necessity or possibility modal, $R(w)$ is the set of relevant worlds associated with the modal Δ at w, and $[\![p]\!]$ is the set of worlds in which p is true:

$$\begin{aligned} \ulcorner\Delta(p \vee q)\urcorner \text{ is true at } w &\Rightarrow R(w) \cap [\![p]\!] \neq \varnothing \\ &\Rightarrow R(w) \cap [\![q]\!] \neq \varnothing \end{aligned}$$

The Diversity inferences, not supported by the orthodox semantics, promise to explain our intuitive judgments in both puzzles. For free choice, this is trivial: the inference just is a special case of Diversity. For the Ross inference, the thought is usually that because of the Diversity inferences, the conclusion $\Box(M \vee B)$ conveys $\Diamond B$, and since the premise $\Box M$ does not guarantee this, the argument is judged to be invalid.[12] I will call the theory that the Diversity inferences capture the unorthodox content conveyed by modals with disjunctive complements, and that these inferences explain the gap between the orthodox semantics and our intuitions about the two puzzles, the 'Diversity

12. See, for example, Wedgwood (2006) and von Fintel (2012).

analysis'.

Unfortunately, the Diversity analysis cannot explain all of the relevant data. In particular, there are two pieces of data it cannot explain: what I will call *extended Ross's puzzle*; and what I will call *independence conditional inferences*. First, let me introduce the former. The Diversity analysis says that the argument in (3.1) is (or merely appears) invalid because the conclusion entails (or implicates) the permissibility of the added disjunct — a fact not guaranteed by the truth of the premise. If that analysis were correct, then adding a premise that explicitly guarantees the permissibility of the added disjunct should change whether or not the argument is judged to be invalid. However, this prediction is not borne out. Even in cases where both disjuncts are guaranteed to be permissible by the premises, the inference pattern seems invalid (see Sayre-McCord (1986) and Fusco (2015) for this point). For example:

Extended Ross's Puzzle.

(3.8) a. Alicia ought to mail the letter. $\Box M$

b. Alicia may use the phone. $\Diamond P$

c. # So, Alicia ought to either mail the letter or use the phone. $\Box(M \vee P)$

The conclusion in (3.8c) does not seem to follow. But the premises guarantee that the Diversity inferences are supported.[13] So the defectiveness of the inference in (3.8) cannot be a result of the failure of the Diversity inferences to be licensed by the premises. Notice that like the original Ross inference, the extended pattern is also defective for non-deontic modals. For example, consider the case of epistemic 'must' and 'might':

(3.9) a. Alicia must have mailed the letter. $\Box M$

b. Alicia might have used the phone. $\Diamond P$

13. I am making the standard assumption that modals, like other quantifiers, presuppose that their domains are non-empty. Thus, $\Box M$ entails that there is a relevant world where M is true. This means that whenever $\Box M$ is true, $\Diamond M$ is also true.

c. # So, Alicia must have either mailed the letter or used the phone. $\Box(M \vee P)$

The second type of data that the Diversity analysis cannot explain arises from the interaction between modals and conditionals.[14] As is well known, conditionals of the form 'If ϕ, then $\Delta(\psi)$' (where Δ is a modal) often give rise to 'restriction' readings, where the semantic function of the antecedent clause (ϕ) seems to be to restrict the domain of worlds relevant for the modal Δ in the consequent to the subset of relevant worlds where the antecedent is true ($R(w) \cap [\![\phi]\!]$).[15] When modals capable of such restriction have disjunctive complements, their acceptance appears to license the inference of what I call *independence conditionals*:

Independence Conditionals.

(3.10) a. Alicia ought to either write an essay or give a presentation. $\Box(E \vee P)$

b. So, if she doesn't write an essay, she ought give a presentation. $\neg E \rightarrow \Box P$

c. So, if she doesn't give a presentation, she ought to write an essay. $\neg P \rightarrow \Box E$

(3.11) a. Alicia may either write an essay or give a presentation. $\Diamond(E \vee P)$

b. So, if she doesn't write an essay, she may give a presentation. $\neg E \rightarrow \Diamond P$

c. So, if she doesn't give a presentation, she may write an essay. $\neg P \rightarrow \Diamond E$

Given the standard assumption that a necessity modal like 'ought' presupposes that its domain is non-empty, the restriction readings of these conditionals have the following truth conditions:

(3.10b) True iff there are some relevant worlds where Alicia doesn't write an essay, and in all of these, she gives a presentation ($R(w) \cap [\![\neg E]\!] \neq \varnothing$ and $R(w) \cap [\![\neg E]\!] \subseteq [\![P]\!]$).

14. See also Fusco (2015) for this data in the case of 'ought', and an alternative account of it.

15. See Kratzer (2012a, 2012c) for the classic theory of this phenomenon in the case of non-attitude modals, and Blumberg and Holguín (2019) for recent work on restriction effects in the case of attitude verbs.

(3.10c) True iff there are some relevant worlds where Alicia doesn't give a presentation, and in all of these, she writes an essay ($R(w) \cap [\![\neg E]\!] \neq \varnothing$ and $R(w) \cap [\![\neg E]\!] \subseteq [\![P]\!]$).

(3.11b) True iff in some relevant world where Alicia does not write an essay, she gives a presentation ($(R(w) \cap [\![\neg E]\!]) \cap [\![P]\!] \neq \varnothing$).

(3.11c) True iff in some relevant world where Alicia does not give a presentation, she writes an essay ($(R(w) \cap [\![\neg P]\!]) \cap [\![E]\!] \neq \varnothing$).

Neither orthodox semantic theories nor the Diversity analysis predicts that (3.10b-3.10c) should follow from (3.10a) or that (3.11b-3.11c) should follow from (3.11a). To see this, suppose that (3.10a) and (3.11a) are true. Orthodox semantic theories allow that (3.10a) and (3.11a) can be true if there are no relevant B-worlds, or if there are no relevant M-worlds. In the former case, the conditionals (3.10b) and (3.11b) would be false, while in the latter, (3.10c) and (3.11c) would be false. Thus, orthodox semantic theories do not validate these conditional inferences. But the Diversity inferences cannot explain their plausibility either. The truth conditions of (3.10b-3.10c) and (3.11b-3.11c) require not just that there is some relevant E-world and some relevant P-world. Rather, (3.10b) and (3.11b) each require more specifically that there is a relevant P-and-not-E-world, and (3.10c) and (3.11c) each require that there is a relevant E-and-not-P-world.[16] In other words, these conditionals require that E and P are *independently* relevant alternatives, and, I claim, the acceptance of (3.10a) and (3.11a) appears to guarantee this.

In fact, this requirement of *independent* relevance also offers a plausible explanation of what goes wrong in the extended Ross's puzzle above. The conclusion (3.8c) does not merely convey that using the phone is allowed. It conveys that Alicia may fulfill her obligation(s) by using the phone *independently* of whether she mails the letter. The premises of (3.8) do not guarantee this. If (3.9) is true at w, then there is some world v in $R(w)$ where Alicia uses the phone. But given (3.8a), it follows that Alicia *also* mails

16. Again, in the case of (3.10b) and (3.10c), this assumes that natural language necessity modals like 'ought' presuppose that their domains are non-empty.

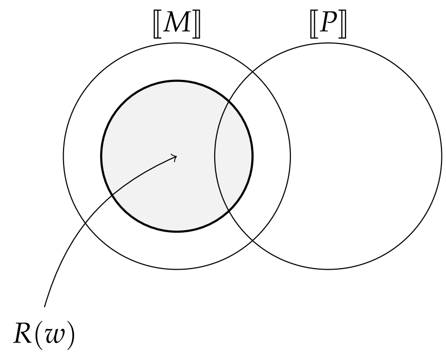

Figure 3.1: Diversity without Independence: $R(w) \cap [\![M]\!] \neq \varnothing$ and $R(w) \cap [\![P]\!] \neq \varnothing$

the letter at v. This situation is illustrated in Figure 3.1, where the relevant set of worlds, $R(w)$, makes the premises (3.8a) and (3.9) true. The problem, according to me, is that while using the phone is permissible, it is not an independent option on a par with mailing the letter. In fact, the premises of (3.8) entail that using the phone is precisely *not* an independent way for Alicia to fulfill her obligation(s), for if they are true, Alicia may use the phone only if she *also* mails the letter. For using the phone to constitute an independent way for Alicia's obligation(s) to be fulfilled, as I am claiming the conclusion in (3.8c) requires, there must also be a relevant world where she uses the phone without mailing the letter. That is just to say the conclusion (3.8c) requires both $\Diamond(M \wedge \neg P)$ and $\Diamond(P \wedge \neg M)$ to be true. In other words, the set of relevant worlds has to intersect the relative complements of the disjuncts, as illustrated in Figure 3.2.

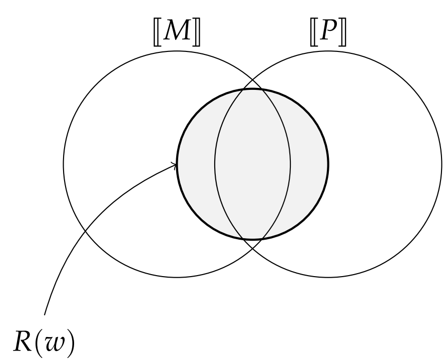

Figure 3.2: Independence: $R(w) \cap [\![M \wedge \neg P]\!] \neq \varnothing$; $R(w) \cap [\![P \wedge \neg M]\!] \neq \varnothing$

The Independence requirement I am proposing has some precedent in Menéndez-

Benito (2010), which argues that a similar condition governs the interaction between possibility modals and free choice 'any' (and Spanish 'cualquiera').[17] Suppose we are playing a card game and before you is the whole suit of diamonds on one side, and the whole suit of hearts on the other. If the rules of the game require that you now take one of these suits, then saying 'You may take any diamond card' is deeply misleading, as Menéndez Benito argues. For while it is true that for each diamond card, there is a permissible possibility in which you take it, the sentence seems to convey the stronger permission to take each diamond card *independently* of taking the others.[18] We may adapt the example to show something similar for the interaction between possibility modals and *disjunctive* complements. Suppose now that you have only one option: to take the pair consisting of the ace and ten of diamonds. Orthodox semantic treatments of 'may' would say that the following sentence is true:

(3.12) # You may take the ace or the ten of diamonds.

But (3.12) seems to misdescribe things. Orthodox semantic theories, by themselves, do not have the resources to explain why, since taking the pair is a relevant possibility, and doing so makes the embedded disjunction true. But the Diversity analysis cannot explain why (3.12) is misleading either, since taking the pair involves taking the ace and taking the ten, making each disjunct true. By contrast, an extension of our analysis of (3.8) does offer an explanation for why (3.12) is misleading: the disjuncts do not pick out *independent* alternatives. Since the only relevant option is to take the whole pair, you cannot take the ace of diamonds *without* taking the ten of diamonds, or vice versa. On my view, then, (3.12) is infelicitous because the alternatives it describes are not independently possible.[19]

17. I will not discuss free choice 'any' in this chapter, but I think it will be clear enough how my account could be extended to that case in order to capture the data Menéndez Benito puts forward.

See also Aloni and Ciardelli (2013), which applies Menéndez Benito's insight to the case of imperatives.

18. The term used by Menéndez-Benito (2005, 2010)) is not *independence* but *exclusivity*. I have opted for 'independence' here since I do not want to suggest any erroneous connections to 'exclusive' disjunction.

19. In fact, on the theory I go on to develop in this chapter, (3.12) will be neither true nor false in the described situation. I discuss the status of (3.12) further in §8.

In sum, I argue that both possibility and necessity modals license what I will call *Independence* inferences. Accepting $\Delta(p \vee q)$ disposes us to accept that p and q are each independent alternatives in the relevant domain of possibilities: i.e., that both (i) p-without-q, and (ii) q-without-p are compatible with the relevant set of worlds. For necessity and possibility modals with lexicalized duals, we can schematize these inferences as follows:

Independence Inferences:

$$\begin{aligned}\Box(p \vee q) &\Rightarrow \Diamond(p \wedge \neg q)\\ &\Rightarrow \Diamond(q \wedge \neg p)\\ \Diamond(p \vee q) &\Rightarrow \Diamond(p \wedge \neg q)\\ &\Rightarrow \Diamond(q \wedge \neg p)\end{aligned}$$

For modals without lexicalized duals like 'want', or 'intend', we can give a meta-linguistic characterization of Independence as follows:

Meta-Linguistic Independence Inferences:

Where Δ is a necessity or possibility modal, $R(w)$ is the set of relevant worlds associated with the modal Δ at w, and $[\![p]\!]$ is the set of worlds in which p is true:

$$\begin{aligned}\ulcorner\Delta(p \vee q)\urcorner \text{ is true at } w &\Rightarrow R(w) \cap ([\![p]\!] \backslash [\![q]\!]) \neq \varnothing\\ &\Rightarrow R(w) \cap ([\![q]\!] \backslash [\![p]\!]) \neq \varnothing\end{aligned}$$

Note that, since $\Diamond(p \wedge \neg q)$ entails $\Diamond p$, the Independence inferences are strictly stronger than the Diversity inferences.

Of course, the Independence inferences only make sense if the embedded disjunction $p \vee q$ satisfies 'Hurford's constraint' against redundant disjuncts — that is, if neither p

nor *q* entails the other (see Hurford (1974) and Gazdar (1979)). I will assume, following recent work on the topic, that felicitous disjunctions which appear to flout Hurford's constraint at the level of surface grammar are interpreted at the level of logical form via 'exhaustification' operators that enforce conformity to the constraint.[20] This means that for the interpreted structures we are interested in, we may safely assume Hurford's constraint holds and neither *p-and-not-q* nor *q-and-not-p* are contradictions.

In the rest of this chapter, I will provide a positive account of the interactions between modals and disjunctions that predicts the Independence inferences. Before moving on to that project, I want to note that it begins with an important theoretical choice: whether to explain the Independence inferences pragmatically or semantically. Presenting a conclusive argument in favor of one or the other strategy would not be possible here. Instead, I will limit myself to noting some problems that pragmatic accounts face. I think these problems suggest that we should opt for a semantic account. First, if hearers make the Independence inferences pragmatically (as Wedgwood (2006) and von Fintel (2012) argue for the Diversity inferences), then it is unclear why Ross inferences should seem so persistently *invalid*. Pragmatic inferences in general are usually drawn when they seem plausible. But a pragmatic account of the Independence inferences hypothesizes that when encountering the argument in (3.1), a hearer draws the Independence inferences precisely when they are *implausible*, given the premise. Indeed, on a pragmatic account, the hearer is supposed to draw the Independence inferences, sense their implausibility,

20. For example, take:

(3.1) Alicia must have either some or all of the ice cream.

The surface grammar of the embedded disjunction, 'Alicia had either some or all of the ice cream,' flouts Hurford's constraint, since having *all* of the ice cream entails having *some* of it. This means that the modal claim (3.1) could not possibly license both of the Independence inferences, since one of them would be:

(3.2) Alicia may have all but not some of the ice cream.

(3.2) cannot be true, since the embedded conjunction, 'Alicia has all but not some of the ice cream,' is a contradiction. The assumption I will make is that in recognizing this, a hearer assigns to (3.1) a logical form roughly equivalent to, 'Alicia must have either *merely* some, or all of the ice cream" (where the exhaustification operator *merely* has the semantic function of denying that Alicia has all of the ice cream. See Simons (2001), Katzir and Singh (2013), Meyer (2013, 2014), and Ciardelli and Roelofsen (2017b) for recent discussion.

and rather than withdraw them, mistake a semantically valid argument for an invalid one.[21]

Furthermore, recent empirical research on pragmatic approaches to the free choice inferences suggests they may have some trouble explaining data concerning processing times (Chemla and Bott 2014) and the interaction between these inferences, on the one hand, and non-monotonic contexts or presuppositions, on the other (see Romoli and Santorio (2019) and Gotzner, Romoli, and Santorio (2020)). A semantic account of the Independence inferences has the potential to do better on all of these points. For these reasons, the solution to our puzzles that I outline below will derive the validity of the Independence inferences as part of the semantics of sentences of the form $\Delta(p \vee q)$.

3.5 Independence and Minimal Covers

Now that we have seen that the Independence inferences are needed to explain the puzzles, what consequence does this have for the semantics of modals with embedded disjunctions? First, let us consider necessity modals. Validating the Independence inferences means that the truth conditions of $\Box(p \vee q)$ will involve two requirements. First, there is the one that orthodox accounts predict: that $\Box(p \vee q)$ is true only if $p \vee q$ is true in all of the relevant worlds. Second, p and q must pick out independent alternatives among the relevant set of worlds — that is, the relative complements of the truth sets of the disjuncts (the set of p-without-q worlds, the set of q-without-p worlds) must each

21. An anonymous reviewer suggests that the independence inferences might be computed by the combination of the Diversity inferences together and an *exclusive* reading of the embedded disjunction. On this hypothesis, the Independence inferences would only be licensed in cases where it is reasonable to suppose that there is no relevant world that makes the *conjunction* of the disjuncts true. While I cannot provide a conclusive argument against this hypothesis here, I think that the apparent invalidity of examples like (3.8) and (3.9) provide evidence against it. In those cases, the premises can be true together, and when they are both true, there is a relevant world that the conjunction of the disjuncts true — where Alicia mails the letter and uses the phone. In such a case, it should be natural for speakers to opt for an *inclusive* reading of the disjunction embedded in the conclusion, and recognize that it follows on the orthodox semantics (even supplemented with the Diversity inferences). But both inferences seem just as invalid as the original (3.1). This suggests to me that the Independence inferences do not arise only on an exclusive reading of the disjunction involved, and thus that the Independence inferences are not computed in the way this hypothesis claims.

have a non-empty intersection with the set of relevant worlds. Together, these two requirements mean that in order for $\Box(p \vee q)$ to be true, the truth sets of the disjuncts must form a *minimal cover* of the set of relevant worlds (as illustrated by Figure 3.2 above).[22] A minimal cover C of a set S is a collection of sets such that (i) their union contains S (they *cover* S); and (ii) no proper subset of C is such that its union contains S (they do so *minimally*):

$$\begin{aligned} C \text{ is a } \textbf{cover} \text{ of } S \text{ iff } \; & S \subseteq \bigcup C \\ C \text{ is a } \textbf{minimal cover} \text{ of } S \text{ iff } \; & C \text{ is a cover of } S \text{ and} \\ & \text{there is no } C' \subset C : C' \text{ is a cover of } S \end{aligned}$$

To break this down: the first requirement associated with the truth of $\Box(p \vee q)$ — that $p \vee q$ be true throughout the relevant set of worlds — is equivalent to the requirement that the truth sets of the disjuncts ($\{[\![p]\!], [\![q]\!]\}$) form a *cover* of the relevant set of worlds. The second requirement — that among the relevant worlds, there are both p-without-q-worlds and q-without-p-worlds — is equivalent to the requirement that the cover be a *minimal* one: neither the truth set of p nor that of q suffices on its own to cover the set of relevant worlds.

We can also use the notion of a minimal cover to explain what validating the Independence inferences will mean for the truth conditions for sentences of the form $\Diamond(p \vee q)$. First, there is the part that orthodox semantics predict: since $\Diamond(p \vee q)$ is a

22. Simons (2005a) also uses a *covering* relation as a helpful way of summarizing the modal/disjunction interaction. For Simons, modals with disjunctive complements truth conditionally require that the disjuncts form a *supercover* of the relevant set of worlds. C is a supercover of S iff it is a cover of S and every member of C has a non-empty intersection with S. Note that every minimal cover is a super cover but not vice versa; and that a supercover semantics validates the Diversity, but not Independence, inferences. See also Nygren (2019), which systematically explores the logic of a supercover semantics.

At the end of her paper (§6), Simons briefly considers various pragmatic 'add-on' requirements that she thinks may govern the felicity of disjunctions in certain contexts. One of the three requirements she outlines resembles the minimal covering relation I define here. However, she does not explore this pragmatic constraint in much detail, and clearly does not think, as I argue here, that it is part of the literal, truth conditional semantics of modals with embedded disjunctions.

possibility claim, its truth requires that there is at least one relevant world where $p \vee q$ is true. Equivalently, the truth sets of the disjuncts must form a cover of a *non-empty* subset of relevant worlds. Second, to support the Independence inferences, the relative complements of the contents of the disjuncts (the set of p-without-q worlds, the set of q-without-p worlds) must each have a non-empty intersection with the set of relevant worlds $R(w)$. Together, these requirements mean that the truth sets of the disjuncts must form a minimal cover of some non-empty subset of relevant worlds. This is illustrated in Figure 3.3, where $\{[\![p]\!], [\![q]\!]\}$ forms a minimal cover of the shaded subset of relevant worlds, R'.

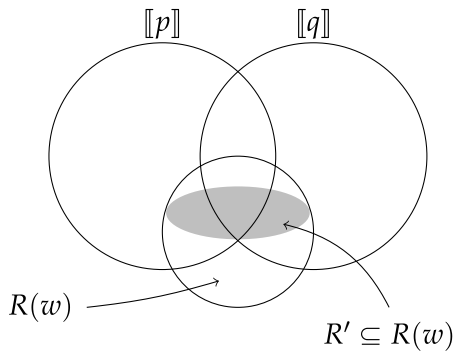

Figure 3.3: Independence for possibility modals: the shaded region R' contains both p-without-q worlds, and q-without-p worlds.

Using the notion of a minimal cover, our target truth conditions for sentences of the form $\Box(p \vee q)$ and $\Diamond(p \vee q)$ are as follows:

$$\Box(p \vee q) \text{ is true at } w \text{ iff } \{[\![p]\!], [\![q]\!]\} \text{ is a minimal cover of } R(w)$$

$$\Diamond(p \vee q) \text{ is true at } w \text{ iff there is a non-empty } R' \subseteq R(w) \text{ such that } \{[\![p]\!], [\![q]\!]\} \text{ is a minimal cover of } R'$$

Since these truth conditions make reference to the semantic values of the disjuncts of the complement clause, generating these truth conditions compositionally requires that we adopt a framework in which we can recover, from the semantic value of a disjunction, the semantic values of its disjuncts. The traditional possible worlds framework, on which the semantic value of a sentence is simply its truth set, makes this impossible. For example, given a disjunction that denotes $\{w_1, w_2, w_3\}$, there is no way to tell whether it was composed from disjuncts denoting $\{w_1, w_2\}$ and $\{w_3\}$, or from disjuncts denoting $\{w_1\}$ and $\{w_2, w_3\}$. The traditional framework thus ensures that propositional operators like modals and attitude verbs are blind to the disjunctive structure of their arguments.

Two related frameworks that allow a modal to see the disjunctive structure of its argument are *alternative semantics* (2002, Alonso-Ovalle (2006), Aloni (2007)) and *inquisitive semantics* (Ciardelli and Roelofsen (2011) Aloni and Ciardelli (2013), Ciardelli, Groenendijk, and Roelofsen (2018), Ciardelli, Roelofsen, and Theiler (2017)). For reasons I leave to a footnote,[23] I will use a version of inquisitive semantics.

I will present the model in two stages.[24] First, in §3.6, I will outline a semantics for modals in a fragment of propositional inquisitive semantics, where propositions consist of alternatives relevant for their *truth*. This will allow me to highlight the basics of the interaction between modals and disjunctions, and show how my semantics accounts for the original two puzzles as well as the data from §3.4. Then, I will note in §3.7 that this implementation of the theory faces two problems resulting from its interaction with negation: (i) it no longer supports impossibility and unnecessity distribution inferences over disjunctions (examples (3.17) and (3.18) below); and (ii), it gives up the duality between possibility and necessity modals. In response to these issues, I extend the account to a *bilateral* system, where propositions consist of *two* kinds of alternatives: those relevant for their truth and those relevant for their falsity.[25]

23. One of the key differences between the two frameworks arises in cases when the set of worlds where 'q' is true is a subset of the worlds where 'p' is true ($[\![q]\!] \subseteq [\![p]\!]$). On the traditional possible worlds analysis of propositions and disjunction, in this case the proposition denoted by '$p \vee q$' is identical to the one denoted by just 'p'. In inquisitive semantics, the same is true, the propositions denoted by 'p' and '$p \vee q$' come out the same. But standard versions of alternative semantics distinguish between these propositions (Roelofsen (2013), Ciardelli, Roelofsen, and Theiler (2017)). This means that alternative semantics, but not inquisitive semantics, gives up the traditional explanation of Hurford's constraint in terms of redundancy (Ciardelli and Roelofsen 2017b).

24. A formal summary of the semantic framework and results is contained in Appendix A.

25. The basic behavior of the connectives in the bilateral approach is formally similar to the 'radical

3.6 Minimal Covering Semantics

I will use a simple formal language to model the semantics I am arguing for, consisting of atomic formulae $p, q, ...$, the standard connectives $\neg, \wedge, \vee$, a possibility and a necessity modal $\Diamond, \Box$, and two special connectives: $\rightarrow$, a restrictor conditional; and $!$, the issue-cancelling operator of inquisitive semantics.

In standard possible worlds semantics, the semantic value of a sentence, a *classical proposition*, is just a set of worlds (or its characteristic function). Inquisitive semantics adds some complexity: an *inquisitive proposition* is a set of sets of worlds. The semantic value of an atomic formula is the set of all sets that only include worlds where it is true. In other words, it denotes its truth set, plus all of the subsets of that set. Let me illustrate with an example of two atomic sentences, p and q. Suppose p is true at w_1 and w_2, while q is true at w_2 and w_3. Then, where $[\phi]$ is the inquisitive proposition denoted by ϕ, we have:

$$[p] = \{\{w_1, w_2\}, \{w_1\}, \{w_2\}, \varnothing\}$$

$$[q] = \{\{w_2, w_3\}, \{w_2\}, \{w_3\}, \varnothing\}$$

In inquisitive semantics, conjunctions and disjunctions of propositions are treated just as in traditional possible world semantics: they denote the operations of set intersection and union. The proposition denoted by $p \wedge q$, then, is just the set containing everything in common between the proposition denoted by p and the one denoted by q. The proposition denoted by $p \vee q$ is just the set containing everything that is included either in the

inquisitive semantics' of Groenendijk and Roelofsen (2010) and Aher (2012), the dual update semantics in Willer (2018), the bilateral 'state-based' semantics of Aloni (2018), and the bilateral truthmaker semantics of Yablo (2014) and Fine (2017a, 2017b). While these other theorists share a similar semantic framework, and some also share an interest in our two puzzles, none of these accounts offers a theory that supports the *Independence* inferences we are interested in in this chapter.

one denoted by p or the one denoted by q. As applied to our simple example, we have:

$$[p \wedge q] = [p] \cap [q] = \{\{w_2\}, \varnothing\}$$
$$[p \vee q] = [p] \cup [q] = \{\{w_1, w_2\}, \{w_2, w_3\}, \{w_1\}, \{w_2\}, \{w_3\}, \varnothing\}$$

Two pieces of information made available by inquisitive propositions are important for our purposes. The first is the *informative content* of a sentence ϕ, written $\mathsf{info}(\phi)$. This is just the set of worlds where ϕ is true, corresponding directly to the traditional possible worlds proposition associated with ϕ. We can recover the informative content of any inquisitive proposition simply by taking the set of all of the worlds that are members of any element of the proposition (equivalently, the union of all sets of worlds in the proposition, $\mathsf{info}(\phi) = \bigcup[\phi]$). So, to continue our example, we have:

$$\mathsf{info}(p) = \bigcup[p] = \{w_1, w_2\}$$
$$\mathsf{info}(q) = \bigcup[q] = \{w_2, w_3\}$$
$$\mathsf{info}(p \wedge q) = \bigcup[p \wedge q] = \{w_2\}$$
$$\mathsf{info}(p \vee q) = \bigcup[p \vee q] = \{w_1, w_2, w_3\}$$

The second piece of information that the inquisitive semantics framework makes available is what sets it apart from the traditional possible worlds framework. This is the notion of the *alternatives* offered by a proposition ϕ, written $\mathsf{alt}(\phi)$. These correspond to the *largest* sets included in the inquisitive proposition denoted by a sentence, and will give us what we are after: the ability to recover from the semantic value of a disjunction the semantic values of each disjunct.[26] In our example, for instance, the alternatives

26. 'Largest sets' here means the sets in the proposition such that there is no proper superset also in the proposition. Officially:

$$\mathsf{alt}(\phi) = \{s \in [\phi] \mid \text{for every } t \in [\phi], \text{ if } s \subseteq t \text{ then } s = t\}$$

offered by $p \vee q$ are the largest sets in $[p \vee q]$, i.e. $\{w_1, w_2\}$ and $\{w_2, w_3\}$. By no accident, these sets are identical to the informative content, or truth sets, of the two disjuncts, p and q:

$$\mathsf{alt}(p \vee q) = \{\{w_1, w_2\}, \{w_2, w_3\}\} = \{\mathsf{info}(p),\ \mathsf{info}(q)\}$$

Compare the case of $[p \wedge q]$, which has only one largest member, the singleton containing w_2, where both p and q are true. In this basic propositional fragment of inquisitive semantics, *only* disjunctions offer multiple alternatives.[27] Every sentence without a disjunction offers only a single alternative, namely, its informative content (i.e., for non-disjunctive ϕ, $\mathsf{alt}(\phi) = \{\mathsf{info}(\phi)\}$).

Using the notion of the alternatives offered by a proposition, then, we can give a semantics for modals that is sensitive to the disjunctive structure of their complement clauses. We will say that a necessity modal claim of the form $\Box\phi$ is true at w just in case the alternatives offered by ϕ form a minimal cover of the relevant set of worlds, $R(w)$. Similarly, a possibility modal claim of the form $\Diamond\phi$ will be true at w just in case the alternatives offered by ϕ form a minimal cover of a non-empty subset R' of the relevant set of worlds, $R(w)$:

$\Box\phi$ is true at w iff $\mathsf{alt}(\phi)$ forms a minimal cover of $R(w)$

$\Diamond\phi$ is true at w iff there is a non-empty $R' \subseteq R(w)$

such that $\mathsf{alt}(\phi)$ forms a minimal cover of R'

These truth conditions, seemingly tailored to the case when ϕ is a disjunction, reduce to the orthodox semantics for modals when ϕ contains no disjunction. Let me explain why. As mentioned, when ϕ contains no disjunction, it offers a single alternative, corresponding to its truth conditions, i.e. $\mathsf{alt}(\phi) = \{\mathsf{info}([\phi])\}$. Now, for a singleton set

27. In richer versions of inquisitive semantics, other expressions like existential quantifiers and interrogative operators also introduce multiple alternatives.

like $\{\mathsf{info}([\phi])\}$ to form a *minimal* cover of a (non-empty) set S is just for it to cover S *simpliciter*, i.e. for $S \subseteq \mathsf{info}([\phi])$. Thus, for a necessity modal $\Box$, when ϕ contains no disjunction, it follows that $\Box\phi$ is true at w iff ϕ is true at every relevant world, i.e. iff $R(w) \subseteq \mathsf{info}(\phi)$. For possibility modals, when ϕ contains no disjunction, $\Diamond\phi$ is true at a world w iff there is a non-empty subset R' of relevant worlds such that R' is a subset of $\mathsf{info}(\phi)$. In other words, $\Diamond\phi$ is true iff there is at least one relevant world where ϕ is true.

In order to discuss the payoffs of my semantics, we must first talk about the semantics of negation ($\neg$) and the restrictor conditional ($\rightarrow$). In basic inquisitive semantics, the negation $\neg\phi$ of a proposition ϕ denotes the set of all sets of worlds that have no overlap with $\mathsf{info}(\phi)$. To illustrate by way of our running example, suppose there is just one more world, w_4, where both atoms p, q are false. Then $[\neg p]$ is the set of all sets of worlds that do not overlap with $\mathsf{info}(p) = \{w_1, w_2\}$. This means we have:

$$[\neg p] = \{\{w_3, w_4\}, \{w_3\}, \{w_4\}, \varnothing\}$$

I will use the conditional $\rightarrow$ solely to express conditional restriction effects. So we will say that $\phi \rightarrow \psi$ is true at w iff ψ is true at w when the set of relevant worlds $R(w)$ is altered so as to only include worlds where ϕ is true: $R(w) \cap \mathsf{info}(\phi)$.[28]

Now let us turn to the results, for which we will suppose that for atomic p, q, $p \vee q$ satisfies Hurford's constraint, i.e. neither p nor q entails the other.

Independence Inferences. Suppose that either $\Diamond(p \vee q)$ or $\Box(p \vee q)$ is true at w. Either way, it follows that $\{\mathsf{info}(p), \mathsf{info}(q)\}$ forms a minimal cover of some non-empty subset of the relevant worlds, R' ($R' = R(w)$ in the case of $\Box(p \vee q)$). Then there are at least two relevant worlds in R', call them v, u, such that one of them, say v, is included in $\mathsf{info}(p) \backslash \mathsf{info}(q)$, while the other, u, is included in $\mathsf{info}(q) \backslash \mathsf{info}(p)$. Now, $\Diamond(p \wedge \neg q)$ is true iff the alternatives offered by $p \wedge \neg q$ form a minimal cover of a non-empty subset

28. See the dynamic model update conditional in van Ditmarsch, van Der Hoek, and Kooi (2008), or Appendix A, for precise versions of such a conditional.

of relevant worlds. Since $p \wedge \neg q$ just offers a single alternative, corresponding to the worlds where it is true, and v is a $\neg q$-world, it follows that $\{v\}$ is a subset of relevant worlds that is minimally covered by the alternatives offered by $p \wedge \neg q$, i.e. $\mathsf{info}(p \wedge \neg q)$. So $\Diamond(p \wedge \neg q)$ is true. A similar argument shows that $\{u\}$ makes the other Independence inference, $\Diamond(q \wedge \neg p)$, true.[29]

Independence Conditionals. I will spell out just the necessity modal case, but the extension to possibility modals is straightforward. Assume that $\Box(p \vee q)$ is true. Then the alternatives offered by $p \vee q$, i.e. $\{\mathsf{info}(p), \mathsf{info}(q)\}$, form a minimal cover of the relevant set of worlds. This means again that every relevant world is either a p-world or a q-world, and further, that there are at least two relevant worlds, $v \in \mathsf{info}(p) \backslash \mathsf{info}(q)$ and $u \in \mathsf{info}(q) \backslash \mathsf{info}(p)$. On our semantics for the restrictor conditional, $\neg p \rightarrow \Box q$ is true iff $\Box q$ is true when we ignore all relevant worlds except the not-p worlds. Doing leaves us only with q-and-not-p-worlds like u. Since all such worlds are in the truth set of q, $\{\mathsf{info}(q)\}$ forms a minimal cover of them. Thus, $\Box q$ is true on the restriction to the $\neg p$ worlds, so $\neg p \rightarrow \Box q$ is true on our original assumption. A parallel argument shows $\neg q \rightarrow \Box p$ is true.

Free Choice. The free choice inference is predicted to be valid. Suppose $\Diamond(p \vee q)$ is true. Then the set of alternatives offered by $p \vee q$, namely $\{\mathsf{info}(p), \mathsf{info}(q)\}$, forms a minimal cover of some subset of the relevant worlds. Since the cover is *minimal*, it follows that there is at least one relevant p-and-not-q-world, call it v, and one q-and-not-p-world, call it u. Thus, there is a non-empty subset of relevant worlds, namely $\{v\}$, that is minimally covered by $\{\mathsf{info}(p)\}$. So $\Diamond p$ is true. Likewise, there is a non-empty subset of relevant worlds, namely $\{u\}$, that is minimally covered by $\{\mathsf{info}(q)\}$. So $\Diamond q$ is true. Thus, the free choice inference is validated: $\Diamond(p \vee q)$ entails $\Diamond p$ and $\Diamond q$.

Ross Inference. The Ross inference, by contrast, is invalid. Suppose p stands for 'Alicia mails the letter', and further, that $\Box p$, 'Alicia ought to mail the letter', is true. Then, the

29. In our running example, the truth of $\Diamond(p \vee q)$ or $\Box(p \vee q)$ requires that w_1, w_3 be among the relevant alternatives.

alternatives offered by p, namely $\{\mathsf{info}(p)\}$, cover the relevant set of worlds, $R(w)$, where Alicia fulfills her obligation. Now, consider the truth conditions of $\Box(p \vee q)$, where q is an independent disjunct like, 'Alicia burns the letter'. $\Box(p \vee q)$ is true iff the alternatives offered by $p \vee q$, namely $\{\mathsf{info}(p), \mathsf{info}(q)\}$, form a minimal cover of $R(w)$. But, they cannot. Since $\Box p$ is true, $\{\mathsf{info}(p)\}$ is a strictly smaller cover of $R(w)$. Thus, $\Box(p \vee q)$ is not true, and the argument is invalid.

Extended Ross's Puzzle. Suppose $\Box p$ and $\Diamond q$, are true. Because $\Box p$ is true, $\{\mathsf{info}(p)\}$ covers the relevant worlds. And since $\{\mathsf{info}(p)\}$ is a minimal cover of the relevant worlds, $\{\mathsf{info}(p), \mathsf{info}(q)\}$ is not. Thus, the conclusion of the extended Ross's puzzle, $\Box(p \vee q)$, is not true. It does not matter whether q is compatible with the relevant alternatives; the inference is invalid because q does not pick out an *independent* alternative relative to p.

Flexibility. Before turning to some difficulties faced by the present minimal covering semantics in the next section, I want to note that standard inquisitive semantics posits the linguistic resources to allow for some flexibility when it comes to drawing the Independence inferences. Indeed, standard inquisitive semantics posits the existence of an 'issue-cancelling' operator (denoted by '!'), which has the effect of eliminating the distinction between alternatives in an inquisitive proposition: any previously distinguished alternatives are collapsed into a single, undifferentiated one. This operator has been posited to distinguish between the semantic values of two questions one can ask using a sentence like 'Does Alicia speak Hindi or Tamil?':

(3.13) a. Does Alicia speak Hindi-or-Tamil$^{\uparrow}$?

b. Does Alicia speak Hindi$^{\uparrow}$ or Tamil$^{\downarrow}$?

(where $\uparrow$/$\downarrow$ indicate rising/falling intonation. The intended reading of the first question, (3.13a), is polar; it can be resolved with *Yes* (Alicia speaks at least one of Hindi and Tamil) or *No* (Alicia speaks neither language). By contrast, the second question, (3.13b), is not polar, and the conventional ways to resolve it directly are to either say that Alicia speaks Hindi, or to say that she speaks Tamil. In inquisitive semantics, the difference in the conventional resolutions of these questions is explained by a difference in the *alternatives* they offer. Most relevant for our purposes, (3.13b) treats Hindi and Tamil as distinct alternatives, but (3.13a) treats Hindi-or-Tamil as a single alternative. In order to generate *different* alternatives for sentences with the same surface structure, inquisitive semantics hypothesizes that the polar question (3.13a) contains an operator (represented with '!') that erases the distinction between the Hindi and Tamil alternatives. For our purposes, the semantics of ! is important insofar as it gives rise to the following identities. For any sentence ϕ:[30]

$$\mathsf{alt}(!\phi) = \{\mathsf{info}(\phi)\}$$

$$\mathsf{info}(!\phi) = \mathsf{info}(\phi)$$

So $!\phi$ is true iff ϕ is, but its proposition always offers a single alternative corresponding to its truth conditions. Above, I explained that for a non-disjunctive formula ϕ, my minimal covering semantics assigns orthodox truth conditions to $\Box\phi$ and $\Diamond\phi$. Since the ! operator makes even a disjunction offer just a single alternative, the same reasoning extends to arbitrary formulae of the form $!\phi$. Even when ϕ contains disjunction, my semantics assigns orthodox truth conditions to $\Box!\phi$ and $\Diamond!\phi$.

With this operator, we can explain how in certain special contexts, the Independence inferences may not be licensed. For example, in an epistemology class, one might hear the following:

(3.14) a. Given the evidence, Smith must own a Ford.

b. So, given the evidence, Smith must either own a Ford or live in Barcelona.

On my view, the premise of (3.14) rules out that *Smith lives in Barcelona* is an inde-

30. The official semantics for the operator is:

$$[!\phi] = \wp(\mathsf{info}(\phi))$$

pendent alternative relative to his owning a Ford. This would normally ensure that the conclusion is not true. In order to make the discourse coherent, then, interpreters may insert the issue-cancelling operator (!) just above the disjunction. When the inference in (3.14) is heard as acceptable, this is because it is not actually of the form $\Box p \Rightarrow \Box(p \vee q)$. Rather, it has the form $\Box p \Rightarrow \Box!(p \vee q)$. In support of this hypothesis, notice that just as with the question (3.13a), a monotonous intonation seems to encourage acceptability:

(3.15) a. Given the evidence, Smith must own a Ford.

b. So, given the evidence, Smith must either own-a-Ford-or-live-in-Barcelona.

In contrast, an alternating intonation, as with (3.13b), makes it harder to accept:

(3.16) a. Given the evidence, it must be that Smith owns a Ford.

b. # So, Smith must either own a Ford$^{\uparrow}$ or live in Barcelona$^{\downarrow}$.

Note that this form of flexibility is very different from a pragmatic account of the Independence inferences. On a semantic account like mine, the unorthodox content captured by the Independence inferences is part of the default, literal meaning of modals with disjunctive complements, and *not* drawing these inferences requires special interpretive work.

3.7 A Bilateral Version

The minimal covering semantics improves upon the orthodox semantics when it comes to predicting the meanings of unembedded modal claims. But, like some other semantic accounts of the puzzles, it is thereby worse than the orthodox semantics at predicting the meanings of modal claims embedded under downward entailing operators like negation.[31] By assigning stronger-than-orthodox truth conditions to a bare disjunctive modal claim ($\Delta(p \vee q)$), the theory assigns weaker-than-orthodox falsity conditions

31. For example, as Alonso-Ovalle (2006) shows, the semantics of Simons (2005a) suffers these problems with negation. See Aloni (2018) and Willer (2018) for alternative bilateral solutions to these problems, and Aloni (2007) for a unilateral response to these issues based on ambiguity.

to it, and thus assigns weaker truth conditions to its negation $\neg\Delta(p \vee q)$. In particular, it gives up the validity of the following two patterns of inference, both of which are validated by the orthodox semantics:

Impossibility Distribution.

(3.17) a. Alicia may not mail the letter or burn it. $\neg\Diamond(M \vee B)$

b. So, Alicia may not mail the letter. $\neg\Diamond M$

c. So, Alicia may not burn the letter. $\neg\Diamond B$

Unnecessity Distribution.

(3.18) a. Alicia doesn't have to either mail the letter or burn it. $\neg\Box(M \vee B)$

b. So, Alicia doesn't have to mail the letter. $\neg\Box M$

c. So, Alicia doesn't have to burn the letter. $\neg\Box B$

A second problem is that the basic minimal covering semantics gives up predicting the *duality* of possibility and necessity modals. Duality (the truth conditional equivalence of $\neg\Box\phi$ and $\Diamond\neg\phi$ on the one hand, and $\neg\Diamond\phi$ and $\Box\neg\phi$ on the other) is one of the fundamental virtues of the orthodox analysis, which derives the duality of necessity and possibility modals from the duality of the universal and existential quantifiers. To illustrate where duality fails on the minimal covering semantics of the previous section, consider the following pair:

Duality.

(3.19) a. Alicia may not mail the letter or burn it $\neg\Diamond(M \vee B)$

b. Alicia must not mail the letter or burn it $\Box\neg(M \vee B)$

Intuitively, these two sentences are equivalent. Orthodox semantic theories predict as much. The basic minimal covering semantics I have sketched, however, strengthens the truth conditions of $\Diamond(M \vee B)$, weakening the truth conditions of $\neg\Diamond(M \vee B)$. It says that (3.19a) is true iff the alternatives offered by $M \vee B$ (Alicia's mailing the letter, Alicia's burning the letter) do not form a minimal cover of the relevant set of worlds. This would hold, for example, in case in all of the relevant worlds, Alicia burns the letter (i.e. if $\Box B$ were true). Meanwhile, it assigns orthodox truth conditions to (3.19b): it is true iff in every relevant world, M and B are both false (so, e.g., it entails $\Box\neg M$ and $\Box\neg B$). Thus, while the truth of (3.19b) prohibits Alicia from burning the letter, the basic minimal covering semantics allows (3.19a) to be true even if Alicia *must* burn the letter. Clearly, the predicted truth conditions of (3.19a) are far too weak.

Fortunately, a natural extension of the semantics solves these problems. The key is to define the falsity conditions of our sentences independently of their truth conditions, and to allow for truth value gaps.[32] I will use the same language as before, but now assign *bilateral inquisitive propositions* to sentences. These bilateral inquisitive propositions will be modeled as *pairs* of regular inquisitive propositions that have only one member in common: the empty set. To return to our simple example of four worlds, we now assign to p a *positive* part of its bilateral proposition (written $[p]^+$), corresponding to the unilateral inquisitive proposition from before:

$$[p]^+ = \{\{w_1, w_2\}, \{w_1\}, \{w_2\}, \varnothing\}$$

and we define the negative part (written $[p]^-$), as the set of worlds where p is false, plus all of its subsets (corresponding to what was previously the unilateral value of $\neg p$):

$$[p]^- = \{\{w_3, w_4\}, \{w_3\}, \{w_4\}, \varnothing\}$$

32. See Groenendijk and Roelofsen (2010), Aloni (2018) and Willer (2018) for other versions of bilateral inquisitive semantics with similar motivations.

Extending these principles to our other atomic sentence, q, we have:

$$[q]^+ = \{\{w_2, w_3\}, \{w_2\}, \{w_3\}, \varnothing\}$$
$$[q]^- = \{\{w_1, w_4\}, \{w_1\}, \{w_4\}, \varnothing\}$$

In order to make use of these negative parts, I will treat negation as an involution — it swaps the negative part of a proposition for its positive part, and vice versa:

$$[\neg\phi]^+ = [\phi]^-$$
$$[\neg\phi]^- = [\phi]^+$$

I will treat the positive contribution of conjunction and disjunction the same as before, namely as set intersection and union, respectively:

$$[\phi \wedge \psi]^+ = [\phi]^+ \cap [\psi]^+$$
$$[\phi \vee \psi]^+ = [\phi]^+ \cup [\psi]^+$$

As for the negative contributions of these connectives, I will, for simplicity, draw on the De Morgan equivalences. The negative part of a conjunction will be the union of the negative parts of its conjuncts, and the negative part of a disjunction will be the intersection of the negative parts of its disjuncts:[33]

$$[\phi \wedge \psi]^- = [\phi]^- \cup [\psi]^-$$
$$[\phi \vee \psi]^- = [\phi]^- \cap [\psi]^-$$

I will also extend definitions of the informative content of a sentence, and the alternatives offered by a sentence, to the bilateral model. Where we used to have just the positive alternatives offered by a sentence, we now have its positive ($\mathsf{alt}([\phi]^+)$) and negative ($\mathsf{alt}([\phi]^-)$) alternatives, defined in the same way as before. And where we used to have just the positive informative content of a sentence, we will now have its positive ($\mathsf{info}([\phi]^+)$) and negative ($\mathsf{info}([\phi]^-)$) informative contents, which correspond to the sets of worlds where it is true and false, respectively.

Figure 3.4 illustrates some simple examples of propositions. In each subfigure, the four circles are worlds corresponding to the classical valuations of the atomic sentences p, q (with $\overline{p}$ indicating p is false), and correspond to our informal model, starting with w_1 in the upper right, and w_2, w_3, w_4 moving counterclockwise around the square. Solid lines represent positive alternatives, while dotted lines represent negative alternatives.

Now, let us return to our modals. We define the *truth* conditions of the modal propositions as we did previously, though now with the detail that they depend on the *positive*

33. It is easy to see that given the semantics of negation, conjunction, and disjunction I outline here, the semantics predicts that each of the classical De Morgan equivalences holds. Some of these equivalences are controversial, especially when embedded under modals or conditionals. The theory I give here thus inherits some of this controversy. For example, it validates *Dual Free Choice*: when p and q are atoms that obey Hurford's constraint, $\neg\Box(p \wedge q) \vDash \Diamond\neg p \wedge \Diamond\neg q$. For given duality between $\Box$ and $\Diamond$, $[\neg\Box(p \wedge q)] = [\Diamond\neg(p \wedge q)]$. Then, given the De Morgan equivalences, $[\Diamond\neg(p \wedge q)] = [\Diamond(\neg p \vee \neg q)]$. Finally, given the validity of free choice, clearly $\Diamond(\neg p \vee \neg q)$ entails $\Diamond\neg p$ and $\Diamond\neg q$. The obvious culprit appears to be the De Morgan equivalence between $\neg(p \wedge q)$ and $\neg p \vee \neg q$, which are not equivalent on the unilateral model of the previous section or in standard inquisitive semantics. One way to modify the present system in order to invalidate dual free choice would be to change the rule for the negative part of conjunction, so that it is not equivalent to a disjunction of negations. Since this issue independent of the original puzzles and the data discussed in §3.4, I do not want to take a stand on it here. For simplicity and completeness, I have opted to validate all of the De Morgan equivalences. For further discussion of the De Morgan equivalences in modal and conditional contexts, see Fox (2007), Chemla (2009), Ciardelli, Zhang, and Champollion (2018), Romoli and Santorio (2019), and Marty et al. (ms.).

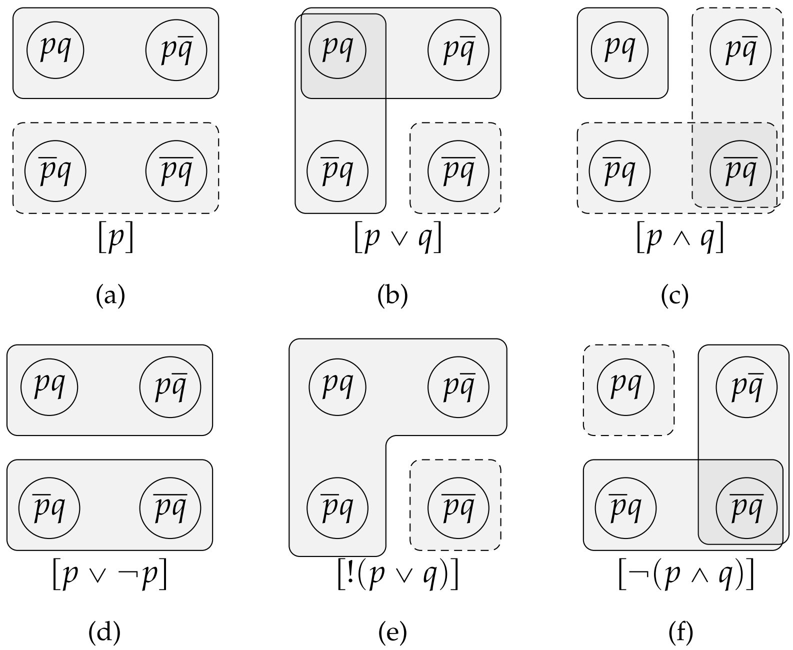

Figure 3.4: Examples of bilateral propositions

alternatives offered by the proposition:

$$\Box\phi \text{ is true at } w \text{ iff } \mathsf{alt}([\phi]^+) \text{ forms a minimal cover of } R(w)$$

$$\Diamond\phi \text{ is true at } w \text{ iff there is a non-empty } R' \subseteq R(w)$$
$$\text{such that } \mathsf{alt}([\phi]^+) \text{ forms a minimal cover of } R'$$

Ensuring duality means making $\neg\Box\phi$ equivalent to $\Diamond\neg\phi$, and $\neg\Diamond\phi$ equivalent to $\Box\neg\phi$. Since the negative alternatives of a proposition are identical to the positive alternatives of its negation (i.e. $\mathsf{alt}([\phi]^-) = \mathsf{alt}([\neg\phi]^+)$), duality leads to the following *falsity* conditions:

$$\Box\phi \text{ is false at } w \text{ iff there is a non-empty } R' \subseteq R(w)$$
$$\text{such that } \mathsf{alt}([\phi]^-) \text{ forms a minimal cover of } R'$$

$$\Diamond\phi \text{ is false at } w \text{ iff } \mathsf{alt}([\phi]^-) \text{ forms a minimal cover of } R(w)$$

Duality. Now let me unpack this semantics by showing how it recovers the truth conditional equivalence of (3.19a) and (3.19b). Recall those sentences:

(3.19) (3.19a) Alicia may not mail the letter or burn it $\neg\Diamond(M \vee B)$

(3.19b) Alicia must not mail the letter or burn it $\Box\neg(M \vee B)$

I will show that these sentences are true and false in the same circumstances. First, truth. By our semantics for negation, $\neg\Diamond(M \vee B)$ is true just in case $\Diamond(M \vee B)$ is false. $\Diamond(M \vee B)$ is false iff the negative alternatives offered by $M \vee B$ form a minimal cover of the set of relevant worlds. Since the negative alternatives offered by $M \vee B$ are equal to the positive alternatives offered by $\neg(M \vee B)$, it follows that $\Diamond(M \vee B)$ is false exactly whenever $\Box\neg(M \vee B)$ is true. Thus, $\neg\Diamond(M \vee B)$ is true iff $\Box\neg(M \vee B)$ is true.

Now for falsity. $\neg\Diamond(M \vee B)$ is false iff $\Diamond(M \vee B)$ is true, i.e. when the positive alternatives of $M \vee B$ form a minimal cover of a non-empty subset of relevant worlds. $\Box\neg(M \vee B)$, on the other hand, is false iff the negative alternatives of $\neg(M \vee B)$ form a minimal cover of a non-empty subset of relevant worlds. Since the negative alternatives of $\neg(M \vee B)$ are identical to the positive alternatives of $M \vee B$, the two sentences are false in exactly the same worlds.

By ensuring duality, we once again predict the validity of impossibility and unnecessity distribution over disjunction (3.17)-(3.18). Consider impossibility distribution, repeated here:

Impossibility Distribution.

(3.17a) Alicia may not mail the letter or burn it. $\neg\Diamond(M \vee B)$

(3.17b) So, Alicia may not mail the letter. $\neg\Diamond M$

(3.17c) So, Alicia may not burn the letter. $\neg\Diamond B$

By Duality, the premise $\neg\Diamond(M \vee B)$ is equivalent to $\Box\neg(M \vee B)$, which by the De Morgan laws is equivalent to $\Box(\neg M \wedge \neg B)$, which is true just in case $\mathsf{alt}([\neg M \wedge \neg B]^{+}) =$

$\{\mathsf{info}([M]^-) \cap \mathsf{info}([B]^-)\}$) is a minimal cover of $R(w)$. If that latter condition holds, then $\{\mathsf{info}([M]^-)\} = \mathsf{alt}([M]^-)$ is also a minimal cover of $R(w)$, meaning that $\Diamond(M)$ is false at w, so $\neg\Diamond M$ is true. Similarly for $\neg\Diamond B$.

A similar argument works for unnecessity distribution, repeated here:

Unnecessity Distribution.

(3.18a) Alicia doesn't have to either mail the letter or burn it. $\neg\Box(M \vee B)$

(3.18b) So, Alicia doesn't have to mail the letter. $\neg\Box M$

(3.18c) So, Alicia doesn't have to burn the letter. $\neg\Box B$

When (3.18a) is true, $\Box(M \vee B)$ is false. It follows that the negative alternatives of $M \vee B$, i.e. the singleton containing all worlds where both M and B are false, cover a non-empty set of relevant worlds. If that's the case, then the larger set of worlds where just M is false *also* covers that set. The same goes for the larger set of worlds where just B is false. But then, this means that $\Diamond\neg M$ and $\Diamond\neg B$ are true. By Duality, these are equivalent to $\neg\Box M$ and $\neg\Box B$.

Moving to this bilateral system thus allows us to regain the orthodox predictions about the truth conditions of disjunctive modal claims embedded under negation and other downward entailing operators.

3.8 Truth Value Gaps

Adopting the bilateral framework I have outlined here requires accepting some truth value gaps, but I think that they are well placed.[34] First, if we assume atomic formulae are either true or false in every world, then it is only sentences containing modals that generate truth value gaps. Further, there are no truth value gluts. So, the bilateral system ensures that the non-modal fragment behaves classically.

34. The bilateral, dynamic inquisitive semantics of Willer (2018) also postulates some truth value gaps for modal sentences.

In semantics, truth value gaps are most often taken to indicate presupposition failure. But other kinds of gaps have been postulated. In particular, predicates of plurals seem to generate gaps as a result of 'homogeneity' effects.[35] Consider the following pair:

(3.20) a. Alicia and Bulmaro saw the movie. They liked it.

b. Alicia and Bulmaro saw the movie. They did not like it.

The second sentence of (3.20a) is true iff *both* Alicia and Bulmaro liked the movie; while its negation in (3.20b) is true iff *neither* Alicia nor Bulmaro liked it. There is a gap between these truth conditions: if Alicia but not Bulmaro liked the movie, both (3.20a) and (3.20b) seem to misdescribe the situation. On some theories of plural predication, this is because the second sentence of each example is neither true nor false in that situation. The truth value gap arises because the predicate 'liked it' expects its plural arguments to be homogeneous with respect to it: either *all* or *none* of the individuals in a collection satisfy it.

For possibility modals, there appears to be a similar kind of truth value gap: when $\Diamond(p \vee q)$ is true, p and q are both possible (by free choice). When $\Diamond(p \vee q)$ is false (i.e. when $\neg\Diamond(p \vee q)$ is true), p and q are both *impossible* (by impossibility distribution). If that is correct, then $\Diamond(p \vee q)$ is neither true nor false when only one of p, q is possible. And this is exactly what my account predicts. In this respect, my account concurs with Goldstein (2019b)'s theory of possibility modals, which writes homogeneity directly into their semantics. In contrast to Goldstein's homogeneity account, my theory posits a second source of truth value gap for sentences like $\Diamond(p \vee q)$. On my account, $\Diamond(p \vee q)$ is also neither true nor false in case p and q are both possible but not *independently* so: that is, when $\Diamond p$ and $\Diamond q$ are true, but either $\Diamond(p \wedge \neg q)$ or $\Diamond(q \wedge \neg p)$ is not. This second source of truth value gap helps explain the anomalous status of (3.12) in the described scenario. To recall, you are playing a card game, and your only option is to pick up the

35. See for example, Schwarzschild (1993) and Križ (2015, 2016). It is controversial whether homogeneity effects give rise to gaps, and if they do, whether these gaps should be thought of as presuppositions or not.

pair consisting of the ace and the ten of diamonds. Someone says:

(3.12) You may either take the ace of diamonds or the ten of diamonds.

(3.12) seems to misdescribe things — it appears to say that there are at least two (independent) options. But its negation seems equally off the mark:

(3.12′) You may not take the ace of diamonds or the ten of diamonds.

My bilateral semantics can explain why both (3.12) and (3.12′) seem to misdescribe things: neither of them are true.

For necessity modals, truth value gaps arise precisely in the cases at issue in Ross's puzzle. Suppose 'Alicia ought to mail the letter' ($\Box M$) is true at w. The truth of the Ross inference conclusion, 'Alicia ought to either mail the letter or burn it' ($\Box(M \vee B)$) requires that M and B are independent alternatives in the relevant set of worlds. If the Ross premise $\Box M$ is true, then B is not an independent alternative, so $\Box(M \vee B)$ is not true. But neither is it false. By duality and one of the De Morgan equivalences, it is false just in case $\Diamond(\neg M \wedge \neg B)$ is true. The truth of that latter sentence obviously requires that there are relevant not-M-worlds. But since $\Box M$ is true, there are none. My account thus predicts that the premise does not merely *fail* to ensure the truth of the conclusion; it ensures the conclusion is *not* true.[36]

If my account is correct in this prediction, then it offers a diplomatic resolution of the disagreement over Ross's puzzle between orthodox and revisionary semantic accounts. Proponents of revisionary semantic accounts have insisted that $\Box M$ does not entail $\Box(M \vee B)$. Defenders of the orthodox semantics sense that it must, since $\neg\Box(M \vee B)$ entails $\neg\Box M$. If these sentences are always true or false, then at most one of these claims can be correct. But my account rejects this assumption. In particular, when $\Box M$ is true, $\Box(M \vee B)$ is neither true nor false, so $\Box M$ does not entail $\Box(M \vee B)$, even though $\neg\Box(M \vee B)$ entails $\neg\Box M$. Thus, we may agree with both parties on these points.

36. This may be why Ross inferences are so strongly repugnant: the premise not only fails to ensure the truth of the conclusion: it ensures the conclusion is *neither true nor false*.

3.9 Conclusion: Relation to Collectivity

I have argued that both Ross's puzzle and the puzzle of free choice permission should be explained in terms of the licensing of *Independence inferences*, which are stronger than the standard Diversity analysis predicts. I have also given a bilateral inquisitive semantics for the interaction between modals and disjunctions that generates the validity of these inferences, and showed how it can explain the full range of data discussed in this chapter.

The bilateral minimal covering semantics was motivated by the data concerning the status of Extended Ross Arguments and Independence Conditional Arguments. Over and above this data, however, it would be desirable to have some explanation for *why* the Independence inferences are licensed. In this section, I want to propose what I think is a plausible interpretation of the semantic interaction between modals and disjunctions that might provide such an explanation. In particular, I want to suggest that we might think of disjunctions as denoting something like *pluralities* of propositions, and that we might think of modals as behaving like *collective predicates* of these pluralities.

Connections between puzzles surrounding the disjunction-modal interaction and the plural term-predicate interaction have been suggested before.[37] To some extent, this is unsurprising, since most semantic solutions to the puzzles surrounding the disjunction-modal interaction adopt semantic frameworks in which a disjunction denotes something like a set of multiple alternatives, each corresponding to a proposition denoted by a disjunct. Most treatments of plural noun/determiner phrases adopt a framework in which such terms denote sets (or sums) of multiple individuals.[38] The two frameworks thus make comparison tempting. Those who have gone in for the comparison, however, have usually focused on comparing the logic of possibility modals and disjunctions, on

37. See, for example, Simons (2005b) which draws the connection to homogeneity in the plural domain, Goldstein (2019b) which works out this connection systematically, and Santorio (2018) for connections between disjunctions, modal accounts of conditionals, and the logic of plurals.

38. See for example, Link (1983) and a vast amount research following.

the one hand, with *distributive* predicates and plural terms, on the other (see the works cited in fn. 37). Since necessity modals do not distribute over disjunctions, the focus on distributivity ensures that the comparison between the two domains is narrow. I wish to expand the comparison by considering non-distributive, or *collective*, predicates, and to draw some analogies with the semantics of modals.[39]

To begin, I want to note that there is a connection between the notion of a minimal cover and the notion of a state that obtains only at the *collective* or *group* level. Suppose C is a set of several elements, and that C is a minimal cover of S. Then there is a sense in which C's covering of S obtains essentially collectively: it is accomplished by the elements of C *only* when they are taken *together*. Another way to put this is that each member of C plays a necessary role in making it so that C is a cover of S. This idea — that each member of the group must play a necessary role for the performance of an action or the obtaining of a state to be *collective* — has been a common way of cashing out the notion of collectivity in distinct literatures: for example, on the logic of agency (see the notion of 'strictly stit' in Belnap and Perloff 1993, 41, explored further by Sergot (2021)); and on the semantics of 'together' (see, e.g., Lasersohn (1990) and Schwarzschild (1993), and Moltmann (2004) for criticism and an alternative).

Additionally, notice that while semantic solutions to Ross's Puzzle that adopt the Diversity Analysis must make reference to the multiple *alternatives* denoted by a non-Hurford disjunction $p \vee q$, they still allow that the set of relevant worlds contains a *single* world witnessing these multiple alternatives. If $R = \{w\}$ and p and q are both true at w, then $\Box(p \vee q)$ is true and the Diversity inferences are supported. A minimal covering semantics like the one I outlined in the last section, by contrast, requires that when $\Box(p \vee q)$ is true, there are at least *two* worlds in R: one which makes p but not q true, and one which makes q but not p true. In this sense, then, a minimal covering semantics requires that the set of relevant worlds with respect to which $\Box(p \vee q)$ is evaluated

39. 'Collective' is used in many different ways in the literature on plurals. Here I adopt one of its weakest uses: to pick out predicates that are not as a rule distributive when applied to plural-denoting terms.

contains as many worlds as there are alternatives offered by the complement clause. This means that only licensing the *Independence* inferences ensures that the 'plurality' of the embedded disjunction is mirrored by a 'plurality' of relevant worlds.

Besides these conceptual connections between a minimal covering semantics, collectivity, and plurality, I want to note that the logic of sentences like $\Box(p \vee q)$ shares some characteristic marks of collective predicates applied to plural terms.

First, collective predicates applied to plural terms are often non-distributive (indeed, this is sometimes taken as the hallmark of collectivity):

(3.21) a. Alicia and Bulmaro performed *Happy Days*. $P(a+b)$

b. $\not\Rightarrow$ Alicia performed *Happy Days*. $P(a)$

c. $\not\Rightarrow$ Bulmaro performed *Happy Days*. $P(b)$

Similarly, when sentences of the form $\Box(p \vee q)$ are true, the necessity modal does not distribute over its embedded disjuncts:

(3.22) a. Pim must trudge or talk. $\Box(p \vee q)$

b. $\not\Rightarrow$ Pim must trudge. $\Box(p)$

c. $\not\Rightarrow$ Pim must talk. $\Box(q)$

Given the validity of free choice inferences, of course, possibility modals *do* distribute over disjunction. But this does not mean that they do not behave like collective predicates, since some 'mixed' predicates, which can be true either collectively or distributively, do distribute:[40]

(3.23) a. The group [Alicia and Bulmaro] weighs less than 500 pounds.

b. $\Rightarrow$ Alicia weighs less than 500 pounds.

c. $\Rightarrow$ Bulmaro weighs less than 500 pounds.

40. This is one reason that (Moltmann 2004) argues we should move away from thinking of collectivity as anti-distributivity.

Second, even though collective predicates do not usually distribute over their plural terms, there are usually some 'involvement' or 'participation' inferences one can derive that *do* distribute (Dowty (1987) and Link (1983))

(3.24) a. Alicia and Bulmaro performed *Happy Days*. $P(a+b)$

b. $\Rightarrow$ Alicia played a role in a performance of *Happy Days*. $P^*(a)$

c. $\Rightarrow$ Bulmaro played a role in a performance of *Happy Days*. $P^*(a)$

Similarly, the Independence inferences $\Box(p \vee q)$ licenses can be thought of as the inference that p and q each are represented in, or participate in covering, the relevant set of worlds:

(3.25) a. Pim must trudge or talk. $\Box(p \vee q)$

b. $\Rightarrow$ Pim may trudge (without talking). $\Diamond(p \wedge \neg q)$

c. $\Rightarrow$ Pim may talk (without trudging). $\Diamond(q \wedge \neg p)$

Third, it is plausibly precisely because of such involvement inferences that predicates that can apply collectively tend to exhibit what I will call *upward failure*:

(3.26) a. Alicia lifted the piano. $P(a)$

b. $\not\Rightarrow$ Alicia and Bulmaro lifted the piano. $P(a+b)$

The analogous inference for disjunctions and modals is just a Ross inference, which also fails:

(3.27) a. Alicia must lift the piano. $\Box p$

b. $\not\Rightarrow$ Alicia or Bulmaro must lift the piano. $\Box(p \vee q)$

(3.28) a. Alicia may lift the piano. $\Box p$

b. $\not\Rightarrow$ Alicia or Bulmaro may lift the piano. $\Box(p \vee q)$

Fourth, both collective predicates applied to plural terms and modals applied to disjunctions exhibit strong negations (the negation of ϕ entails but is not equivalent to the failure of the truth of ϕ). Consider the case of a collective predicate applied to a plural term:

(3.29) a. Alicia and Bulmaro didn't perform *Happy Days*. $\neg P(a + b)$

b. $\Rightarrow$ Alicia didn't perform *Happy Days*. $\neg P(a)$

c. $\Rightarrow$ Bulmaro didn't perform *Happy Days*. $\neg P(b)$

(3.29a), the negation of (3.24a), does not seem to allow that Alicia but not Bulmaro performed the play, which would be expected if the negation were simply the denial of the truth. Rather, it seems to require that neither Alicia nor Bulmaro performed the play.

Likewise, $\neg\Box(p \vee q)$ does not allow that p is necessary but not q as would be expected if it simply required that $\Box(p \vee q)$ was not true. Rather, it seems to make the stronger claim that neither p nor q is necessary:

(3.30) a. Pim does not have to either trudge or talk. $\neg\Box(p \vee q)$

b. $\Rightarrow$ Pim does not have to trudge. $\neg\Box p$

c. $\Rightarrow$ Pim does not have to talk. $\neg\Box q$

To summarize, I think a number of features common to the modal-disjunction interaction, on the one hand, and the collective predicate-plural term interaction, on the other, recommend that we interpret the bilateral minimal covering semantics as a framework in which disjunctions denote *pluralities* of propositions, and necessity modals with disjunctive complements behave like *collective predicates* applied to those pluralities. Of course, there are important differences between the two domains, and these differences could ultimately strain the analogy. Still, the several respects in which they are similar means, I believe, that it may be fruitful to explore the comparison in more detail in future research.

Chapter 4: The Samaritan Paradox

4.1 Introduction

Natural language necessity modals like English 'must', 'should', and 'ought', seem to allow us to speak truly even when we underspecify what is necessary. For example, if my obligations make it true that *I must bake a gluten-free cookie*, then in situations where the recipe is irrelevant, I might simply tell you that *I must bake a cookie*. This latter sentence is less specific than the former about what I have to do, but it is not for that reason false or merely figurative. This is clearly a useful feature of the semantics of ordinary language necessity modals: often our obligations are specific enough that having to spell them out completely every time we make reference to them would take up a lot of time.

Standard semantic theories of necessity modals luckily predict that underspecific descriptions of what is necessary are true. On these theories, the possibility of underspecification is underwritten by a property called *upward monotonicity*: if a sentence ϕ entails a sentence ψ, then where '$\Box$' is an arbitrary necessity modal, $\ulcorner \Box\phi \urcorner$ entails $\ulcorner \Box\psi \urcorner$. As applied to the above example, I am able to truly and underspecifically describe my obligations by saying 'I must bake a cookie' since the sentence 'I bake a gluten-free cookie' entails the sentence 'I bake a cookie'.

Despite everything in favor of monotonicity, there are some well-known and puzzling counterexamples. This chapter is centrally concerned with one in particular, which trades on presuppositional features of natural language sentences. Generally, the phrase 'the person who was robbed' is thought to carry the presupposition that there is a unique salient or familiar person who was robbed, and without the interference of other phrases or pragmatic factors, sentences that contain this phrase inherit this presupposition. Among other things, this means that such sentences will also *entail* that there is such a person.[1]

(4.1) a. Someone helped the person who was robbed.

b. So, there was a person who was robbed. ✓

Prior (1958) noticed that a consequence of this is that upward monotonic semantics for necessity modals will validate patterns of inference like the following:

(4.2) a. It ought to be that someone helps the person who was robbed.

b. So, it ought to be that there is a person who was robbed. ✗

(4.3) a. Given the moral law, Alice must care for her hamster.

b. So, given the moral law, Alice must have a hamster. ✗

Reconciling upward monotonicity with the intuition that these latter inference is invalid is the challenge known as the *Paradox of the Good Samaritan*, or, for brevity, the *Samaritan Paradox*.

4.2 Background

In this section, I explain in more detail why the Samaritan paradox arises on standard semantics for necessity modals; in particular, for the theories of Saul Kripke and Angelika Kratzer. I then outline what I consider the best candidate for a 'received' solution to the paradox in natural language semantics, based on Kratzer's semantics for modals.

On the Kripke semantics for necessity modals, claims of the form $\ulcorner\Box\phi\urcorner$ are true at a world w in a model $\mathcal{M}$ (written $\mathcal{M}, w \vDash \Box\phi$) when the subset relation holds between the set of worlds accessible from w (R_w) and the set of worlds that make the prejacent clause true relative to the model ($[\phi]_{\mathcal{M}}$). Thus, given a model $\mathcal{M} = \langle W, R, V\rangle$ (where W is a set of worlds, R is a function taking each world to a set of accessible worlds R_w, and V is

1. Of course, on a Russellian theory of definite descriptions, a simple negation of such a sentence will not carry this entailment. It is enough for my purposes, however, that the positive form does.

a valuation function assigning truth-sets to atomic formulas) and a world in that model $w \in W$:[2]

Definition (Kripke Semantics)**.**

$$\mathcal{M}, w \vDash \Box\phi \quad \text{iff} \quad R_w \subseteq [\phi]_{\mathcal{M}}$$

Kripke's semantics for necessity modals makes them *upward monotonic*:

Definition (Upward Monotonicity)**.** A propositional operator Δ is upward monotonic iff for any proposition-denoting terms ϕ and ψ:

$$\text{If } \phi \text{ entails } \psi\text{, then } \ulcorner\Delta(\phi)\urcorner \text{ entails } \ulcorner\Delta(\psi)\urcorner$$

On the Kripke semantics, the relevant set of worlds R_w for evaluating a modal claim $\Box\phi$ is fixed by the world of evaluation and the model. This means that in order to model different 'flavors' of modality — epistemic, deontic, logical, etc. — our language will need to include separate operators for each, and different accessibility relations corresponding to them. In natural languages, however, there is a cross-linguistically attested ability to use the same words, e.g., English 'have to', 'must', 'ought', 'should', to express different flavors of modality on different occasions. In order to account for this variation, the dominant approach to natural language necessity modals like 'must', 'ought', 'have to', etc., due to Kratzer (1977, 1991, 2012d) proposes that instead of being tied to their lexical semantics, the relevant sets of worlds for such modals are determined by the context of utterance. Thus, on the Kratzer semantics, a model $\mathcal{M} = \langle W, V \rangle$ is simpler: it is just a pair of a set of worlds W and a valuation function assigning sets of worlds to atomic propositions V.[3] Formulas are evaluated relative to both a model and a *context* that determines which worlds are relevant for the evaluation of a modal claim like $\ulcorner\Box\phi\urcorner$.

2. See, for example, Kripke (1963) and Chellas (1980)
3. I should note that Kratzer does not actually discuss models explicitly.

Besides making modals sensitive to the context, the Kratzer semantics splits the work of determining which worlds are relevant or accessible into two separate contextual parameters: a *modal base*, and an *ordering source*. A modal base at a world $f(w)$ is a set of propositions, and the semantics is partly a function of their *conjunction* ($\bigcap f(w)$). Since we will mostly be concerned with this conjunction, I will often refer to *it* as the 'modal base', rather than $f(w)$ proper. An ordering source at a world ($g(w)$) is a set of propositions that determines a partial order over worlds:

$$w' \geqslant_{g(w)} w'' \Leftrightarrow \forall p \in g(w): \text{ if } w'' \in p \text{ then } w' \in p$$

This partial order in turn determines a set of *best* or *undominated* worlds within a given particular set of worlds S, which the semantics relies on:

$$\mathsf{best}(g(w), S) := \{w' \in S \mid \text{ for all } v \in S: \text{ if } v \geqslant_{g(w)} w' \text{ then } v = w'\}$$

Putting these elements together, we can state the Kratzer semantics as follows:[4]

Definition (Kratzer Semantics)**.**

$$\mathcal{M}, f, g, w \vDash \Box\phi \quad \text{iff} \quad \mathsf{best}(g(w), \bigcap f(w)) \subseteq [\phi]_{\mathcal{M},\langle f,g\rangle}$$

Despite the complexity that the Kratzer semantics introduces, it retains a (contextually-sensitive) version of the upward monotonicity feature of the Kripke semantics:[5]

Definition (Contextual Upward Monotonicity)**.** A propositional operator Δ is upward monotonic iff for any proposition-denoting terms ϕ, ψ, and any context $\langle f, g\rangle$:

If ϕ entails ψ at a context $\langle f, g\rangle$, then $\ulcorner\Delta(\phi)\urcorner$ entails $\ulcorner\Delta(\psi)\urcorner$ at $\langle f, g\rangle$.

4. Where $[\phi]_{\mathcal{M},\langle f,g\rangle}$ is the set of worlds w that make ϕ true relative to $\mathcal{M}, f, g$.

5. If ϕ entails ψ, then $[\phi]_{\mathcal{M},\langle f,g\rangle} \subseteq [\psi]_{\mathcal{M},\langle f,g\rangle}$. Thus, for any point $\mathcal{M}, f, g, w$, if $\mathsf{best}(g(w), \bigcap f(w)) \subseteq [\phi]_{\mathcal{M},\langle f,g\rangle}$, it follows that $\mathsf{best}(g(w), \bigcap f(w)) \subseteq [\psi]_{\mathcal{M},\langle f,g\rangle}$. So if $\Box\phi$ is true, so is $\Box\psi$.

All of the Samaritan inferences we have so far considered trade on sentences related by entailment, even with the context fixed. Thus, even though the Kratzerian theory gives rise to a slightly more complex version of upward monotonicity, it still generates all of the counterintuitive predictions we have discussed.

4.3 The 'Received' Solution

The Samaritan Paradox has received much discussion in the literature on deontic logic, but the paradox also arises for attitude verbs that seem to have nothing to do with deontic notions like obligation, norms, or even actions and the future. Consider, for example, 'weak speech reports', which also seem to be upward monotonic (Brasoveanu and Farkas 2007; Abreu Zavaleta 2019), but which do not validate Samaritan inferences:

(4.4) a. Anna said her wardrobe is spacious and clean.

b. So, Anna said her wardrobe is spacious. ✓

c. Anna said she has a wardrobe. ✗

(4.4a) seems to entail (4.4b) but not (4.4c). Since the Samaritan Paradox seems to arise for verbs that have nothing to do with normative or deontic issues, it is somewhat surprising that it has received more attention in deontic logic than in natural language semantics. To the extent that there is a 'received solution' in natural language semantics, it must be cobbled together from various sources and to some extent, extrapolated. See Heim (1992) for a view much like the one I outline here applied to the paradox as it arises for 'want'; see von Fintel and Heim (2011) for the suggestion that something like Heim's (1992) approach should apply more generally; see Portner (2009) for suggestive comments; and Carr (2014) for discussion of similar views on a closely related paradox that arises for conditionals. With these caveats, what I will call the 'received' Kratzerian solution says that necessity modals like 'ought' and 'must' carry two presuppositions concerning the modal base. First, that the presuppositions of the prejacent clause are true throughout the modal base. Call this the *Satisfaction Presupposition*. Second, necessity modals presuppose that the modal base includes both worlds where the prejacent clause is true and worlds where it is not. Call this the *Divergence Presupposition*. So, where $\mathsf{dom}(\phi)$ (the *domain* of ϕ) is the set of worlds that satisfy the presuppositions of ϕ:

$$\bigcap f(w) \subseteq \mathsf{dom}(\phi) \qquad (\textit{Satisfaction})$$

$$\bigcap f(w) \cap [\phi] \neq \varnothing \text{ and } \bigcap f(w) \cap [\neg\phi] \neq \varnothing \qquad (\textit{Divergence})$$

Let me illustrate how these presuppositions are meant to solve the puzzle. The complement of (4.2a) includes the phrase 'the person who was robbed', due to this, the sentence as a whole presupposes that there was someone who was robbed. By Satisfaction, this means that (4.2a) as a whole presupposes there was someone robbed in every world in $\bigcap f(w)$. Furthermore, by Divergence, (4.2a) presupposes that there are worlds in the modal base where someone helps the person who was robbed and also worlds where it is not the case that someone helps the person who was robbed. Finally, if (4.2a) is true, then both presuppositions are satisfied and all of the maximally preferred worlds given g and $f(w)$ are worlds where someone helps the person who was robbed. Now consider (4.2b). By Divergence, (4.2b) presupposes that the modal base with respect to which it is evaluated, $\bigcap f'(w)$ includes some worlds where there is no person who was robbed. Since all worlds in $\bigcap f(w)$ are worlds where someone was robbed, it follows that $\bigcap f'(w) \neq \bigcap f(w)$. In other words, (4.2a) and (4.2b) cannot be true with respect to the same modal base. If (4.2a) is true, (4.2b) cannot be true — it suffers from presupposition failure. More generally, the Satisfaction presupposition of (4.2a) ensures that the Divergence presupposition of (4.2b) fails. In this sense, then, the argument is invalid.[6]

6. Of course, relative to some other notions of validity, such as 'Strawson entailment' (von Fintel 1999), the argument will be valid since there is no point of evaluation where the premise is true and the conclusion is false. In my opinion, cases like these, where the truth of the premise ensures that the conclusion is not true, provide good reason to think that Strawson entailment is not the appropriate notion of validity for natural language. See Sharvit (2017) for an independently motivated refinement of Strawson entailment that avoids counting arguments of this type as valid.

Although this hypothesis gets the correct result in cases like (4.3), I will now present data to argue that it is too strong, given two background assumptions about how presuppositions are determined. The first assumption is that presuppositions introduced by words like 'must' and 'ought' are determined by the lexical semantics of these words, not by context or other factors. The second assumption is that the difference between deontic and epistemic flavors of these modals is determined by different properties of the modal base and/or ordering source.[7] For example, if 'must' on epistemic, but not deontic, readings validates the (M) schema — i.e. Must(ϕ), therefore ϕ — this can be explained by appeal to the kinds of modal bases and ordering sources available on epistemic or deontic interpretations. In particular, constraints on epistemic parameters may ensure that for any world w, w is among the best worlds (i.e. $w \in \mathsf{best}(g(w), \bigcap f(w))$). If that is so, then the (M) schema is valid on epistemic readings of 'must.' By contrast, if deontic parameters allow that for some worlds w, w is not among the best worlds (i.e. $w \notin \mathsf{best}(g(w), \bigcap f(w))$), then the (M) schema is not valid for deontic readings of 'must.' This is perhaps one of the principle virtues of the Kratzerian, contextualist analysis of necessity modals — it gives a parsimonious account of how the same modal can take on various flavors of meaning and license different inferences without positing rampant ambiguity.

Together, these assumptions mean that if necessity modals carry the Divergence and Satisfaction Presuppositions, they do so for both deontic and epistemic interpretations. This leads to the following prediction: *ceteris paribus*, the status of Samaritan inferences should be the same between deontic and epistemic readings of the modals. This prediction, however, does not seem to be borne out. Consider the following minimal pairs. Suppose we are explaining the societal norms of Anna's time:

(4.5) Given the norms of Anna's milieu:

7. This assumption has been criticized, but see Hacquard (2010) for what I believe is a compelling response.

a. Anna had to love her sister.

b. So, Anna had to have a sister. ✗

Anna's being obliged to love her sister does not, I think, entail that she was obliged to *have* a sister. As just explained, the received solution to the Samaritan paradox predicts this correctly: the truth of (4.5a) requires that the modal base $\bigcap f(w)$ satisfies the presupposition that Anna has a sister (Satisfaction). Assuming (4.5b) is evaluated with respect to the same modal base, this ensures that the Divergence presupposition of (4.5b) fails. Since it suffers from presupposition failure, it is not true. So far, so good. But now suppose we have been conducting an investigation into facts about Anna, and we state our conclusions:

(4.6) Given all that we have established about Anna at the time:

a. Anna had to love her sister.

b. So, Anna had to have a sister. ✓

In my judgment, this inference seems valid. Epistemic readings of other modals also seem to validate Samaritan inferences. Consider:

(4.7) a. Given what we knew yesterday, he must have already caught the murderer.

b. So, given what we knew yesterday, there must have been a murderer. ✓

(4.8) a. Given what we know, Maxwell should have cleaned his silver hammer.

b. So, given what we know, Maxwell should have a silver hammer. ✓

The received solution to the Samaritan paradox does not predict that these inferences are valid. If (4.6a) is true, this requires that the modal base $\bigcap f(w)$ satisfies the presupposition that Anna has a sister (Satisfaction). Assuming (4.6b) is evaluated relative to the same modal base, this ensures that the Divergence presupposition of (4.6b) fails. Since it suffers from presupposition failure, it is not true.

In fact, this points to a respect in which the problem is even more pressing for the received solution. It does not merely predict that epistemic Samaritan inferences are *invalid*; it predicts that the premise and the conclusion are *incompatible* in the sense that they cannot both be true at the same context.

To add some indirect support to this data, we may observe a parallel difference in the status of Samaritan inferences between cognitive and conative attitude verbs, which are often theorized in tandem with epistemic and deontic readings of modals. Cognitive attitude verbs often license Samaritan inferences,[8] as illustrated by the following examples:[9]

(4.9) a. Anna believes that her wardrobe is clean.

b. So, Anna believes that she has a wardrobe. ✓

(4.10) a. Anna knows that her wardrobe is clean.

b. So, Anna knows that she has a wardrobe. ✓

Conative attitude verbs, by contrast, often do not license Samaritan inferences, as illustrated by the following examples:[10]

(4.11) a. Anna wants to clean her wardrobe.

b. So, Anna wants to have a wardrobe. ✗

(4.12) a. Anna hopes to get rid of her wardrobe.

b. So, Anna hopes to have a wardrobe. ✗

In my opinion, the difference that we observe in the status of Samaritan inferences for different kinds of attitude verbs mirrors the status of such inferences for differently flavored readings of traditional modals like 'must,' 'have to,' 'should,' and 'ought.'

The lesson I will draw from these examples is that whether or not a necessity modal licenses a Samaritan inference depends on its flavor. Generally speaking, deontically

8. There are some interesting exceptions: *Anna sees that her wardrobe is clean* does not appear to entail that *Anna sees that she has a wardrobe*.

9. Of course, drawing conclusions from this data requires that we control for some confounding factors, such as *de re* readings of phrases that include presupposition triggers.

10. Here again, there are some interesting exceptions. To my ear, the following inference is neither obviously correct nor incorrect: *Anna plans to clean her wardrobe*, so *Anna plans to have a wardrobe*.

flavored readings seem to invalidate Samaritan inferences, while epistemically flavored readings seem to validate them. If that is correct, it poses a problem for the received solution to the Samaritan paradox, which invalidates all Samaritan inferences by virtue of the lexical semantics of modals, independent of the flavor of interpretation they receive in a context. What we need, instead, is a theory that treats Samaritan inferences rather like the (M) schema discussed above: one on which modals have a single lexical semantics, and either validate or invalidate Samaritan inferences depending on features of the contextual parameter(s) their semantics is sensitive to.

4.4 Outline of a new solution

The Kratzerian solution to the Samaritan paradox in the deontic case works because the semantics makes a three-way distinction between worlds relevant for the evaluation of $\ulcorner\Box\phi\urcorner$. First, there are the worlds in the modal base $\bigcap f(w)$ that are undominated with respect to $g(w)$, i.e. the worlds in $\mathsf{best}(g(w), \bigcap f(w))$. To abstract from the details of how the Kratzerian semantics makes the three-fold distinction, call these the 'good' worlds. The truth of $\ulcorner\Box\phi\urcorner$ requires that the good worlds satisfy the presuppositions of ϕ and make ϕ true. Then there are the worlds in the modal base that are *dominated*: $\bigcap f(w) \backslash \mathsf{best}(g(w), \bigcap f(w))$ — call these the 'bad' worlds. Given the Divergence presupposition, the truth of $\ulcorner\Box\phi\urcorner$ requires that the bad worlds include some world where ϕ is false. Finally, there are worlds that are outside the modal base altogether: $W \backslash (\bigcap f(w))$ — call these the 'irrelevant' worlds.

Allow me the use of subscripts to syntactically represent presuppositions. ϕ_ψ is the sentence with the truth conditions of ϕ and which presupposes the truth of ψ. Let 'p' abbreviate 'Anna has a parking ticket' and 'q' abbreviate 'Anna pays any parking tickets she has'. Then 'q_p' stands for something like, 'Anna pays her parking ticket'; it suffers from presupposition failure when Anna has no parking ticket, is true iff she has a parking ticket and pays it, and is false if she has a parking ticket but does not pay it.

Notice that we have not said whether q_p is false (as on a bivalent account) or neither-true-nor-false (as on a trivalent account) when it suffers from presupposition failure, since I take the theory I am developing here to be compatible with many different ways of modelling presuppositions.

On the received solution, when you know that $\Box(q_p)$ is true, you know that satisfying the relevant norm(s) — being in a good world — requires that p and q are both true. If Anna must pay her parking ticket, then satisfying the norm is possible only if Anna has a parking ticket and she pays it. The insight of the received view of the deontic case is that it also says that if you know that $\Box(q_p)$ is true, then given the Divergence presupposition, you *also* know that it is possible to *fail* to satisfy those norms — to be in a bad world — in case p is true and q is false. If Anna must pay her parking ticket, then having a parking ticket but *not* paying it is a way of frustrating the body of norms given by the ordering source: it is a way of being in a bad world. When $\Box(q_p)$ is true in this case, however, worlds where p is false — where Anna doesn't have a parking ticket — are irrelevant; the relevant body of norms ($g(w)$) and body of information ($\bigcap f(w)$) that require q_p don't apply. Since the truth of $\Box p$ requires that it is possible (given the relevant information in the modal base) to frustrate the relevant norms by making p false, but it is not, $\Box p$ is not true (as illustrated in Figure 1).

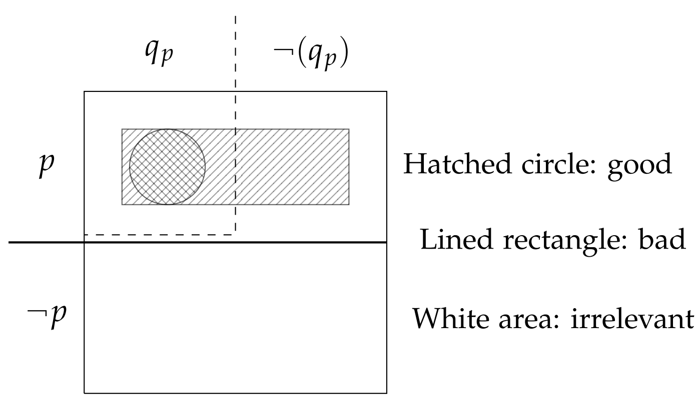

Figure 4.1: $\Box(q_p)$ is true (since there is a bad $\neg(q_p)$-world); but $\Box p$ is not true (since there is no bad $\neg p$-world). Note that it is allowed that not every p-world is relevant for $\Box(q_p)$ to be true

Now, suppose that *no* worlds were irrelevant, and in particular, that all the formerly irrelevant worlds were *bad* worlds. Then $\Box(q_p)$ is still true, since all good worlds are q_p-worlds, and that there is still a bad world where q_p is false. What about $\Box p$? Since all the good worlds are q_p-worlds, they are also p-worlds. Furthermore, since the $\neg p$-worlds are now bad worlds, there is a bad world where p is false. So $\Box p$ is true. Thus, whether

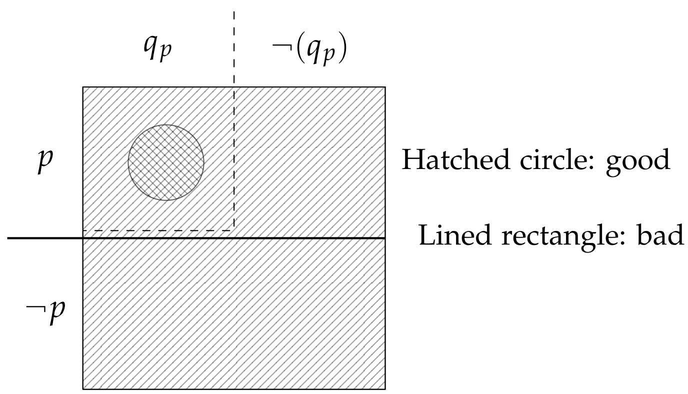

Figure 4.2: $\Box(q_p)$ and $\Box p$ are true (there are bad $\neg(q_p)$- and $\neg p$-worlds).

or not a Samaritan inference is valid with respect to a point of evaluation depends, in part, on the boundary between bad and irrelevant worlds.

This is basis of the solution to the Samaritan Paradox I wish to propose in this chapter. We will say, like the Kratzerian semantics, that modals are evaluated with respect to parameters that may make a threefold distinction between worlds. When a reading of a necessity modal validates a Samaritan inference, as in the case of our epistemic examples (4.6-4.8) above, it does so because the parameter divides the domain of worlds into just two categories: the good and the bad (as in Figure 2). When a reading does not support a Samaritan inference, as in the case of our deontic examples (4.5) above, it is because the parameter divides the domain of worlds into three categories: the good, the bad, and the irrelevant (as in Figure 3).

In my view, this is a philosophically appealing solution. Suppose *A* has normative authority over *B* and gives the following order to *B*:

(4.13) Take off your fedora!

If this is all *A* says, we can characterize the content of the normative requirements placed on *B* as follows. They are satisfied only if *B* takes off their fedora, frustrated if *B* does not take off their fedora, and irrelevant or moot if *B* has no fedora.[11]

Now suppose *A* tells *C*:

(4.14) *B* must be wearing their fedora.

If this is all *A* says, we can characterize the content of the information *A* represents themselves as possessing to as follows: it is information that is correct iff *B* is wearing their fedora. Having this information necessarily involves having the information that *B has* a fedora. Thus, the body of information *A* possesses is correct iff *B* has a fedora and is wearing it; it is incorrect if *B* is not wearing their fedora, but it is *also* incorrect if *B* does not have a fedora.

The way I described the relevant contents in the two cases assumes a sort of asymmetry in line with what I have been suggesting: bodies of norms often make a threefold distinction between worlds, whereas bodies of information make a twofold distinction. Of course, I could have described things differently. On the one hand, I could have said that the command given in (4.13) was satisfied iff *A* takes off their fedora, and *not* satisfied otherwise. That would be to emphasize that at some level, both kinds of contents make only a twofold distinction — call this the symmetric-2 description. On the other hand, I could have said that the information *A* represented themselves as possessing was correct iff *B* is wearing their fedora, incorrect iff *B* is not wearing their fedora, and *moot* if *B* does not have a fedora. That would be to emphasize that at some level, both kinds of contents make a threefold distinction. Call this the symmetric-3 option. Why should we privilege asymmetric way I initially described the situation over these symmetric

11. This is not to say that worlds where *B* has no fedora are irrelevant with respect to *every* normative requirement — clearly worlds where *B* does not murder are morally preferable to those in which *B* does murder, whether or not *B* has a fedora. Rather, what I am trying to suggest here is that deontic readings of modals can express claims about less-than-total bodies of norms or correctness. When it comes to the clearly less-than-total normative conditions placed on *B* by *A*'s command in (4.13), for example, worlds where *B* has no fedora are simply irrelevant, even though a more complete view of the normative demands on *B* distinguishes between worlds where *B* has no fedora.

alternatives?

The first reason to think that adopting the asymmetric description helps us latch onto a meaningful difference in the kinds of contents the two states possess is that theorists working in the Kratzerian tradition in some sense already tend to assume as much. In discussions of epistemic readings of necessity modals on a Kratzerian semantics, ordering sources are often ignored altogether.[12] Sometimes, it is explicitly assumed that epistemic necessities are (or at least can be) true relative to empty ordering sources (for example, Kratzer (1991)). If an epistemic modal is evaluated with respect to a modal base and an empty ordering source, this amounts to saying it is evaluated with respect to parameters that divide the domain of worlds into just *two* categories: the worlds within the modal base, and the worlds without it. Deontic readings of modals, by contrast, are taken to *essentially* involve a non-empty ordering source, since the ordering source is supposed to be the parameter that can take on a set of norms or preferences as its value, and deontic readings are supposed to tell us something about the relevant norms or preferences in play. This amounts to saying that deontic modals *always* divide the domain of worlds into three categories: the undominated worlds in the modal base, the dominated worlds in the modal base, and those outside the modal base. Thus, the kind of asymmetry in the contents had by cognitive and conative states I am positing is fairly standard within contemporary semantic theory. As I discuss in §4.6.2, what will be non-standard in my view is the interpretation of these distinctions and their representation in the pragmatics of the theory.

A second reason to take the asymmetric way of thinking of things seriously is that whether we try to stick to one or the other of these symmetric ways of describing things, there will still be reasons to account for a serious asymmetry between the two kinds of case. For convenience, let's give some storied labels to the two categories of contentful states we are discussing: (i) *cognitive* states which include acceptance, belief, and

12. For a few examples, see von Fintel and Gillies (2007), von Fintel and Gillies (2011) and Anand and Hacquard (2013).

knowledge; and (ii) *conative* states which include obligations, intentions, and desires. If we were to adopt the symmetric-2 option, many of our practices will still require us to distinguish between two different ways of not-satisfying a conative state, practices that are deeply connected with the nature of the state. For example, take a command backed by sanctions that q_p, given by A to B. If p and q are both true, then the command is fulfilled and positive sanctions like rewards or praise should obtain for B. If p is true and q is false, then the command is not-satisfied by being *violated*, and negative sanctions like punishment or blame should obtain for B. If p is false, on the other hand, the command is not-satisfied in a second important way: the command was unfulfillable and unviolatable, and neither positive nor negative sanctions should obtain for B. Here, something has gone wrong with A's giving the order at all, and the application of any sanctions would be unfair. We will need similar distinctions to make sense of someone's emotions regarding, say, a desire for q_p that in one case meets the fact that p is true and q is false, and in the other, meets the fact that p is not true. In the former case, they may feel frustrated, in the latter, they may feel delusional. Conative states bear systematic relationships to different kinds of sanctions, and these often track such a threefold distinction.

If, on the other hand, we adopt the symmetric-3 option, we will still need, for many of our practices intimately connected with the natures of cognitive states, to lump together the two ways of being not-correct. A characteristic part of the point of giving testimony is to represent things correctly. If my testimony represents q_p as true, then it represents things correctly iff q_p is true, and incorrectly otherwise. Clearly, if q_p is false because p is true and q is false, then my report is incorrect. But, if p is false and so my report has a false presupposition, then my report is still incorrect since it is not true. Either way, I have not succeeded in the essential function of giving testimony: to represent things correctly. Thus, even if we adopt the more fine-grained symmetric-3 way of describing things, many of our concerns for cognitive states will carve things more coarsely into the two categories.

4.5 The Basic Account

In this section, I develop a simple propositional modal logic that implements the ideas discussed in the last section. In particular, modals will be sensitive to parameters that distinguish between relevant and irrelevant worlds, and within the relevant worlds, those that are good or bad. I will work with a language that includes negation, the representation of presuppositions by use of subscripts, and a necessity modal.

The theory I present attempts to stay as neutral as possible on substantive issues in the theory of presupposition. It adopts a 'two-dimensional' approach to semantic composition (like the approach of Karttunen and Peters (1979) and subsequent work like Mandelkern (2016)), where a sentence ϕ is assigned two dimensions of meaning: its presupposition satisfaction conditions (or *domain*), $[\phi]^d$, and its truth conditions, $[\phi]^+$. The two-dimensional framework allows me to be neutral about whether presupposition failure gives rise to truth value gaps or not. Of course, I do assume that some presupposition projection behavior can be accounted for compositionally, and should not be wholly relegated to the theory of Gricean pragmatics. For example, I assume that as a matter of convention, negations inherit all of the presuppositions of their negatum.[13]

Definition (Language)**.** Our vocabulary consists of:

- $\mathcal{A}$, a countable set of atomic sentences.
- Constants: $\neg, \Box$.

The set of wffs $\mathcal{L}$ is built from the following grammar (where $p \in \mathcal{A}$):

$$\phi ::= p \mid \neg\phi \mid \phi_\phi \mid \Box\phi$$

13. See footnote 15 below for a mechanism that could intervene to explain putative counterexamples.

Models are simple, as on the Kratzerian theory.

Definition (Model)**.** A model $\mathcal{M}$ is a pair $\langle W, V \rangle$ where W is a set, and V is a function from atoms to sets of worlds ($V : \mathcal{A} \mapsto \wp(W)$).

Modals will be evaluated relative to contextually-given accessibility functions that determines which worlds are relevant, and which subset of relevant worlds are good:

Definition (Accessibility Functions)**.** An accessibility function R is a function that takes a world $w \in W$ to a pair of sets of worlds $\langle R^d_w, R^+_w \rangle$, such that $R^+_w \subseteq R^d_w$. R^d_w is meant to represent the domain of worlds *relevant* from w, while R^+_w is meant to represent those worlds in the domain that are *good*. Worlds in $R^d \backslash R^+_w$ are *bad*.

Definition (Total Accessibility Functions)**.** Given a model $\mathcal{M}$ and an accessibility function R on $\mathcal{M}$:

$$R \text{ is } \textit{total} \text{ iff for every } w \in W: \ R^d_w = W$$

Definition (Context)**.** A *context* in a model c, R is a pair of accessibility functions where c is a total function, with c^+_w representing the set of worlds compatible with the common ground of a conversation at w. R represents the accessibility function relevant for the interpretation of the necessity modal.

Epistemic readings of modals will be evaluated with respect to the total accessibility function representing the common ground, so in a context where the necessity modal $\Box$ gets an epistemic reading, $R = c$. Deontic readings of modals, by contrast, will be evaluated with respect to parameters that only count *those worlds compatible with the common ground* as relevant. So in contexts where $\Box$ gets a deontic reading, $R^d_w \subseteq c^+_w$.

Assumption 1 (Context Properties)**.** *To summarize:*

- *c is always a total accessibility function.*

- *If a context c, R is deontic, the accessibility function R only deems relevant the worlds that are compatible with the common ground, i.e., for every world w:*

$$R^d_w \subseteq c^+_w$$

- *If a context is epistemic, then $R = c$.*

- *R is always either deontic or epistemic.*

Semantic Values. Relative to a model $\mathcal{M}$ and a context c, R, a sentence ϕ will be assigned (i) a *domain* (a set of worlds that satisfy the presuppositions of ϕ), written $[\phi]^d_{\mathcal{M},c,R}$, and (ii) a set of *truth conditions*, written $[\phi]^+_{\mathcal{M},c,R}$. I will usually suppress reference to the model. The only constraint we place on the relationship between these entities in general is that $[\phi]^+_{c,R} \subseteq [\phi]^d_{c,R}$: in other words, the truth of ϕ requires the satisfaction of its presuppositions.

Semantic values are defined recursively. For simplicity, I assume atoms ($p \in \mathcal{A}$) presuppose nothing, and are true when the valuation function V says so:

$$[p]^d_{c,R} = W \qquad\qquad [p]^+_{c,R} = V(p)$$

Negations inherit the presuppositions of their negatum, and are true just in case these presuppositions are satisfied but the truth conditions of the negatum are not met:

$$[\neg\phi]^d_{c,R} = [\phi]^d_{c,R} \qquad\qquad [\neg\phi]^+_{c,R} = [\phi]^d_{c,R} \backslash [\phi]^+_{c,R}$$

Sentences of the form ψ_ϕ are true just in case ϕ and ψ are true. They inherit whatever presuppositions ψ already has, but also presuppose that ϕ is true:

$$[\psi_\phi]^d_{c,R} = [\phi]^+_{c,R} \cap [\psi]^d_{c,R} \qquad\qquad [\psi_\phi]^+_{c,R} = [\phi]^+_{c,R} \cap [\psi]^+_{c,R}$$

Now for our necessity modals. I will assume that they have three main presuppositions: (i) they presuppose at a world w that the common ground c^+ satisfies the presuppositions of their prejacent clauses; (ii) they presuppose that there are some 'good' worlds (so R^+_w is non-empty); and (iii) they carry a new Divergence presupposition: that there is a world in the domain of the accessibility relation where the prejacent is true, and a world where its negation is true.[14]

$$\begin{aligned}
[\Box\phi]^d_{c,R} &= \{w \mid c^+_w \subseteq [\phi]^d_{c,R}\} && \text{(Satisfaction)}\\
&\cap \{w \mid R^+_w \neq \varnothing\} && \text{(Non-emptiness)}\\
&\cap \{w \mid [\phi]^+_{c,R} \cap R^d_w \neq \varnothing \text{ and } [\neg\phi]^+_{c,R} \cap R^d_w \neq \varnothing\} && \text{(Divergence)}
\end{aligned}$$

As for their truth conditions, they are the usual ones: that the good worlds are a subset of the truth conditions of the prejacent. We take the intersection of this set with the domain of $\Box\phi$ in order to ensure that its truth requires satisfaction of its presuppositions:

$$\begin{aligned}
[\Box\phi]^+_{c,R} &= \{w \in W \mid R^+_w \subseteq [\phi]^+_{c,R}\} && \text{(Entailment)}\\
&\cap [\Box\phi]^d_{c,R} && \text{(Presupposition satisfaction)}
\end{aligned}$$

Definition (Truth)**.** A *point* is a sequence $(\mathcal{M}, c, R, w)$ of a model $\mathcal{M}$, a context c, R, and a world $w \in W_{\mathcal{M}}$. We say that ϕ is true relative to a point ($\mathcal{M}, c, R, w \vDash \phi$) as follows:

$$\mathcal{M}, c, R, w \vDash \phi \text{ iff } w \in [\phi]^+_{c,R}$$

We say that a point satisfies the presuppositions of ϕ ($\mathcal{M}, c, R, w \pm \phi$) as follows:

$$\mathcal{M}, c, R, w \pm \phi \text{ iff } w \in [\phi]^d_{c,R}$$

14. If a reader is concerned that it is not good practice to define the Divergence presupposition with reference to the semantic value of a formula $\neg\phi$ that is not a constituent of $\Box\phi$, this is easily rectified by lifting the semantics of negation to a space of semantic values constituted by elements $P = \langle P^d, P^+\rangle$, where $P^+ \subseteq P^d \subseteq W$. Then $-P = \langle P^d, P^d \backslash P^+\rangle$.

Notice that we have not said when ϕ is false at a point ($\mathcal{M}, c, R, w \dashv \phi$). There are two clear options. We could use a bivalent notion, on which a formula is false just in case it is not true:

$$\mathcal{M}, c, R, w \dashv \phi \text{ iff } w \notin [\phi]^{+}_{c,R}$$

Or, we could adopt the gappy or trivalent way of thinking of presuppositions, and say that a formula ϕ is false just in case its presuppositions are met but it is not true:

$$\mathcal{M}, c, R, w \dashv \phi \text{ iff } w \in [\phi]^{d}_{c,R} \backslash [\phi]^{+}_{c,R}$$

This amounts to admitting that a formula ϕ may be neither true nor false at a point.

Happily, we may avoid commitment to either notion, since the results concerning the status of Samaritan inferences will not depend on taking a stand on this issue.[15]

A Samaritan inference is an argument of the form ⌜$\Box(\psi_\phi)$, therefore $\Box(\phi)$⌝. In order to judge the results of the framework just outlined for the status of Samaritan inferences in epistemic and deontic contexts, we need to define a notion of entailment meant to model the 'therefore' relation.

I plan to count Samaritan inferences as *valid* in epistemic contexts because the truth of the premise guarantees the truth of the conclusion. Furthermore, I hope to ground the *invalidity* of Samaritan inferences in deontic contexts in the failure of the Divergence presupposition for the conclusion; in other words, I hope to predict that in a deontic

15. One might object that the bivalent version of our theory is still not classical, especially given our treatment of negation as a presupposition 'hole'. In particular, $\neg\phi$ is not in general true iff ϕ is false, and there are worlds where ϕ and $\neg\phi$ are both false. Thus, while every formula ϕ is either true or false at every point, it's not true that for every formula ϕ, either ϕ or $\neg\phi$ is true at every point.

I take this as a feature, not a bug, of a compositional treatment of the negation in our language and its presupposition projection behavior. If one is worried that our object language should have *some* way of using negation to just mean *not true*, we might borrow the the technique of Beaver and Krahmer (2001) and include a presupposition suspension operator A, where $A(\phi)$ has the same truth conditions as ϕ but no presuppositions:

$$[A(\phi)]^{d}_{c,R} = W \qquad\qquad [A(\phi)]^{+}_{c,R} = [\phi]^{+}_{c,R}$$

Then we might formalize ⌜p is not true ⌝ in the object language as follows $\neg A(\phi)$. On the bivalent version of the present theory, $\neg A(\phi)$ is true iff ϕ is false.

context, when $\Box(\psi_\phi)$ is true, $\Box(\phi)$ is undefined. This might suggest that we use classical, truth-conditional entailment as the proper way of distinguishing between the status of the inference pattern in the two kinds of context.

However, using classical truth-conditional entailment will mean that simple generalizations of the system predict our necessity modals to be non-monotonic in simple cases where they seem to license upward monotonic inferences. For example, suppose we added conjunction $\wedge$ to our language, and gave it a semantics such that for two atoms p, q: $[p \wedge q]^+ = V(p) \cap V(q)$ and $[\neg(p \wedge q)]^+ = W \backslash (V(p) \cap V(q))$. Further, suppose 'you must sweep and mop' ($\Box(p \wedge q)$) is true. Then by Divergence, *you sweep and mop* ($p \wedge q$) and *you do not sweep and mop* ($\neg(p \wedge q)$) are both compatible with the relevant domain of worlds, R^d. Now, we would like to predict that $\Box(p)$ ('you ought to sweep') is an entailment of $\Box(p \wedge q)$. The truth of $\Box(p)$ requires, however, that $\neg p$ is compatible with the relevant set of worlds, which is not guaranteed by the truth of $\Box(p \wedge q)$, since $\neg(p \wedge q) \equiv \neg p \vee \neg q$ might be compatible with R^d only by virtue of $\neg q$ ('you do not mop') being compatible. Thus, it would be possible for the Divergence presupposition to be satisfied in a deontic context for $\Box(p \wedge q)$, but not for $\Box p$, since R^d may overlap with $[\neg q]^+ = W \backslash V(q)$ (and therefore, with $[\neg(p \wedge q)]^+ = W \backslash (V(p) \cap V(q))$), but not with $[\neg p]^+ = W \backslash V(p)$. In this case, $\Box(p \wedge q)$ would be true, but $\Box p$ would not be true, since it suffers from presupposition failure.

What we need to solve this problem is a generalization of classical entailment where ϕ may entail ψ even if at some points of evaluation, ϕ is true but the presuppositions of ψ are unmet. A well-known example of such a generalization is 'Strawson entailment' introduced by von Fintel (1999) to handle cases exactly like this one. Put simply, ϕ Strawson-entails ψ iff whenever ϕ is true and the presuppositions of ϕ and ψ are met, ψ is true. Clearly, adopting Strawson-entailment would solve the problem just discussed for $\Box(p \wedge q)$ entailing $\Box p$, since cases in which the Divergence presupposition failed for $\Box p$ would be disregarded. However, adopting Strawson entailment would also undermine the explanation I am trying to give for the status of Samaritan inferences in the deontic contexts: for if deontic contexts where $\Box(\psi_\phi)$ is true ensure that $\Box(\phi)$ is undefined, then $\Box(\psi_\phi)$ vacuously Strawon-entails $\Box(\phi)$, since it is never the case that $\Box(\psi_\phi)$ is true and the presuppositions of both $\Box(\psi_\phi)$ and $\Box(\phi)$ are satisfied.

Rather than a flaw of the present framework, I think this points to a serious flaw in using Strawson-entailment in order to represent intuitive entailment. For Strawson-entailment counts any argument in which the truth of the premise ensures that the presupposition conditions of the conclusion are unmet as valid. This means that the following absurd argument counts as Strawson-valid:

(4.15) a. There is no king of France.

b. So, the king of France is bald.

Fortunately, it is simple to refine Strawson-entailment in order to rule out the validity of cases like these. We need only add the requirement that the truth of the premise is *compatible* with the satisfaction of the presuppositions of the conclusion. Clearly, adding this requirement will deliver a better verdict for cases like (4.15). It will also serve my purposes by occupying a middle ground between truth-conditional entailment and Strawson-entailment. Like Strawson entailment, it may deem an argument valid even if it is possible for the premise to be true and for the presuppositions of the conclusion to be unsatisfied. The discussion of conjunction above suggested this sort of flexibility would be desirable. But unlike Strawson-entailment, we will predict that since $\Box(\psi_\phi)$ ensures that the presuppositions of $\Box\phi$ are unsatisfied in deontic but not epistemic contexts, Samaritan arguments are invalid in those contexts.

This, then, is the notion of entailment I will adopt in order to account for the differences in the status of Samaritan inferences.

Definition (Entailment $\models$)**.** Given a set $\mathbb{C}$ of sequences of models and contexts in those models $\langle \mathcal{M}, c, R \rangle$, we say that ϕ entails ψ relative to $\mathbb{C}$ ($\phi \models_{\mathbb{C}} \psi$) as follows:

$$\phi \models_{\mathbb{C}} \psi \Leftrightarrow \text{there is some } \langle \mathcal{M}, c, R \rangle \in \mathbb{C} \text{ such that for some } w \in W_{\mathcal{M}} : \; w \in [\phi]^{+}_{c,R} \cap [\psi]^{d}_{c,R}$$
$$\text{and for every } \langle \mathcal{M}, c, R \rangle \in \mathbb{C} \text{ and } w \in W_{\mathcal{M}} : \; \text{ if } w \in [\phi]^{+}_{c,R} \cap [\psi]^{d}_{c,R} \text{ then } w \in [\psi]^{+}_{c,R}$$

If ϕ entails ψ with respect to $\mathbb{C}$, I will say that $\mathbb{C}$ validates the argument ϕ therefore ψ. If ϕ does not entail ψ relative to $\mathbb{C}$, then I say that a Samaritan inference is *invalid* with respect to $\mathbb{C}$.

Given the Divergence presupposition, $\Box\phi$ can be true only if ϕ is *contingent* in the following sense:

Definition (Contingency)**.** Relative to a model and context:

$$\phi \text{ is contingent iff } \exists w, v \in W : \; w \in [\phi]^{+}_{c,R} \text{ and } v \in [\neg\phi]^{+}_{c,R}$$

This is likely to raise some objections. For epistemic readings of modals, at least, it seems clear enough that $\Box\phi$ can be true when ϕ expresses a logically or metaphysically necessary truth. I think this sort of worry can be met. First, notice that it also arises for the received view, given the original Divergence presupposition: if there is no world where $\neg\phi$ is true, then there can be no such world within the modal base. In any case, I think one should not assume that 'contingency' in the sense I have defined it here is meant to align with genuine metaphysical or logical contingency. Rather, we might think of it as a sort of 'hyperintensional' notion of contingency, on which the set of worlds includes some that are logically or metaphysically impossible; in particular, worlds where the negations of logical or metaphysical necessities are true. See Berto and Jago (2019) for discussion of the notion of impossible worlds as applied to problems of hyperintensionality.[16] Assuming there is some satisfactory way to respond to these

16. There are significant challenges for the use of impossible worlds to model hyperintensionality in

worries, the following results for Samaritan inferences will concern sentences ϕ that are contingent in the sense defined.

Proposition 1 (Epistemic Contexts Validate Contingent Samaritan Inferences)**.** *Let* $\mathbb{E}_\phi$ *be the class of models with epistemic contexts on them such that* ϕ *is contingent. Then* $\Box(\psi_\phi) \therefore \Box\phi$ *is valid with respect to* $\mathbb{E}_\phi$*. In other words:*

$$\Box(\psi_\phi) \models_{\mathbb{E}} \Box\phi$$

Proof. First, I prove that if the premise is true and the conclusion's presuppositions are met, then the conclusion is true. Let $\mathcal{M}$ be a model and c, R be an epistemic context. Then $R = c$. Let w be a world such that $w \in [\Box(\psi_\phi)]^+_{c,R} \cap [\Box(\phi)]^d_{c,R}$. Then we want to show that $w \in [\Box(\phi)]^+_{c,R}$. Recall the truth conditions for $\Box\phi$:

$$[\Box\phi]^+_{c,R} = [\Box\phi]^d_{c,R} \cap \{w \in W \mid R^+_w \subseteq [\phi]^+_{c,R}\}$$

By assumption, $w \in [\Box\phi]^d_{c,R}$. So we just need to show $R^+_w \subseteq [\phi]^+_{c,R}$. Since $w \in [\Box(\psi_\phi)]^+_{c,R}$, we know that $R^+_w \subseteq [\psi_\phi]^+_{c,R}$. Since $[\psi_\phi]^+_{c,R} \subseteq [\phi]^+_{c,R}$, it follows that $R^+_w \subseteq [\phi]^+_{c,R}$. Thus, $w \in [\Box(\phi)]^+_{c,R}$.

Now we need to show that there are indeed models and epistemic contexts c, R where for some world w, $w \in [\Box(\psi_\phi)]^+_{c,R} \cap [\Box(\phi)]^d_{c,R}$. Take a world and epistemic context with the following properties:

- $c^+_w \subseteq [\psi_\phi]^+_{c,R}$ (Satisfaction and Entailment)
- $c^+_w \neq \varnothing$. (Non-emptiness)
- $c^d_w \cap [\neg(\psi_\phi)]^+_{c,R} \neq \varnothing$
- $c^d_w \cap [\neg\phi]^+_{c,R} \neq \varnothing$ (Divergence)

general (see, for example, Bjerring and Schwarz (2017) for discussion), but my needs here are fairly modest: we need only include an absurd world $\bot$ where all negations are true.

Then, $w \in [\Box(\psi_\phi)]^+_{c,R} \cap [\Box(\phi)]^d_{c,R}$.

Thus:

$$\Box(\psi_\phi) \models_{\mathbb{E}} \Box\phi$$

$\square$

Proposition 2 (Deontic Contexts Invalidate Contingent Samaritan Inferences)**.** *Let* $\mathbb{D}$ *be the set of sequences of models with deontic contexts.* $\mathbb{D}$ *invalidates Samaritan inferences.*

Proof. I will show that $\mathbb{D}$ invalidates Samaritan inferences since there is no $\langle \mathcal{M}, c, R \rangle \in \mathbb{D}$ such that for some world $w \in W_{\mathcal{M}}$, $w \in [\Box(\psi_\phi)]^+_{c,R} \cap [\Box(\phi)]^d_{c,R}$.

Let $\langle \mathcal{M}, c, R \rangle \in \mathbb{D}$ and suppose $w \in [\Box(\psi_\phi)]^+_{c,R}$. Then by the Satisfaction presupposition, $c^+_w \subseteq [\psi_\phi]^d_{c,R}$. Since $[\psi_\phi]^d_{c,R} = [\psi]^d_{c,R} \cap [\phi]^+_{c,R}$, it follows that $c^+_w \subseteq [\phi]^+_{c,R}$. Since R is deontic, $R^d_w \subseteq c^+_w$. Therefore, $R^d_w \subseteq [\phi]^+_{c,R}$. Clearly, then, $R^d_w \cap [\neg\phi]^+_{c,R} = \varnothing$, so the Divergence presupposition of $\Box\phi$ is unsatisfied at w. Therefore, $w \notin [\Box\phi]^d_{c,R}$.

Since w was arbitrary, the same holds for every world in $W_{\mathcal{M}}$. Since $\langle \mathcal{M}, c, R \rangle$ was an arbitrary, there is no element $\langle \mathcal{M}, c, R \rangle$ of $\mathbb{D}$, such that for some $w \in W_{\mathcal{M}}$: $w \in [\Box(\psi_\phi)]^+_{c,R} \cap [\Box(\phi)]^d_{c,R}$. Thus, Samaritan inferences are invalid with respect to $\mathbb{D}$. $\square$

Given Propositions 1 and 2, then, the present theory shows that it is possible to give an account on which necessity modals are given a uniform semantics, their flavors are contextually (not lexically) determined, and Samaritan inferences are valid for epistemic contexts but not deontic ones.

4.6 Further Issues

4.6.1 Weak Speech Reports

I noted above that Samaritan inferences also seem invalid for verbs of *saying*: *say*, *claim*, etc. It is not clear how to give such verbs a semantics in the Kratzer style (relative to a modal base and a non-empty ordering source), so it is not clear whether the received solution to the Samaritan paradox could apply to them.

By contrast, it is simple to see how my proposal could extend to them. Suppose Anna utters, "Anna's wardrobe is clean." Let 'q_p' abbreviate 'Anna's wardrobe is clean', let 'p' abbreviate 'Anna has a wardrobe', and let '$\Box$' stand for 'Anna said'. The content of Anna's assertion speech act was $[q_p]$ — we might say she presupposed that p was true, asserted q, and denied that $\neg q$. We model this fact by supposing that at the actual world w, $R^d_w = [q_p]^d_R$, and $R^+_w = [q_p]^+_R = [p]^+ \cap [q]^+$ is the set of worlds compatible with what Anna asserted, and $R^d_w \backslash R^+_w = [\neg q_p]^+_R = [p]^+ \backslash [q]^+$ is the set of worlds compatible with what she denied. Then $\Box(q_p)$, is true, since $R^+_w \subseteq [q_p]^+$ and the Divergence presupposition is satisfied. However, since $R^d_w \subseteq [p]^+$, or, on our interpretation, Anna presupposed that she has a wardrobe, it follows that $\neg p$ is not compatible with what she presupposed. Thus, $R^d_w \cap [\neg p]^+ = \varnothing$, so $\Box p$ is not true: it's not true that Anna said she had a wardrobe.

4.6.2 Relation to the Kratzerian Theory

The solution to the Samaritan paradox I have argued for in this chapter began from an insight of the Kratzerian way of modeling the failure of Samaritan inferences in the deontic case: that it divided the relevant set of worlds into three substantive groups: the good, the bad, and the irrelevant. This raises a question: is my solution to the Samaritan paradox compatible with the Kratzerian theory?

Yes and no. On the standard interpretation of the Kratzerian semantics, the union of what I am calling the 'good' and the 'bad' worlds would be the worlds within modal base, and this parameter is supposed to be something like the Stalnakerian common ground: it consists of a set of propositions taken for granted in the conversation. If $\Box p$ is true, on my account, the good worlds are all p-worlds, but there is also a bad world where p is false. Since bad worlds would be part of the modal base, it would follow that $\Box p$ is true but there is a world compatible with what we are taking for granted that makes p false. On a deontic reading, this seems intuitive enough — that's how we came to the idea — but on an epistemic reading, this is quite a bad result. We should be able to say things like the following, which sounds near nonsensical to my ear:

(4.16) It must be raining, but given what we are taking for granted, it is left open that it is not raining. ??

Therefore, accommodating my solution to the Samaritan paradox to the Kratzerian theory would seem to require a significant change in how the Kratzerian understands the nature of the two parameters. While we could interpret deontic flavors much as the Kratzerian already does, we would have to say that for epistemic flavors, modal bases are empty and it is *ordering sources* instead that contain all the relevant propositions we take for granted in the conversation. This would generate the result we are after: that in the epistemic case, all worlds are divided into the good and the bad, with the good being those worlds compatible with the common ground. The goals of the Kratzerian theory are many, and I am not sure if this rethinking of the interpretation of the parameters would sit well with the pursuit of all of them.

4.7 Conditional Samaritan Paradoxes

There are several puzzles surrounding the interactions between conditionals and modals that are sometimes labeled Samaritan paradoxes.[17] I will discuss one version here that I think most deserves that label, and outline an extension of the theory given above that may solve it.

On the Kratzerian theory of indicative conditionals, they function to add the information that the antecedent is true to any modal base involved in interpreting the consequent. So $\ulcorner \phi \rightarrow \Box\psi \urcorner$ is equivalent to $\ulcorner \Box\psi \urcorner$ evaluated with respect to a modal base f that includes the proposition denoted by ϕ. Thus, every world in the modal base relevant for

17. For example, Kratzer (1991) discusses one.

the evaluation of $\Box\psi$, dominated or undominated, will be one where ϕ is true. This has the result of making every sentence of the form $\ulcorner\phi \to \Box\phi\urcorner$ true.

This seems to be acceptable for epistemic readings of $\Box$, as illustrated by the following example, which I take to sound intuitively trivial:

(4.17) If Anna has a wardrobe, then (given what we've established) she must have a wardrobe.

On the other hand, for deontic readings of modals, such sentences do not seem trivial:

(4.18) If Anna cleans her wardrobe, then (given the rules) she must clean her wardrobe.

Call conditionals of the form $\phi \to \Box\phi$ *Samaritan conditionals*. As with the original Samaritan paradox, given Kratzerian assumptions about the lexicon, existing semantics for modals and conditionals predict that Samaritan conditionals should have the same status, no matter the flavor of the modal in the consequent. See Frank (1997), Zvolenszky (2002), and Carr (2014) for discussion of the paradox.

I will now outline an extension of the theory I have given in this chapter that can account for the different status of Samaritan conditionals in terms of the flavor of the modals involved. The extension rests on the idea that conditional antecedents add information to the context, and their consequents are evaluated with respect to the resulting context.

Definition (Language)**.** Our vocabulary consists of:

- $\mathcal{A}$, a countable set of atomic sentences.
- Constants: $\neg, \to, \Box$.

The set of wffs $\mathcal{L}$ is built from the following grammar (where $p \in \mathcal{A}$):

$$\phi ::= p \mid \neg\phi \mid \phi'_{\phi} \mid \phi \to \phi' \mid \Box\phi$$

Models, accessibility functions, contexts, and related notions are defined as before.

The non-conditional fragment has the same semantics as in §4.5. The semantics of our conditional requires the notion of a context update.

Definition (Context Update)**.** Given an accessibility function R and a set of worlds S, $R(S)$ is the accessibility function that for each w:

$$R(S)^d_w = R^d_w \text{ and } R(S)^+_w = R^+_w \cap S$$

Allow me the following abbreviation. Relative to a model and a context $\mathcal{M}, c, R$:

$$c(\phi) := c([\phi]^+_{c,R}) \text{ and } R(\phi) := R([\phi]^+_{c,R})$$

$$(c, R) + \phi := c(\phi), R(\phi)$$

I will assume that conditionals presuppose that the common ground satisfies the presuppositions of the antecedent, and that the common ground in the context updated by the antecedent satisfies the presuppositions of the consequent:

$$[\phi \to \psi]^d_{c,R} = \{w \mid c^+_w \subseteq [\phi]^d_{c,R}\} \cap \{w \mid c([\phi]^+_{c,R})^+_w \subseteq [\psi]^d_{c,R+\phi}\}$$

A conditional is true iff its presuppositions are satisfied, and the common ground updated with the antecedent makes the consequent true:

$$[\phi \to \psi]^+_{c,R} = [\phi \to \psi]^d_{c,R} \cap \{w \mid c([\phi]^+_{c,R})^+_w \subseteq [\psi]^+_{c,R+\phi}\}$$

The framework just outlined will distinguish between the status of deontic vs. epistemic interpretations of simple Samaritan conditionals like $p \to \Box p$ for contingent $p \in \mathsf{At}$. Since Samaritan conditionals are single statements, and not arguments, I will have to adapt my notion of entailment (Def. 4.5) to serve my different purposes here. While I will not give a formal definition, I will first show that the presuppositions of $p \to \Box p$ are satisfied at a point only if the updated context $(c, R) + p$ is epistemic. We might think of this result as showing that $p \to \Box p$ fails a null-premise version of the *compatibility* requirement for entailment in deontic contexts. Second, I will show that if the presuppositions of $p \to \Box p$ are satisfied at a point (and so the context is epistemic), it is true. I then outline a model in which it is true, which might be thought of as a demonstration that the sentence satisfies a null-premise version of the compatibility requirement for entailment in epistemic contexts.

Proposition 3 (If the presuppositions of $p \to \Box p$ are satisfied at a point $\mathcal{M}, c, R, w$, then $c, R + p$ is epistemic)**.**

Proof. Let $\mathcal{M}, c, R, w$ be a point, and suppose $w \in [p \to \Box p]^d_{c,R}$. Given the presuppositions of conditionals, we know that $c([p]^+_{c,R})^+_w \subseteq [\Box p]^d_{(c,R)+p}$. Since the semantics of atomic sentences is unaffected by updates to the context, we know that:

$$[p]^+_{(c,R)+p} = V(p) \text{ and } [\neg p]^+_{(c,R)+p} = W \backslash V(p)$$

Clearly, then, $c([p]^+_{c,R})^+_w = c(V(p))^+_w$. Let $v \in c(V(p))^+_w$. Our presupposition conditions for conditionals mean that if $p \to \Box p$ is defined at w, then $v \in [\Box p]^d_{(c,R)+p}$. By the Divergence presupposition, we know that $\neg p$ is compatible with the domain of R relative to v in the context updated with p, i.e.:

$$R([p]^+_{(c,R)+p})^d_v \cap [\neg p]^+_{(c,R)+p} \neq \varnothing$$

or, put more simply given what we've established:

$$R(V(p))^d_v \cap (W \backslash V(p)) \neq \varnothing$$

Now, suppose for contradiction that $(c, R) + p$ were deontic. Then the set of relevant worlds in the updated context is a subset of the resulting common ground ($R(V(p))^d_v \subseteq c(V(p))^+_v$). By Divergence, then, $\neg p$ should be compatible with the resulting common ground ($c(V(p))^+_v \cap (W \backslash V(p)) \neq \varnothing$). But this is impossible, since the common ground that results from updating with p entails p:

$$c(V(p))^+_v = c^+_v \cap V(p) \subseteq V(p)$$

In sum, if $(c, R) + p$ were deontic, then $R(p)$ could only count p-worlds as relevant, since the common ground in the updated context entails p. Thus, in the updated context the Divergence presupposition of $\Box p$ on a deontic interpretation would fail. So, $(c, R) + p$ cannot be deontic; if $p \rightarrow \Box p$ is defined relative to c, R, then $(c, R) + p$ must be epistemic.

□

Proposition 4 (If $p \rightarrow \Box p$ is defined, then it is true)**.**

Proof. Let $\mathcal{M}, c, R, w$ be a point and suppose $w \in [p \rightarrow \Box p]^d_{c,R}$. The conditional is true at w iff:

$$c([p]^+_{c,R})^+_w \subseteq [\Box p]^+_{(c,R)+p}$$

Let $v \in c([p]^+_{c,R})^+_w$. Since the conditional's presuppositions are satisfied, $v \in [\Box p]^d_{(c,R)+p}$. Thus, we need to show only that the good worlds from the perspective of v entail p:

$$R([p]^+_{(c,R)+p})^+_v \subseteq [p]^d_{(c,R)+p}$$

As before, since the semantics of atomic sentences is unaffected by the update of a context, this is equivalent to showing:

$$R(V(p))^+_v \subseteq V(p)$$

Showing this is trivial, since $R(V(p))^+_v = R^+_v \cap V(p) \subseteq V(p)$. □

In sum, given my assumptions, a Samaritan conditional $p \to \Box p$ is never defined when $(c, R) + p$ is deontic; when the formula is defined, $(c, R) + p$ is epistemic and the formula is true. In future research, I hope to explore generalizations of these results along at least two dimensions: (i) concerning the status of conditionals of the form $\phi \to \Box\phi$ with more complex substitutions of ϕ; and (ii) concerning more relaxed assumptions about the natures of the parameters with respect to which modals are evaluated.

I want to close by giving two concrete illustrations of the sorts of plausible situations in which conditionals with epistemic or deontic modals in the consequent can be true and false. Consider a basic four world model representing the classical possible evaluations of atoms p, q ($W = \{w_1, w_2, w_3, w_4\}$):

Four Worlds

	p	q
w_1	1	1
w_2	1	0
w_3	0	1
w_4	0	0

For simplicity, I will discuss the verdicts of the present system for conditional-modal formulas relative to *uniform* contexts, where for any two worlds w, v, $c_w = c_v$ and $R_w = R_v$; and the *null* common ground; i.e. where $c_w^+ = W$. Thus, I will specify a context simply by stating the values of R^d and R^+ (since the contexts I will discuss are *uniform*, I drop the subscript specifying the world associated with these values.

1. Epistemic R. Let $R = c$. Take an arbitrary world $w \in W$, and consider the status of $p \to \Box p$. First, the formula's presuppositions are satisfied under the following conditions:

$$w \in [p \to \Box p]^d_{c,R} \text{ iff } c_w^+ \subseteq [p]^d_{c,R} \text{ and } c([p]^+_{c,R})^+_w \subseteq [\Box p]^d_{c,R+p}$$

Clearly, since atoms presuppose nothing, the common ground satisfies p's presuppositions ($c_w^+ \subseteq [p]_{c,R}^d$). Now let's check that the context updated with the antecedent satisfies the presuppositions of $\Box p$. Let $v \in c([p]_{c,R}^+)_w^+$. Again, since atoms presuppose nothing, the Satisfaction presupposition is trivially satisfied. Since c, R is epistemic and c is uniform and null, $R(p)_v^+ = c([p]_{c,R}^+)_v^+ = \{w_1, w_2\}$, which is not empty. Second, since c, R is epistemic, $R(p)_v^d = W$, so the Divergence presupposition is satisfied. Finally, $R(p)_v^+ = R_v^+ \cap [p]_{c,R}^+ \subseteq V(p)$, so the conditional is true.

2. R **as a conditional norm**. Let the relevant worlds be the p-worlds, and the good ones be those where q is also true. So $R^d = \{w_1, w_2\}$ and $R^+ = \{w_1\}$. This, I take it, is the sort of situation in which a sentence like $p \to \Box q$ would intuitively be true, but $p \to \Box p$ not. The present system predicts this. Consider the common ground updated with the antecedent: $c([p]_{c,R}^+)^+ = \{w_1, w_2\}$ and let v be an arbitrary member of that set. Then $R(p)_v^+ = \{w_1\} \cap \{w_1, w_2\} = \{w_1\}$, so it is non-empty. But $R(p)_v^d = R_v^d = \{w_1, w_2\}$, which witnesses Divergence for $\Box q$ (since q is true at w_1 and $\neg q$ is true at w_2), but not for $\Box p$, since $\neg p$ is not true at any world in $\{w_1, w_2\}$. Thus, $p \to \Box p$ is undefined, but $p \to \Box q$ is defined. Finally, since $R(p)_v^+ = \{w_1\}$ is a subset of the q-worlds, it follows that $p \to \Box q$ is true.

4.8 Conclusion

On the theory just outlined, the original Samaritan paradox and its conditional version have a common diagnosis. Modals carry a Divergence presupposition that requires both the prejacent and its negation to be compatible with the relevant set of worlds. When the accessibility function is epistemic, *all* worlds are relevant, so this presupposition is satisfied for contingent formulas. This means that in epistemic contexts, $\Box q_p$ entails $\Box p$, and if $p \to \Box p$ is defined, it is true. When the context is deontic, on the other hand, the relevant worlds may be restricted so that the Divergence presupposition fails.

In particular, in deontic contexts, $\Box q_p$ does not entail $\Box p$, and the presuppositions of $p \rightarrow \Box p$ are never met.

Chapter 5: Supposition, Presupposition, and Trivalent Conditionals

5.1 Introduction

In the last chapter, I argued that the parameters with respect to which modals are evaluated can make a three-fold distinction between worlds. The semantics I gave for the interaction between modals, presuppositions, and conditionals suggests that presuppositions and conditionals give us the linguistic resources for communicating information not just about what it takes for a world to be relevant and 'good' with respect to a modal state, but also non-trivial information about how it is possible for a world to be relevant and 'bad.' When $\Box\phi$ is true relative to a modal state, I argued, this is not just because all the relevant and good worlds satisfy ϕ's presuppositions and make ϕ true, but also that some of the bad worlds are worlds that satisfy ϕ's presuppositions and make ϕ false. Similarly, when an indicative conditional like $\psi \rightarrow \Box\phi$ is true, this means not only that all of the relevant and good worlds are ψ-and-ϕ worlds, but also that some of the 'bad' worlds are ψ-and-not-ϕ worlds. On the theory of the last chapter, then, speakers may leverage both presuppositions and conditional antecedents (which we might call *suppositions*) to communicate information about conditions met both by the good and the bad worlds. Both kinds of expression can provide information about the distinctions that modal states draw between worlds that are relevant or irrelevant.

In this chapter, I explore the prospects of a theory that explores an analogy between presuppositions and conditional antecedents in its most literal form: presuppositions and suppositions are semantically equivalent ways of placing conditions on the truth or falsity of the sentences they attach to. In particular, both give rise to truth-value gaps at worlds where their pre(suppositions) fail.

Apart from my own motivations for exploring the prospects of this sort of theory, there is independent reason to do so. The trivalent theory of indicative conditionals – on which 'if *A*, *B*' is true if *A* and *B* are true, false if *A* is true and *B* is false, and neither true nor false when *A* is false – has a lot going for it. Yet even the most sympathetic theorists have thought the theory comes at a high price, given how it appears to interact with the trivalent theory of presupposition. I will argue in this chapter that by distinguishing between the characteristic effects of conditional and unconditional assertions on a context, we can avoid paying this price. I then develop a model of presupposition-carrying and conditional sentences consisting of (i) a static semantics that assigns trivalent truth conditions to sentences of both types, and (ii) a dynamic pragmatics that distinguishes between two types of assertion. The resulting package distinguishes between supposition and presupposition, even though both phenomena give rise to the same kinds of truth value gaps.

5.2 Tri-Curiosity

I bet you $1 that *if Daria went to the party last night, she had a bad time*. You accept the bet, thinking that the party was bound to be fun, and that Daria is more sociable than her reputation allows. Later, we meet Daria and she tells us that she did not end up attending the party — something came up. Who won the bet? The most natural response, I take it, is *no one*, for the bet only concerned what else would happen *if Daria had gone to the party*.

Considerations like these have motivated the *trivalent* theory of indicative conditionals,[1] on which 'if *A*, *B*' is:

1. true if *A* and *B* are true;

2. false if *A* is true and *B* is false; and,

1. See, for example, de Finetti (1936), McDermott (1996), and Milne (1997). In this chapter, I will continue the storied tradition of not theorizing about indicative and subjunctive conditionals at the same time.

3. neither true nor false if A is false.

Of course, these clauses do not cover all possibilities: they do not say what happens when A or B is neither true nor false. On the most well-known version of the view due to Bruno de Finetti (de Finetti 1936), the conditional 'if A, B' is neither true nor false in each of these cases. Alternatively, on the 'Cooper-Cantwell' semantics due to Cooper (1968) and Cantwell (2006, 2008), 'if A, B' inherits whatever value B has in these other cases (see Égré, Rossi, and Sprenger (2021a, 2021b) for discussion and comparison). In this chapter, I will be concerned with issues that are mostly independent of the choice between these versions of the trivalent semantics for indicative conditionals. For the sake of simplicity, I will focus on the de Finettian theory.

The trivalent semantics can explain the natural response to the question of who won the bet, given the following natural principle connecting winning/losing with truth/falsity:

> *Betting Bridge*. One who bets on p wins iff p is true; one who bets against p wins iff p is false.

The bet about Daria was on a conditional proposition and the antecedent turned out to be false, so, on the trivalent theory, the conditional as a whole is neither true nor false. Thus, neither of us won.

In contrast, any bivalent semantics for conditionals, combined with *Betting Bridge*, entails that someone won the bet. For example, on the *material implication* semantics for indicative conditionals (on which 'if A, B' is true iff either A is false or B is true, and is false otherwise),[2] I won the bet just because Daria did not go to the party (if she also had a bad time last night, my winning is over-determined).

While bets are the canonical inspiration for trivalent semantics for conditionals, other relations to propositions provide similar motivation. Consider the following sentences:

2. See Grice (1989), Lewis (1973), and Jackson (1979), among others, for defenses of the material implication semantics for indicative conditionals.

(5.1) a. I want to go out for a beer this evening if I am not too tired then.

b. I intend to go to the party if Daria goes.

McDaniel and Bradley (2008) argue that what they call 'conditional desires', like the one described by (5.1a), are neither satisfied nor frustrated when the antecedent condition is false — in this case, when the speaker *is* too tired in the evening. Rather, in their terminology, the desire ascribed with (5.1a) is *void* in that case. Similarly, Ferrero (2009) argues that what he calls 'conditional intentions', like the one described by (5.1b) are neither fulfilled nor frustrated if Daria ends up not going to the party. In that case, Ferrero says, the intention is *moot*. While I will not defend these views here, I want to point out that just as in the betting case, each of these judgments are generated straightforwardly by the trivalent semantics for indicative conditionals and the following bridge principles:

> *Wanting Bridge*. A want for p is satisfied iff p is true and frustrated iff p is false.

> *Intention Bridge*. An intention for p is fulfilled iff p is true and frustrated iff p is false.

Besides allowing for intuitive distinctions we make about bets, desires, and intentions, there are more theoretical reasons to be curious about the trivalent theory of conditionals — other reasons to be, as it were, *tri-curious*. First, like the restrictor theory — which could fairly be called the 'orthodox' theory of conditionals among linguists — (Lewis (1998), Kratzer (2012b), Heim (1982)), it can give a uniform account of the ability of 'if' clauses to restrict quantifiers (see Belnap (1970), Huitink (2008, 2012)). Second, natural extensions of probability functions to algebras of trivalent propositions avoid generally trivializing the probabilities of conditionals (see Cantwell (2006, 2008), Rothschild (2014), and Lassiter (2019) for discussion).

A third theoretical virtue of the trivalent semantics for conditionals suggests itself. The trivalent theory of conditionals, on the standard construal of its posited truth value gaps, says that conditional sentences denote *partial* truth functions. In semantics, partial functions are most often used to model the denotations of sentences that carry *presuppositions* (see, among many others, Frege (1948), Strawson (1950), Peters (1979), Heim and Kratzer (1998), and Beaver and Krahmer (2001)). In effect, then, the trivalent theory of conditionals proposes to model *supposition*-carrying sentences (conditionals) with the same tools used to model *presupposition*-carrying sentences. Thus, in addition to the aforementioned virtues, the trivalent semantics for conditionals suggests a unified account of suppositional and presuppositional contents: a sentence's carrying a supposition or a presupposition really amounts to the same thing: its denoting a *partial* function from worlds to truth values, one that is defined only among worlds where its (pre-)supposition holds.

Usually, that a theory unifies disparate phenomena is a consideration in its favor. But some tri-curious theorists have actually taken this unity to show that the two trivalent theories are incompatible (see von Fintel (2007), Rothschild (2014), Huitink (2008)). Given that the trivalent theory of presupposition is better understood, the perception of their incompatibility has led many to trade tri-curiosity for ambivalence — or rather, as it were, am-*tri*-valence — about the trivalent theory of indicative conditionals.

In this chapter, I will argue that the two trivalent theories are actually compatible. The key, I claim, is that suppositional and presuppositional sentences are *used* differently, to perform conditional and unconditional assertions, respectively. Distinguishing these patterns of use allows us to combine trivalent semantics for both theories without generating absurd consequences. While there may be other reasons for amtrivalence, I hope to make the case that the interaction with trivalent theories of presupposition need not be among them.

In the next section (§2), I detail the reasons for which some have taken the two trivalent theories to be incompatible. In §3, I respond to these objections, and sketch an account of conditional assertion that distinguishes between sentences that have truth value gaps on the basis of the particular language that generates these gaps. In §4 I outline a simple formal model illustrating the proposal. Finally, in §5, I consider some embeddings of conditionals (under negations and modals), and extend the model of §4 to make plausible predictions about them.

5.3 Amtrivalence

Tri-curious theorists have worried that the trivalent theory of conditionals is incompatible with the trivalent account of presupposition for two main reasons. First, there is the worry that the combination of the two trivalent theories predicts that conditionals *presuppose* their antecedents:

> "[U]sing three-valued semantics for [conditionals] ... precludes using three valued semantics for modelling presupposition (as is often done) at the same time, since clearly the antecedents of conditionals are not presupposed to be true." von Fintel 2011, 1527

Second, there is the worry that unembedded indicative conditionals *do* carry some presuppositions that will go unexplained. In particular, indicative conditionals appear to presuppose that their antecedents are compatible with the common ground of a conversational context (Stalnaker (1975), von Fintel (1997), Starr (2014)) — this is commonly supposed to be one of the distinguishing features of *indicative* (as opposed to *subjunctive*) conditionals. Since the trivalent semantics for conditionals is not context-sensitive, however, there appears to be no way of accounting for this compatibility presupposition.

Both worries fall out of a standard principle connecting truth value gaps and the conventional presuppositions of a sentence ϕ:

> *Logical Presupposition*. ϕ conventionally presupposes π iff:
> 1. if ϕ is true, π is true, and
> 2. if ϕ is false, π is true.

Let me write $A \rightarrow B$ to abbreviate the indicative conditional, 'if A, B'. On the de Finetti semantics for conditionals, when $A \rightarrow B$ is true, A is true. And when $A \rightarrow B$

is false, A is also true. Thus, *Logical Presupposition* entails that $A \to B$ presupposes A. Notably, this result does not carry over in full generality to other trivalent semantics for conditionals, such as the aforementioned Cooper-Cantwell semantics, on which $A \to B$ has whatever value B has when A is neither true nor false. But adopting the Cooper-Cantwell semantics does not help much, for it still entails that for any *bivalent* A, $A \to B$ presupposes A. This result is less bad than the one for the de Finetti semantics, but it is bad enough.

Second, the truth or falsity of $A \to B$ has nothing to do with whether A is compatible with the common ground of the conversation in which $A \to B$ is uttered. Suppose it is common ground in a conversation at w that A is false, but the participants are wrong: both A and B are true at w. Then $A \to B$ is true at w but it is false that A is compatible with the common ground of the conversation at w. Thus, by *Logical Presupposition*, $A \to B$ does not presuppose A is compatible with the common ground.

For some theoretical purposes, *Logical Presupposition* has had the status of a stipulation. In those instances, it may be easy to give up or relax, assuming an adequate alternative is available. More troubling for the assimilation of trivalent theories of conditionals and presuppositions, however, is that at least the right-to-left direction of the biconditional principle has been defended on conceptual grounds.

Somewhat ironically, the best motivation for *Logical Presupposition* is probably to be found in some of Robert Stalnaker's work arguing that notions like it should not be considered explanatory concepts in linguistic theory (Stalnaker 1973; Stalnaker 1999; Stalnaker 2014). Stalnaker famously argues that the fundamental explanatory concept for the theory of presupposition is not a logical relation between sentences or propositions, but rather a pragmatic relation between propositions and *conversational participants*. In any conversation, according to Stalnaker, the participants *pragmatically presuppose* the truth of various propositions. What is commonly presupposed between them determines a set of relevant alternative situations — the 'common ground' or the 'context set'. For

Stalnaker, the linguistic analysis of presupposition should be based not on truth value gaps, but on facts about about the common ground and what conversational participants pragmatically presuppose.

Despite his emphasis on the *pragmatic* notion of presupposition as a relation between conversational participants and propositions, Stalnaker allows that we might make sense of a derivative notion of logical or 'semantic' presupposition, and his explanation of how we might do so contains an argument for the right-to-left part of the biconditional in *Logical Presupposition*. It is worth quoting at length:[3]

> Among the reasons that a pragmatic presupposition might be required by the use of a sentence, by far the most obvious and compelling reason would be that the semantical rules for the sentence failed to determine a truth value for the sentence in possible worlds in which the required presupposition is false. *Since the whole point of expressing a proposition is to divide the relevant set of alternative possible situations — the presupposition set — into two parts, to distinguish those in which the proposition is true from those in which the proposition is false, it would obviously be inappropriate to use a sentence which failed to do this*. Thus, that a proposition is presupposed by a sentence in the technical semantic sense provides a reason for requiring that it be presupposed in the pragmatic sense whenever the sentence is used. Stalnaker 1973, 452, my emphasis.

Agreeing with Stalnaker's argument here means accepting the right-to-left part of the biconditional in *Logical Presupposition*. This direction is sufficient to give rise to our first worry, that trivalent conditionals presuppose their antecedents.

While Stalnaker certainly does not suggest this, one might also accept the further claim that truth-value gaps are the *only* way in which presuppositions can be conventionally associated with a sentence. If so, then one would accept the left-to-right direction of the biconditional in *Logical Presupposition*. Accepting this direction gives rise to our second worry, that trivalent semantics for indicative conditionals cannot explain how they seem to presuppose that their antecedents are open possibilities with respect to the common ground of a conversation.[4]

3. Stalnaker reiterates this line of thought in "Assertion" Stalnaker 1999, 89–90. See Soames (1989) for criticism of this argument.

4. Of course, this only the simplest, not the only, route to our second worry.

This argument for *Logical Presupposition* relies on two points. First, there is an account of the characteristic effect of an assertion. On Stalnaker's account, if the assertion of ϕ is accepted in a context, the conversational participants update the context set (the common ground, or the relevant set of alternative possibilities between which they aim to distinguish in conversation) by keeping the worlds where ϕ is true (I'll write $[\![\phi]\!]^+$), and removing the worlds where ϕ is false ($[\![\phi]\!]^-$):

> *Assertion*. The assertion of ϕ in a context with a context set c (represented as a set of worlds) is acceptable iff $c \subseteq [[\![\phi]\!]^+ \cup [\![\phi]\!]^-]$. When acceptable, accepting ϕ is a matter of changing the context set to $c \cap [\![\phi]\!]^+ = c \backslash [\![\phi]\!]^-$.

By asserting a sentence ϕ that denotes a partial truth function, a speaker in effect presupposes (in Stalnaker's pragmatic sense) that ϕ has a truth value at every world in the context set. Given *Assertion*, this pragmatic presupposition is predictable from the conventions associated with ϕ. There is a sense, then, in which we can think of this presupposition as attaching to the sentence itself. We can generate this derivative notion of sentential presupposition via the following bridge principle:

> *Presupposition Bridge*. A sentence ϕ conventionally presupposes π iff an assertion of ϕ is acceptable only in contexts c such that π is true at every world in c.

Assertion and *Presupposition Bridge* entail *Logical Presupposition*. Suppose the truth of π is a necessary condition on the truth or falsity of ϕ. When a speaker asserts ϕ, they propose to update the context in the way summarized by *Assertion*. For that to be possible, ϕ must have a truth value at every world in the context set. Thus, π must be true at every world in the context set. Since the context was arbitrary, every context in which the assertion of ϕ is acceptable has this feature. By *Presupposition Bridge*, ϕ conventionally presupposes π. For the converse, suppose ϕ conventionally presupposes π. Then by *Presupposition Bridge*, the assertion of ϕ is acceptable only in contexts c such that π is true at every world in c. Since the relevant notion of acceptability holds just in case the ϕ is true or false at every world in a context, it follows that π is a necessary condition on ϕ's truth or falsity.

The argument from *Assertion* and *Presupposition Bridge* to *Logical Presupposition* clearly threatens the trivalent account of conditionals on conceptual grounds. It appears to establish just by reflection on the nature of assertion and the structure of a conversational context that on the trivalent account, $A \to B$ presupposes A and does not presuppose that A is compatible with the context set. In the next section, I argue that we can block this result by limiting the scope of *Assertion* in a principled way.

5.4 Conditional Assertion

There is a long philosophical tradition of distinguishing *conditional* from *unconditional* assertions.[5] Quine famously says, for example:

> "An affirmation of the form 'if p then q' is [not] an affirmation of a conditional [but] a conditional affirmation of the consequent. If, after we have made such an affirmation, the antecedent turns out true, then we consider ourselves committed to the consequent. . . . [If] the antecedent turns out to have been false, our conditional affirmation is as if it had never been made" (Quine 1959, 12)

If conditional and unconditional assertions are distinct in important ways, as many philosophers have suggested, a question arises: does Stalnaker's characterization of the conditions under which an assertion is acceptable apply to both types?

Answering this question requires having some grip on the notion of conditional assertion. The argument for *Logical Presupposition* we are considering assumes the Stalnakerian idea that the presuppositions of conversational participants determine a common ground or context set, and that the essential functions of speech acts like asserting and questioning should be characterized in terms of the effects they have on the common ground. In order to address the argument on its own terms, then, I want to sketch a notion of conditional assertion that hews as closely to this framework as possible.

5. See Edgington (1986, 1995), and also Goldstein (2019a) for recent discussion.

The hypothesis I want to propose is simple: to make a conditional assertion (in the indicative mood) is to propose an update to a *part* of the context set — in particular, the part of the context set that would result from accepting the condition. We can imagine that when a speaker asserts $A \to B$, the conversational participants divide the context set into two parts: that which accepts A and that which does not. They then update the A-part in the normal way with B — e.g., when B is not itself a conditional, by keeping the worlds where B is true, and removing the worlds where B is false. After this update, the participants lump the two parts back together.

This simple sketch of conditional assertion can make sense of the remark made by Quine and many other theorists of conditional assertion that if, after asserting $A \to B$, the participants come to accept A, a commitment to B results. On the present model, that is because when $A \to B$ is accepted, A-part of the context set is updated so that it also accepts B. Assuming updates are monotonic, subsequent acceptance of A will thus necessitate acceptance of B. If, instead, A is *rejected*, then the participants will remove the A-part. Since the latter was the only part affected by the assertion of $A \to B$, the resulting context set is *ceteris paribus* as if the conditional had never been asserted in the first place.

The simple account of conditional assertion I have just sketched allows us to avoid the two worries about the interaction between trivalent theories of presupposition and conditionals. Consider the following sentences:

(5.2) If Isaac has a car, it is clean.

(5.3) Isaac's car is clean.

Suppose we combine the trivalent theory of presuppositional expressions with the trivalent theory of conditionals, and that (5.2) and (5.3) denote the same partial function from worlds to truth values. Both are undefined for worlds in which Isaac does not have a car. They are defined in worlds where Isaac has a car, and true if it is clean, false if it is not clean. On the account of conditional assertion I am outlining here, the assertions of these sentences have different essential effects on a context. Since (5.3) is unconditional, asserting it is to propose to change the context *as a whole* in the way that Stalnaker explains (as summarized in *Assertion*). This means that it is only acceptable in contexts that already accept that Isaac has a car. Since (5.2) is a conditional, on the other hand, asserting it is to propose a change to a *proper part* (a subset) of the context set — the part throughout which it is true that Isaac has a car. Accepting the speaker's proposal means the conversational participants update only that part with the consequent — they keep the worlds where Isaac's car is clean, and remove the worlds where it is not clean. Since the consequent of (5.2) is defined in every world where the antecedent is true, there is no possibility of the usual update operation breaking down — the semantic value of (5.2) gives enough information to sort through all of the worlds where Isaac has a car.

Distinguishing between the two speech acts thus allows us to draw a distinction between contexts in which asserting (5.2) is acceptable and those in which asserting (5.3) is acceptable, even if they denote the same partial function over worlds. The assertion of the unconditional sentence (5.3) is acceptable only if the sentence has a truth value at every world in the context. But the assertion of the conditional sentence (5.2) requires something far weaker: it is acceptable in a context so long as there is a *part* of the context where the antecedent holds, and the consequent is defined for every world in that part. Because their acceptability conditions are different, we may use *Presupposition Bridge* to tease apart their presuppositions. Since (5.3) is unconditional, it is acceptable only in contexts that accept that Isaac has a car. Thus, the sentence conventionally presupposes that Isaac has a car. Since (5.2) is acceptable in some contexts that do not accept that Isaac has a car, it does not conventionally presuppose this. Instead, it is acceptable only in contexts that have *some* worlds where Isaac has a car. This is just to say that it presupposes that its antecedent clause is compatible with the context.

The theory I am outlining here — a trivalent semantics for both presuppositional and suppositional expressions, coupled with a distinction between conditional and unconditional assertions, requires accepting that how the assertion of a sentence affects a context is not completely determined by the proposition it denotes. In determining what speech act a speaker intends to perform, a hearer must not only compute the semantic value of the speaker's sentence, she must also take into account pragmatic conventions governing the language the speaker used to generate it. In these respects, the account I am giving in this chapter is in the spirit of Stalnaker (1998) and Lewis (2012, 2014), which distinguish between semantic content and pragmatic force in a similar manner.

5.5 Proof of Concept

In this section, I outline a formal model on which conditional and unconditional sentences are both given a trivalent semantics, but each generate distinct presuppositions via their association with a distinct type of assertion.

Syntax. From a set of atomic sentences At, formulas ϕ are generated by the following grammar (with $p \in \mathsf{At}$):

$$\phi ::= p \mid \phi \to \psi$$

Partial Propositions. A partial proposition $P = \langle P^+, P^- \rangle$ will be modeled as a pair consisting of a set of worlds where it is true (P^+), and those where it is false (P^-). These pairs as a rule have no overlap (we disallow the possibility of truth value *gluts*), but they may not cover the whole set of worlds in a model. In other words, we do not require that $W \subseteq [P^+ \cup P^-]$.

Semantics. A model is a pair $\langle W, V \rangle$, consisting of a set of worlds W and a valuation function V that maps an atomic sentence $p \in \mathsf{At}$ to a partial proposition, $V(p) = \langle V^+(p), V^-(p) \rangle$.

The partial proposition denoted by a formula ϕ in a model $\mathcal{M}$ ($[\![\phi]\!]_{\mathcal{M}} = \langle [\![\phi]\!]^+_{\mathcal{M}}, [\![\phi]\!]^-_{\mathcal{M}} \rangle$)

is determined recursively as follows:

$$[\![p]\!]^+_{\mathcal{M}} = V^+(p) \quad (\text{for } p \in \mathsf{At})$$
$$[\![p]\!]^-_{\mathcal{M}} = V^-(p) \quad (\text{for } p \in \mathsf{At})$$
$$[\![\phi \to \psi]\!]^+_{\mathcal{M}} = [\![\phi]\!]^+_{\mathcal{M}} \cap [\![\psi]\!]^+_{\mathcal{M}}$$
$$[\![\phi \to \psi]\!]^-_{\mathcal{M}} = [\![\phi]\!]^+_{\mathcal{M}} \cap [\![\psi]\!]^-_{\mathcal{M}}$$

Dynamic Pragmatics.

Definition (Context)**.** Given a model $\mathcal{M}$, a context $c_{\mathcal{M}} \subseteq W_{\mathcal{M}}$ is a set of worlds, representing the worlds compatible with what is mutually accepted by the conversational participants.

Assertion will be modeled as a partial operation on the contexts of a model. The successful assertion of ϕ in a context c results in a new context $d = c + \phi$.

Definition (Assertion)**.** The update operation (+) is composed of two rules, corresponding to two different types of assertion:

Unconditional Assertion. For $p \in \mathsf{At}$, $c + p$ is defined iff $c \subseteq [V^+(p) \cup V^-(p)]$. When defined:

$$c + p := c \cap V^+(p) = c \backslash V^-(p)$$

Conditional Assertion. For formulas of the form $\phi \to \psi$, $c + \phi \to \psi$ is defined iff $(c + \phi) + \psi$ is defined. When defined:

$$c + \phi \to \psi = [c \backslash (c + \phi)] \cup [(c + \phi) + \psi]$$

Thus, the conditional assertion of $\phi \to \psi$ at c is the union of (i) whatever would be discarded by the assertion of ϕ, and (ii) the result of asserting ϕ and then ψ.

The result of updating a context c with ϕ thus depends not just on what partial proposition ϕ denotes, but also on whether ϕ is conditional. Let me illustrate how this system works with the following simple model, (*Four Worlds*) consisting of four worlds and a valuation function over three atomic sentences:

Four Worlds

	A	B	C	$A \rightarrow B$
w_1	1	1	1	1
w_2	1	0	0	0
w_3	0	1	#	#
w_4	0	0	#	#

Notice that C and $A \rightarrow B$ denote the same partial proposition, since we have:

$$[\![C]\!]^+ = [\![A \rightarrow B]\!]^+ = \{w_1\}$$

$$[\![C]\!]^- = [\![A \rightarrow B]\!]^- = \{w_2\}$$

Given their different syntax, however, the assertion of C and $A \rightarrow B$ are defined for different contexts. In particular, the assertion of the unconditional C is possible at a context c only if $c \subseteq [[\![C]\!]^+ \cup [\![C]\!]^-]$. So it is only acceptable in contexts that are subsets of $\{w_1, w_2\}$. $A \rightarrow B$ is more versatile. The contexts for which update with $A \rightarrow B$ is defined are any that *overlap* with $\{w_1, w_2\}$. This includes contexts like the null context, $\{w_1, w_2, w_3, w_4\}$. We can use these differences to systematically account for the differences in what C and $A \rightarrow B$ presuppose.

Definition (Acceptance)**.** A context c accepts ϕ iff $c + \phi = c$.

Definition (Acceptability)**.** A sentence ϕ is *acceptable* in a context c iff $c + \phi$ is defined.

Definition (Presuppositions of a sentence)**.** A sentence ϕ presupposes π iff every context in which ϕ is acceptable is one that accepts π.

The present model is enough to block our first worry: that conditionals might turn out to *presuppose* the truth of their antecedents. Consider the *Four Worlds* model. Since C is only acceptable in contexts that are subsets of $\{w_1, w_2\}$, and for every such context d, $d + A = d$, it follows that C presupposes A. But, even though $A \rightarrow B$ denotes the same partial proposition, it is acceptable some contexts that do not accept A. For example, it is acceptable in the null context $\{w_1, w_2, w_3, w_4\}$, which does not accept A, since $\{w_1, w_2, w_3, w_4\} + A = \{w_1, w_2\}$. Thus, unlike C, $A \rightarrow B$ does not presuppose A.

In one sense, our second worry — that indicative conditionals will not turn out to presuppose their antecedents are *compatible* with the context set — is blocked. For a conditional $\phi \rightarrow \psi$ is only acceptable in a context c if it is possible to accept the antecedent, i.e. only if $c + \phi$ is defined. In another sense, the second worry is not yet blocked, for a conditional like $A \rightarrow B$ is acceptable in the present system even in contexts like $\{w_3, w_4\}$, where A is accepted as *false*. But $A \rightarrow B$ seems to presuppose not just that A is acceptable in this vacuous sense, but rather that A *might be true*.

A simple solution is to add the following proviso: for any context c and sentence ϕ, $c + \phi$ is undefined when computing $c + \phi$ in the way outlined above would deliver the empty set — one cannot accept an assertion of ϕ if doing so would rule out *all* possibilities. With this additional requirement, we block the second worry in a stronger sense; a conditional $\phi \rightarrow \psi$ is acceptable only in contexts in which it is possible to accept ϕ *and not rule out all possibilities*. In other words, those in which ϕ might be true.

The extension of this model in §5.2, further strengthens our response to the second worry. In particular, it will ensure that sentences of the form $\phi \rightarrow \psi$ are only acceptable in contexts that accept $\Diamond\phi$.

5.6 Embeddings: Negations and Modals

The basic model I outlined in the last section can answer the two worries I discussed above for the interaction between the trivalent theories of conditionals and presuppositions. But a related worry looms. We may be left with the inability to distinguish between the way conditional and unconditional sentences embed under other expressions. This worry about *embedded conditionals* is a specific version of a more general worry for theories of indicative conditionals that place a lot of weight on their pragmatics.[6]

I think the trivalent account of conditionals may be able to answer this worry. In this section, I will address two kinds of embeddings that are often discussed in the literature — conditionals embedded under *negations* (§5.1), on the one hand, and *modals* (§5.2), on the other.

5.6.1 Negation

Suppose we were to add a negation operator to our grammar:

$$\phi ::= p \mid \neg\phi \mid \phi \to \psi$$

And suppose we extended the semantics of §4 with a Kleene semantics for negations:

$$[\![\neg\phi]\!]^+ = [\![\phi]\!]^-$$
$$[\![\neg\phi]\!]^- = [\![\phi]\!]^+$$

So $\neg\phi$ is true iff ϕ is false, and $\neg\phi$ is false iff ϕ is true.

How should we model the characteristic effect of asserting $\neg\phi$? We might try to treat negations as categorical, where $c + \neg\phi$ is defined iff it has a truth value in every world in c. When it's defined, $c + \neg\phi = c \cap [\![\phi]\!]^-$.

This semantics/pragmatics package would generate good results for negations of unconditional sentences like $\neg A$. But it generates bad results for negations of conditionals. Let me illustrate by way of our example, *Four Worlds*:

6. See Kolbel 2000 for discussion of the problem of embeddings as it arises for Edgington's theory.

	A	B	C	$A \to B$	$\neg C$	$\neg(A \to B)$
w_1	1	1	1	1	0	0
w_2	1	0	0	0	1	1
w_3	0	1	#	#	#	#
w_4	0	0	#	#	#	#

Our rule for negation entails we cannot update the null context $c = \{w_1, w_2, w_3, w_4\}$ with $\neg C$. Since C does not have a truth value in the worlds w_3 and w_4, $c + C$ is undefined, and thus so is $c \backslash (c + C)$. This seems like a good result: if we do not take for granted that Isaac has a car, then the assertion of 'Isaac's car isn't clean' is just as infelicitous as that of 'Isaac's car is clean.'

But now consider updating the null context with $\neg(A \to B)$. Intuitively, we should be able to do so: the felicity of 'It is not the case that if Isaac has a car, his car is clean' does not seem to depend on our accepting that Isaac has a car. But our rule for negations suggests otherwise: $c + \neg(A \to B)$ is undefined, since c includes worlds w_3 and w_4 where $\neg(A \to B)$ does not have a truth value.

On the present proposal, then, we have succeeded in distinguishing between the effects of asserting C and $A \to B$, but we fail to distinguish between the effects of $\neg C$ and $\neg(A \to B)$.

What *should* the correct result be? As Stalnaker 1981; Stalnaker 2011, Cariani and Goldstein 2018, Cariani and Santorio 2018, and others argue, the negation of a conditional is interpreted like a conditional formed from the same antecedent but with the negated consequent. In pragmatic terms: the denial of a conditional is a conditional denial. A natural remedy, then, is to introduce the operations of conditional and unconditional *denial* to shadow our modes of assertion. We will continue with the language that includes a negation given the semantics above. But now, we will use two update operations on contexts in a model: assertion ($+$) and denial ($-$).

Dynamic Pragmatics.

Definition (Assertion). *Assertion / Denial of Atoms*. For atomic $p \in \mathsf{At}$, $c + p$ and $c - p$ are defined only if p has a truth value at every world in c, i.e. iff:

$$c \subseteq [V^{+}(p) \cup V^{-}(p)]$$

. When defined:

$$c + p = c \cap V^{+}(p) = c \backslash V^{-}(p)$$
$$c - p = c \cap V^{-}(p) = c \backslash V^{+}(p)$$

Assertion / Denial of Negations. When ϕ is of the form $\neg\psi$, then $c + \phi$ is defined iff $c - \psi$ is; and $c - \phi$ is defined iff $c + \psi$ is. When defined:

$$c + \neg\psi = c - \psi$$
$$c - \neg\psi = c + \psi$$

Assertion / Denial of Conditionals. When ϕ is of the form $\psi \rightarrow \chi$, $c + \phi$ is defined iff $(c + \psi) + \chi$ is; while $c - \phi$ is defined iff $(c + \psi) - \chi$ is. When defined:

$$c + \psi \rightarrow \chi = [c \backslash (c + \psi)] \cup [(c + \psi) + \chi]$$
$$c - \psi \rightarrow \chi = [c \backslash (c + \psi)] \cup [(c + \psi) - \chi]$$

The main difference between the previous system and the present one is, of course, what happens with negated conditionals: they are now treated as conditional denials.

Let us revisit the result of asserting $\neg(A \to B)$ in the null context, c, of *Four Worlds*:

$$\begin{aligned} c + \neg(A \to B) &= c - (A \to B) \\ &= [c \backslash (c + A)] \cup [(c + A) - B] \\ &= [c \backslash \{w_1, w_2\}] \cup [\{w_1, w_2\} - B]) \\ &= \{w_3, w_4\} \cup [\{w_1, w_2\} - B]) \\ &= \{w_3, w_4\} \cup [\{w_1, w_2\} \cap \{w_2, w_4\}]) \\ &= \{w_3, w_4\} \cup \{w_2\} \\ &= \{w_2, w_3, w_4\} \end{aligned}$$

With a pragmatics of conditional and unconditional denial, the result is no longer a context that accepts A and accepts not-B. Rather, the result is more appropriately a context that has eliminated all A-worlds that that are also B-worlds, in particular, w_1. In other words, the result is a context that denies B on the condition that A.

5.6.2 Modals

Suppose we expanded the language of §4 to include possibility and necessity modals ($\Diamond, \Box$), with the grammar:

$$\phi ::= p \mid \phi \to \psi \mid \Diamond(\phi) \mid \Box(\phi)$$

For concreteness, I will interpret our modals as epistemic 'must' and 'might'.

Partial propositions and models will be defined as before. But in order to handle our epistemic modals, formulas will now denote partial propositions relative to a model *and a context set*, written $[\![\phi]\!]_{\mathcal{M},c}$, where $c \subseteq W_{\mathcal{M}}$ is a set of worlds representing the context set. For non-modals, the semantic clauses are the same as in §4, but with this new detail:

$$[\![p]\!]^+_{\mathcal{M},c} = V^+(p) \quad (\text{for } p \in \mathsf{At})$$

$$[\![p]\!]^-_{\mathcal{M},c} = V^-(p) \quad (\text{for } p \in \mathsf{At})$$

$$[\![\phi \to \psi]\!]^+_{\mathcal{M},c} = [\![\phi]\!]^+_{\mathcal{M},c} \cap [\![\psi]\!]^+_{\mathcal{M},c}$$

$$[\![\phi \to \psi]\!]^-_{\mathcal{M},c} = [\![\phi]\!]^+_{\mathcal{M},c} \cap [\![\psi]\!]^-_{\mathcal{M},c}$$

Now let's turn to the semantics of our epistemic modals. Suppose we drew on the semantics of Yalcin 2012 for simplicity:[7]

Definition (Simple Domain Clauses)**.**

$$[\![\Diamond\phi]\!]^+_{\mathcal{M},c} = \{v \in W \mid \text{ for some } v' \in c : \; v' \in [\![\phi]\!]^+_{\mathcal{M},c}\}$$

$$[\![\Diamond\phi]\!]^-_{\mathcal{M},c} = W \backslash [\![\Diamond\phi]\!]^+_{\mathcal{M},c}$$

$$[\![\Box\phi]\!]^+_{\mathcal{M},c} = \{v \in W \mid \text{ for all } v' \in c : \; v' \in [\![\phi]\!]^+_{\mathcal{M},c}\}$$

$$[\![\Box\phi]\!]^-_{\mathcal{M},c} = W \backslash [\![\Box\phi]\!]^+_{\mathcal{M},c}$$

Clauses like these lead to an unfortunate result. The semantics of our modals relies solely on the partial proposition denoted by their prejacent clauses. Thus, if C and $A \to B$ denote the same partial proposition, then we will not able to distinguish between the meanings of $\Box C$ and $\Box(A \to B)$. But these sorts of pairs in natural language clearly have different meanings. Consider the following:

(5.4) a. Isaac's car must be clean. $\Box C$

b. It must be that if Isaac has a car, it is clean. $\Box(A \to B)$

(5.4b) can be true even if we have not accepted that Isaac has a car, while (5.4a) requires that we have. What this shows, I think, is that our modal semantics must be sensitive

7. See also Holliday and Icard 2017.

to the conditional/unconditional distinction in just the same way that our pragmatics is. The semantic clauses in Definition 5.6.2 do not allow for this.

In response, I want to propose that for $\Box(A \to B)$ to be true, it has to be the case that B is true in all the worlds included in the context set *updated by supposing* A is true. For $\Box C$ to be true, alternatively, C must be true at every world in the context set. In order to predict these truth conditions within the present framework, I will make the semantics of our modals depend on two functions, DOMID and PROPID defined in terms of our pragmatics, the context set, and the syntax of the prejacent clause. DOMID will help us identify the restrictor domain of the modal — for $\Box(A \to B)$, the context set updated by supposing A, while PROPID will identify the nuclear scope of the modal — for $\Box(A \to B)$, the proposition denoted by B (interpreted relative to our supposition that A). I take these functions to model a sub-sentential pragmatics, used for deriving subordinate contexts and assertions within them.

First, our pragmatics. It will be the same as that of §4, except that we will make explicit that our modal sentences are associated with unconditional assertion.

DYNAMIC PRAGMATICS.

Definition (Assertion)**.** Update will be modeled as a change in context; an operation $(+)$ that takes a context set $c \subseteq W$ in a model $\mathcal{M}$ and a sentence ϕ to a new context set in the same model $d \subseteq W$.

Unconditional Assertion. For atomic ϕ, and ϕ of the form $\Diamond\psi$ or $\Box\psi$, $c + \phi$ is defined iff:

1. The truth of ϕ is compatible with c In other words: $c \cap [\![\phi]\!]^{+}_{\mathcal{M},c} \neq \varnothing$; and,

2. ϕ has a truth value at every world in c. In other words: $c \subseteq [[\![\phi]\!]^{+}_{\mathcal{M},c} \cup [\![\phi]\!]^{-}_{\mathcal{M},c}]$

When defined:

$$c + \phi := c \cap [\![\phi]\!]^{+}_{\mathcal{M},c}$$

Modals are asserted unconditionally, like atoms. Note that we have also made explicit the proviso discussed at the end of §4: the acceptability of an unconditional assertion of ϕ depends on the truth of ϕ being compatible with the context set.[8] Conditional assertion is defined just as in §4:

Conditional Assertion. When ϕ is of the form $\psi \to \chi$, $c + \psi \to \chi$ is defined iff $(c + \psi) + \chi$ is defined. When defined:

$$c + \psi \to \chi = [c \backslash (c + \psi)] \cup [(c + \psi) + \chi]$$

The DOMID function takes a context set c in a model $\mathcal{M}$ and a sentence ϕ to another context set d in that model:

function DOMID(c, ϕ)
  if ϕ is of the form $\psi \to \chi$ **then return** DOMID($c + \psi$, χ)
  else return c
  end if
end function

The PROPID function takes a context set c in a model $\mathcal{M}$ and a sentence ϕ to a partial proposition in the model $[\![\psi]\!]_{\mathcal{M},c}$:

function PROPID(c, ϕ)
  if ϕ is of the form $\psi \to \chi$ **then return** PROPID($c + \psi$, χ)
  else return $[\![\phi]\!]_{\mathcal{M},c}$
  end if

8. This plays a role in strengthening our response to the second worry. See Appendix B.

end function

Let me illustrate how DOMID and PROPID work with a few examples, where $A, B, C \in \mathsf{At}$ and ϕ is an arbitrary formula.

Unconditionals. Atoms and modals are fixed points:

$$\textsc{domid}(c, A) = c \quad \text{since } A \text{ is not of the form } \psi \to \chi$$

$$\textsc{propid}(c, A) = [\![A]\!]_{\mathcal{M},c} \quad \text{since } A \text{ is not of the form } \psi \to \chi$$

$$\textsc{domid}(c, \Box\phi) = c \quad \text{since } \Box\phi \text{ is not of the form } \psi \to \chi$$

$$\textsc{propid}(c, \Box\phi) = [\![\Box\phi]\!]_{\mathcal{M},c} \quad \text{since } \Box\phi \text{ is not of the form } \psi \to \chi$$

Simple Conditional: $A \to B$.

$$\begin{aligned}\textsc{domid}(c, A \to B) &= \textsc{domid}(c + A, B) \quad \text{since } A \to B \text{ is of the form } \psi \to \chi\\ &= c + A \quad \text{since } B \text{ is not of the form } \psi \to \chi\\ &= c \cap [\![A]\!]^{+}_{\mathcal{M},c} \quad \text{if } c + A \text{ is defined}\\ \textsc{propid}(c, A \to B) &= \textsc{propid}(c + A, B) \quad \text{since } A \text{ is not of the form } \psi \to \chi\\ &= [\![B]\!]_{\mathcal{M},c+A} \quad \text{since } B \text{ is not of the form } \psi \to \chi\end{aligned}$$

Simple Conditional Consequent: $A \to (B \to C)$.

$\textsc{domid}(c, A \to (B \to C)) =$

$$\begin{aligned}&= \textsc{domid}(c + A, B \to C) \quad \text{since } A \to (B \to C) \text{ is of the form } \psi \to \chi\\ &= \textsc{domid}((c + A) + B, C) \quad \text{since } B \to C \text{ is of the form } \psi \to \chi\\ &= (c + A) + B \quad \text{since } C \text{ is not of the form } \psi \to \chi\end{aligned}$$

$\textsc{propid}(c, A \to (B \to C)) =$

$$= \textsc{propid}(c + A, B \to C) \quad \text{since } A \to (B \to C) \text{ is of the form } \psi \to \chi$$

$$= \textsc{propid}((c + A) + B, C) \quad \text{since } B \to C \text{ is of the form } \psi \to \chi$$

$$= [\![C]\!]_{\mathcal{M},(c+A)+B} \quad \text{since } C \text{ is not of the form } \psi \to \chi$$

Simple Conditional Antecedent: $(A \to B) \to C$.

$\textsc{domid}(c, (A \to B) \to C) =$

$$= \textsc{domid}(c + (A \to B), C) \quad \text{since } (A \to B) \to C \text{ is of the form } \psi \to \chi$$

$$= c + (A \to B) \quad \text{since } C \text{ is not of the form } \psi \to \chi$$

$$= [c \backslash (c + A)] \cup [(c + A) + B] \quad \text{if } (c + A) + B \text{ is defined}$$

$\textsc{propid}(c, (A \to B) \to C) =$

$$= \textsc{propid}(c + (A \to B), C) \quad \text{since } (A \to B) \to C \text{ is of the form } \psi \to \chi$$

$$= [\![C]\!]_{\mathcal{M},c+(A \to B)} \quad \text{since } C \text{ is not of the form } \psi \to \chi$$

Using these functions, we can now give a semantics for our modals that avoids the problem of identifying the meanings of $\Box C$ and $\Box(A \to B)$ when the prejacent clauses denote the same partial proposition:

Definition (Pragmatic Domain Clauses)**.**

$$[\![\Diamond\phi]\!]^{+}_{\mathcal{M},c} = \{v \in W \mid \text{ for some } v' \in \textsc{domid}(c, \phi) : \ v' \in \textsc{propid}(c, \phi)^{+}\}$$

$$[\![\Diamond\phi]\!]^{-}_{\mathcal{M},c} = W \backslash [\![\Diamond\phi]\!]^{+}_{\mathcal{M},c}$$

$$[\![\Box\phi]\!]^{+}_{\mathcal{M},c} = \{v \in W \mid \text{ for all } v' \in \textsc{domid}(c, \phi) : \ v' \in \textsc{propid}(c, \phi)^{+}\}$$

$$[\![\Box\phi]\!]^{-}_{\mathcal{M},c} = W \backslash [\![\Box\phi]\!]^{+}_{\mathcal{M},c}$$

Let us now consider how the present system distinguishes between the truth conditions of $\Box C$ and $\Box(A \to B)$ even if the prejacent clauses denote the same partial proposition. Consider *Four Worlds* again:

	A	B	C	$A \to B$
w_1	1	1	1	1
w_2	1	0	0	0
w_3	0	1	#	#
w_4	0	0	#	#

Even though $A \to B$ and C denote the same partial proposition, their modal embeddings are different. Let $c = \{w_1, w_3, w_4\}$. Consider $\Box(A \to B)$ interpreted relative to c:

$$\begin{aligned}
[\![\Box(A \to B)]\!]^+_{\mathcal{M},c} &= \{v \in W \mid \text{ for all } v' \in \textsc{domid}(c, A \to B) : \ v' \in \textsc{propid}(c, A \to B)^+\} \\
&= \{v \in W \mid \text{ for all } v' \in \textsc{domid}(c + A, B) : \ v' \in \textsc{propid}(c + A, B)^+\} \\
&= \{v \in W \mid \text{ for all } v' \in c + A : \ v' \in [\![B]\!]^+_{\mathcal{M},c+A}\}^+\} \\
&= \{v \in W \mid \text{ for all } v' \in \{w_1\} : \ v' \in \{w_1\}\} \\
&= W
\end{aligned}$$

Thus, $\Box(A \to B)$ is true at every world when c is the context, since B is true throughout the A-part of the context.

Meanwhile, consider $\Box C$, where C denotes the same partial proposition in our model:

$$\begin{aligned}
[\![\Box C]\!]^+_{\mathcal{M},c} &= \{v \in W \mid \text{ for all } v' \in \textsc{domid}(c, C) : \ v' \in \textsc{propid}(c, C)^+\} \\
&= \{v \in W \mid \text{ for all } v' \in c : \ v' \in [\![C]\!]^+_{\mathcal{M},c}\} \\
&= \{v \in W \mid \text{ for all } v' \in \{w_1, w_3, w_4\} : \ v' \in \{w_1\}\} \\
&= \varnothing
\end{aligned}$$

$\Box C$, then, is not true at any world when c is the context, because some worlds in c do not make C true.

The present framework, then, succeeds in distinguishing between the propositions denoted by $\Box(A \rightarrow B)$ and $\Box C$, even if their prejacent clauses denote the same partial propositions. And it seems to do so in the right way: by assigning intuitively correct, distinct truth conditions to each sentence.

Furthermore, I note (the proof is contained in Appendix B) that the present system strengthens our response to the second worry I discussed for the combination of the two trivalent theories: $c + \phi \rightarrow \psi$ is defined only if c accepts $\Diamond\phi$ (Proposition 6 in Appendix B). Given the presupposition principles in §4, then, sentences of the form $\phi \rightarrow \psi$ officially presuppose $\Diamond\phi$. In other words, the indicative conditional $\phi \rightarrow \psi$ presupposes the object-language statement $\Diamond\phi$.

5.7 Conclusion

I began by noting some reasons theorists have been interested in trivalent truth-functional accounts of indicative conditionals, and two worries they have raised. I argued that we can answer both of these worries by distinguishing between the essential effects of conditional and unconditional assertions. In §4, I outlined a simple formal system that allowed us to make this distinction precise. In §5, I noted that a third worry loomed for the account: that it would not be able to distinguish between the semantics of complex sentences containing presuppositions from those containing conditionals. In response, I outlined two different extensions of the framework to handle embeddings under negation and under a modal.

Many questions remain open. For one, conditionals can embed in many contexts besides the two I have discussed here — can a trivalent theory account for all of them? For another, can the present theory be supplemented with an account of mood in order to generate plausible predictions for subjunctive conditionals? While I will not address these questions further in this dissertation, I do hope that my defense of the trivalent theory of indicative conditionals suggests that they may be worth pursuing in future research.

Conclusion

In this dissertation, I have argued that modal claims can be true even when they underspecify the contents of modal states. As I noted in the introduction, this claim has faced two significant challenges. One was that existing semantic accounts of modals that allow for this form of underspecificity have seemed to deliver incorrect truth-value judgments in cases involving complex forms of uncertainty and preference. The other was that existing accounts of modals were committed to naïve upward monotonicity, which generates several paradoxes for the interaction between modals, on the one hand, and disjunctions, presuppositions, or conditionals, on the other. I have outlined solutions to these paradoxes in chapters 3 and 4 that improve on empirical predictions of existing solutions while still allowing for the possibility of true, underspecific modal claims. Furthermore, in the case of 'want', in particular, I argued in chapter 2 that giving an upward monotonic semantics directly in terms of the contents of an agent's desire(s) allows us to account for truth-value judgments in cases of complex uncertainty and preference without giving up the possibility of underspecificity. While I have not argued for this here, I take the lesson of that chapter to be generalizable to decision-theoretic semantics proposed for 'ought' and other modals.

There are several respects in which further research would be desirable. I have noted some along the way, and I would like to close by mentioning three I have not yet mentioned. I hope to address some of these questions in future work.

First, the semantic accounts I have given to resolve the puzzles of chapters 3 and 4 reject the idea that modals are what I called in the introduction *naïvely* upward monotonic. But the rejection was conservative, still allowing for true, underspecific modal claims, and predicting the data I have discussed in support of some form of upward monotonicity. Thus, one question that deserves further research is whether the semantic frameworks developed in chapters 3 and 4 may help us refine the definition of entailment relevant for explaining the sense in which modals are upward monotonic. For example, the account of chapter 3 shows that whether $\Delta(\phi)$ entails $\Delta(\psi)$ depends not just on whether ϕ truth conditionally entails ψ, but also how the alternatives offered by ϕ and ψ are related. Similarly, chapter 4 shows that whether $\Delta(\phi)$ entails $\Delta(\psi)$ depends on how the presuppositions of ϕ and ψ are related.

A second issue deserving of further research is whether the account of the Samaritan paradoxes in chapter 4 can shed light on the theory of conditional attitudes and the semantics of conditional attitude ascriptions. For example, conditional attitudes are sometimes thought of as special because they make a three-fold distinction between worlds: those that satisfy the state, those that frustrate it, and those that neither satisfy nor frustrate it.[9] Extending my account of the Samaritan paradox to attitude verbs (e.g. 'want') suggests a view on which the attitudes we ascribe with presupposition-triggering complement clauses (e.g., 'I want to catch the burglar') are *also* 'conditional' in this sense, even though we need not use any conditional constructions in reporting them. I think that recognizing these other ways of reporting 'conditional' attitudes may help us clarify the idea of a conditional attitude and help us separate the semantic and metaphysical issues involved in theorizing about them. Furthermore, I think it would be worth exploring whether extending the account I gave of the conditional version of the Samaritan paradox to conditional *attitude* ascriptions might explain some of their perplexing behaviors.[10]

9. See, for example, McDaniel and Bradley (2008) on conditional desires, Ferrero (2009) on conditional intentions, and others.

10. See Jerzak (2019b) and Blumberg and Holguín (2019) for recent discussion.

Finally, I think it would be especially helpful to explore how discourse-level mechanisms of meaning affect the interpretation of underspecific modal claims. Modal claims (including attitude ascriptions) are often interpreted anaphorically, commenting further on possibilities introduced by prior parts of a longer discourse.[11] Since such possibilities are usually introduced with underspecific modal claims, it would be helpful to understand the interpretive resources interlocutors draw upon in order to coordinate on and keep track of these underspecified possibilities. Furthermore, the anaphoric behavior of modals seems to be affected by rhetorical relations between claims.[12] For example, the interpretation of a discourse of the form ⌜$\Delta(\phi)$ and $\Delta(\psi)$⌝ may depend on whether the rhetorical function of the second conjunct, $\Delta(\psi)$ is to elaborate on what else is true, given $\Delta(\phi)$, or to introduce a parallel, comparative remark. These differences may have semantic ramifications. For example, if the rhetorical relation is *elaboration*, the modal parameter relevant for the interpretation of $\Delta(\psi)$ might differ from that which is relevant for the interpretation of the first conjunct $\Delta(\phi)$, while if the relation is *parallel*, both conjuncts might be interpreted with respect to the same parameter. Clearly, a full understanding how we use modal claims to underspecify the modal facts would include an understanding of discourse-level effects like these.

11. See, for example, Roberts (1989), Stone (1999).

12. As argued by Stojnić (2017, 2019, 2021).

References

Abreu Zavaleta, Martín. 2019. "Weak Speech Reports." *Philosophical Studies* 176 (8): 2139–2166.

Aher, Martin. 2012. "Free Choice in Deontic Inquisitive Semantics (DIS)." In *Logic, Language and Meaning*, edited by Maria Aloni, Vadim Kimmelman, Floris Roelofsen, Galit W. Sassoon, Katrin Schulz, and Matthijs Westera, 22–31. Berlin, Heidelberg: Springer Berlin Heidelberg.

Aloni, Maria. 2018. "FC Disjunction in State-Based Semantics."

———. 2007. "Free Choice, Modals, and Imperatives." *Natural Language Semantics* 15 (1): 65–94.

Aloni, Maria, and Ivano Ciardelli. 2013. "A Logical Account of Free Choice Imperatives." In *The Dynamic, Inquisitive, and Visionary Life of ϕ, ?ϕ , and $\Diamond\phi$: a Festschrift for Jeroen Groenendijk, Martin Stokhof, and Frank Veltman*, edited by M. Frank M. Aloni and F. Roelofsen, 1–17.

Alonso-Ovalle, Luis. 2006. *Disjunction in Alternative Semantics.* PhD Dissertation, University of Massachusetts Amherst.

Anand, Pranav, and Valentine Hacquard. 2013. "Epistemics and attitudes." *Semantics and Pragmatics* 6.

Asher, Nicholas. 1987. "A Typology for Attitude Verbs and Their Anaphoric Properties." *Linguistics and Philosophy* 10 (2): 125–197.

Ashwell, Lauren. 2009. *Desires and Dispositions.* PhD Dissertation, Massachussetts Institute of Technology.

———. 2014. "The Metaphysics of Desire and Dispositions." *Philosophy Compass* 9 (7): 469–477.

Barker, Chris. 2010. "Free choice permission as resource-sensitive reasoning." *Semantics and Pragmatics* 3 (10): 1–38.

Beaver, David, and Emiel Krahmer. 2001. "A Partial Account of Presupposition Projection." *Journal of Logic, Language and Information* 10 (2): 147–182.

Belnap, Nuel, and Michael Perloff. 1993. "In the Realm of Agents." *Annals of Mathematics and Artificial Intelligence* 9:25–48.

Belnap, Nuel D. 1970. "Conditional Assertion and Restricted Quantification." *Noûs* 4 (1): 1–12.

Berto, Franz, and Mark Jago. 2019. *Impossible Worlds.* Oxford: Oxford University Press.

Bjerring, Jens Christian, and Wolfgang Schwarz. 2017. "Granularity Problems." *Philosophical Quarterly* 67 (266): 22–37.

Blumberg, Kyle, and John Hawthorne. Forthcoming(a). "A New Hope." *Journal of Philosophy.*

———. Forthcoming(b). "Inheritance: Professor Procrastinate and the Logic of Obligation." *Philosophy and Phenomenological Research.*

———. Forthcoming(c). "Wanting What's Not Best." *Philosophical Studies,* 1–22.

Blumberg, Kyle, and Ben Holguín. 2019. "Embedded Attitudes." *Journal of Semantics* 36 (3): 377–406.

Blumberg, Kyle, and Ben Holguín. 2019. "Embedded Attitudes." Ffz004, *Journal of Semantics,* eprint: http://oup.prod.sis.lan/jos/advance-article-pdf/doi/10.1093/jos/ffz004/28080311/ffz004.pdf.

Brasoveanu, Adrian, and Donka F. Farkas. 2007. "Say Reports, Assertion Events and Meaning Dimensions." *Pitar Mos: A Building with a View. Papers in Honour of Alexandra Cornilescu, Editura Universitatii din Bucuresti, Bucharest,* 175–196.

Bratman, Michael. 1987. *Intention, Plans, and Practical Reason.* Cambridge, MA: Harvard University Press.

Braun, David. 2015. "Desiring, Desires, and Desire Ascriptions." *Philosophical Studies* 172 (1): 141–162.

Buchak, Lara. 2013. *Risk and Rationality.* Oxford University Press.

Cantwell, John. 2008. "Indicative Conditionals: Factual or Epistemic?" *Studia Logica* 88 (1): 157–194.

———. 2006. "The Laws of Non-Bivalent Probability." *Logic and Logical Philosophy* 15 (2): 163–171.

Cariani, Fabrizio. 2013. "'Ought' and Resolution Semantics." *Noûs* 47 (3): 534–558.

Cariani, Fabrizio, and Simon Goldstein. 2018. "Conditional Heresies." *Philosophy and Phenomenological Research,* no. 2, 251–282.

Cariani, Fabrizio, and Paolo Santorio. 2018. "Will Done Better: Selection Semantics, Future Credence, and Indeterminacy." *Mind* 127 (505): 129–165.

Carr, Jennifer. 2014. "The If P, Ought P Problem." *Pacific Philosophical Quarterly* 95 (4): 555–583.

Carston, Robyn. 1996. "Metalinguistic Negation and Echoic Use." *Journal of Pragmatics* 25:309–330.

Chellas, Brian F. 1980. *Modal Logic: An Introduction.* Cambridge University Press.

Chemla, Emmanuel. 2009. "Universal implicatures and free choice effects: experimental data." *Semantics and Pragmatics* 2:1–33.

Chemla, Emmanuel, and Lewis Bott. 2014. "Processing inferences at the semantics/pragmatics frontier: disjunctions and free choice." *Cognition* 130 (3): 380–396.

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen. 2018. *Inquisitive Semantics.* Oxford University Press.

Ciardelli, Ivano, and Floris Roelofsen. 2017a. "Hurford Constraint, the Semantics of Disjunction, and the Nature of Alternatives." *Natural Language Semantics* 25 (3): 199–222.

———. 2017b. "Hurford's Constraint, the Semantics of Disjunction, and the Nature of Alternatives." *Natural Language Semantics* 25 (3): 199–222.

———. 2011. "Inquisitive Logic." *Journal of Philosophical Logic* 40 (1): 55–94.

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler. 2017. "Composing Alternatives." *Linguistics and Philosophy* 40 (1): 1–36.

Ciardelli, Ivano, Linmin Zhang, and Lucas Champollion. 2018. "Two Switches in the Theory of Counterfactuals: A Study of Truth Conditionality and Minimal Change." *Linguistics and Philosophy,* no. 6.

Cooper, William S. 1968. "The Propositional Logic of Ordinary Discourse." *Inquiry: An Interdisciplinary Journal of Philosophy* 11 (1-4): 295–320.

Coppock, Elizabeth, and David Beaver. 2013. "Mere-ology." In *Alternatives in Semantics,* edited by Anamaria Fluş, 150–173. London: Palgrave Macmillan UK. ISBN: 978-1-137-31724-7.

De Finetti, Bruno. 1936. "La logique de la probabilité." *Actes du congrès international de philosophie scientifique* 4:1–9.

Del Pinal, Guillermo. 2018. "Meaning, Modulation, and Context: A Multidimensional Semantics for Truth-Conditional Pragmatics." *Linguistics and Philosophy* 41 (2): 165–207.

Dowty, David. 1987. "Collective Predicates, Distributive Predicates, and *All*." In *Proceedings of the 3rd ESCOL, Ohio State University, Ohio,* edited by Fred Marshall, 97–115. The University of Pittsburgh.

Drucker, Daniel. 2017. "Policy Externalism." *Philosophy and Phenomenological Research* 94 (3).

Edgington, Dorothy. 1986. "Do Conditionals Have Truth-Conditions." *Cr'itica* 18 (52): 3–30.

———. 1995. "On Conditionals." *Mind* 104 (414): 235–329.

Égré, Paul, Lorenzo Rossi, and Jan Sprenger. 2021a. "De Finettian Logics of Indicative Conditionals Part I: Trivalent Semantics and Validity." *Journal of Philosophical Logic* 50 (2): 187–213.

———. 2021b. "De Finettian Logics of Indicative Conditionals Part II: Proof Theory and Algebraic Semantics." *Journal of Philosophical Logic* 50 (2): 215–247.

Fara, Delia Graff. 2013. "Specifying Desires." *Noûs* 47 (2): 250–272.

Fauconnier, Gilles. 1975. "Pragmatic scales and logical structure." *Linguistic Inquiry,* 353–375.

Ferrero, Luca. 2009. "Conditional Intentions." *Noûs* 43 (4): 700–741.

Fine, Kit. 2017a. "A Theory of Truthmaker Content I: Conjunction, Disjunction and Negation." *Journal of Philosophical Logic* 46 (6): 625–674.

———. 2017b. "A Theory of Truthmaker Content II: Subject-Matter, Common Content, Remainder and Ground." *Journal of Philosophical Logic* 46 (6): 675–702.

Fodor, Jerry A. 1978. "Propositional Attitudes." *The Monist* 61 (October): 501–23.

Fox, Danny. 2007. "Free Choice and the Theory of Scalar Implicatures* MIT,"

Frank, Anette. 1997. *Context Dependence in Modal Constructions.* PhD Dissertation, Universitat Stuttgart.

Frege, Gottlob. 1948. "Sense and Reference." *Philosophical Review* 57 (3): 209–230.

Fusco, Melissa. 2015. "Deontic Modality and the Semantics of Choice." *Philosophers' Imprint* 15.

Gazdar, Gerald. 1979. *Pragmatics: Implicature, presupposition, and logical form.* Academic Press.

Giannakidou, Anastasia. 1999. "Affective dependencies." *Linguistics and Philosophy,* 367–421.

Goldstein, Simon. 2019a. "A Theory of Conditional Assertion." *Journal of Philosophy* 116 (6): 293–318.

———. 2019b. "Free Choice and Homogeneity." *Semantics and Pragmatics* 12:1–48.

Gotzner, Nicole, Jacopo Romoli, and Paolo Santorio. 2020. "Choice and Prohibition in Non-Monotonic Contexts." *Natural Language Semantics* 28 (2): 141–174.

Grant, Lyndal, and Milo Phillips-Brown. 2020. "Getting What You Want." *Philosophical Studies,* 1791–1810.

Grice, H. P. 1989. *Studies in the Way of Words.* Cambridge: Harvard University Press.

Groenendijk, Jeroen, and Floris Roelofsen. 2010. "Radical Inquisitive Semantics."

Hacquard, Valentine. 2010. "On the Event Relativity of Modal Auxiliaries." *Natural Language Semantics* 18 (1): 79–114.

Harley, Heidi. 2012. "Lexical Decomposition in Modern Syntactic Theory." In *The Oxford Handbook of Compositionality,* edited by Edouard Machery Wolfram Hinzen and Markus Werning, 328–350. Oxford University Press.

Heim, Irene. 1992. "Presupposition Projection and the Semantics of Attitude Verbs." *Journal of Semantics* 9 (3): 183–221.

———. 1982. "The Semantics of Definite and Indefinite Noun Phrases." PhD diss., UMass Amherst.

Heim, Irene, and Angelika Kratzer. 1998. *Semantics in Generative Grammar.* Blackwell.

Hintikka, Jaako. 1969. "Semantics for propositional attitudes." In *Philosophical Logic,* edited by D.J. Hockney J.W. Davis and W.K. Wilson, 21–45. Dordrecht: Reidel.

Holliday, Wesley H., and Thomas Icard. 2017. "Indicative Conditionals and Dynamic Epistemic Logic." In *Proceedings of the Sixteenth Conference on Theoretical Aspects of Ra-*

tionality and Knowledge (TARK 2017), Liverpool, UK, 24-26 July 2017, edited by Jérôme Lang, 337–351.

Horn, Laurence. 1989. *A Natural History of Negation.* University of Chicago Press.

Huitink, Janneke. 2008. *Modals, Conditionals and Compositionality.* PhD Dissertation, Radboud Universiteit Nijmegen.

———. 2012. "Quantified Conditionals and Compositionality." *Language and Linguistics Compass* 1 (4): 42–53.

Hurford, James R. 1974. "Exclusive or Inclusive Disjunction." *Foundations of Language* 11 (3): 409–411.

Jackson, Frank. 1979. "On Assertion and Indicative Conditionals." *Philosophical Review* 88 (4): 565–589.

Jackson, Frank, and Robert Pargetter. 1986. "Oughts, Options, and Actualism." *Philosophical Review* 95 (2): 233–255.

Jerzak, Ethan. 2019a. "Two Ways to Want?" *Journal of Philosophy* 116 (2): 65–98.

———. 2019b. "Two Ways to Want?" *Journal of Philosophy* 116 (2): 65–98.

Kahneman, Daniel, and Amos Tversky. 1979. "Prospect Theory: An Analysis of Decision under Risk." *Econometrica* 47 (2): 263–291.

Kamp, Hans. 1973. "Free Choice Permission." *Proceedings of the Aristotelian Society* 74 (1): 57–74.

Karttunen, Lauri, and Stanley Peters. 1979. "Conventional Implicature." In *Syntax and Semantics.*

Katzir, Roni, and Raj Singh. 2013. "Hurford disjunctions: embedded exhaustification and structural economy." *Sinn und Bedeutung* 18:201–216.

Kolbel, M. 2000. "Edgington on Compounds of Conditionals." *Mind* 109 (433): 97–108.

Kratzer, Angelika. 2012a. "Conditionals." In *Modals and Conditionals*, 86–108. Oxford University Press.

———. 2012b. "Conditionals." In *Modals and Conditionals,* 86–108. Oxford University Press.

Kratzer, Angelika. 1991. "Modality." In *Semantics: An International Handbook of Contemporary Research,* edited by Arnim von Stechow & Dieter Wunderlich, 639–650. Berlin: de Gruyter.

———. 1981. "Partition and Revision: The Semantics of Counterfactuals." *Journal of Philosophical Logic* 10 (2): 201–216.

———. 2012c. "The Notional Category of Modality." In *Modals and Conditionals,* 27–69. Oxford University Press.

———. 2012d. "The Notional Category of Modality." In *Modals and Conditionals,* 27–69. Oxford University Press.

———. 1977. "What 'Must' and 'Can' Must and Can Mean." *Linguistics and Philosophy* 1 (3): 337–355.

Kratzer, Angelika, and Junko Shimoyama. 2002. "Indeterminate Pronouns: The View from Japanese." *Proceedings of the Third Tokyo Conference on Psycholinguistics.*

Kripke, Saul A. 1963. "Semantical Considerations on Modal Logic." *Acta Philosophica Fennica* 16 (1963): 83–94.

Križ, Manuel. 2015. *Aspects of Homogeneity in the Semantics of Natural Language.* PhD Thesis, University of Vienna.

———. 2016. "Homogeneity, Non-Maximality, Andall." *Journal of Semantics* 33 (3): 493–539.

Ladusaw, William. 1980. "On the notion 'affective' in the analysis of negative polarity items." *Journal of Linguistic Research,* 1–16.

Lasersohn, Peter. 1990. "Group Action and Spatio-Temporal Proximity." *Linguistics and Philosophy* 13 (2): 179–206.

Lassiter, Daniel. 2011. "Nouwen's Puzzle and a Scalar Semantics for Obligations, Needs, and Desires." *Proceedings of SALT* 21:694–711.

———. 2019. "What we can learn from how trivalent conditionals avoid triviality." *Inquiry.*

Levinson, Dmitry. 2003. "Probabilistic model-theoretic semantics for *want*." *Proceedings of SALT* 13:222–239.

Lewis, David K. 1998. "Adverbs of Quantification." In *Papers in Philosophical Logic*, 5–20. Cambridge University Press.

Lewis, David K. 1973. *Counterfactuals.* Blackwell.

Lewis, Karen S. 2012. "Discourse Dynamics, Pragmatics, and Indefinites." *Philosophical Studies* 158 (2): 313–342.

———. 2014. "Do We Need Dynamic Semantics?" In *Metasemantics: New Essays on the Foundations of Meaning,* edited by Alexis Burgess and Brett Sherman, 231–258. Oxford University Press.

Link, Godehard. 1983. "The Logical Analysis of Plurals and Mass Terms: A Lattice-Theoretic Approach." In *Formal Semantics - the Essential Readings,* edited by P. Portner and B. H. Partee, 127–147. Blackwell.

Lycan, William G. 2012. "Desire Considered as a Propositional Attitude." *Philosophical Perspectives* 26 (1): 201–215.

Mandelkern, Matthew. 2016. "Dissatisfaction theory." *Proceedings of SALT* 26.

Marino, Patricia. 2009. "On Essentially Conflicting Desires." *Philosophical Quarterly* 59 (235): 274–291.

Marty, Paul, Jacopo Romoli, Yasutada Sudo, and Richard Breheny. ms. "Negative Free Choice."

Matthewson, Lisa. 2016. "Modality." In *The Cambridge Handbook of Formal Semantics,* edited by Maria Aloni and PaulEditors Dekker, 525–559. Cambridge Handbooks in Language and Linguistics. Cambridge University Press.

McDaniel, Kris, and Ben Bradley. 2008. "Desires." *Mind* 117 (466): 267–302.

McDermott, Michael. 1996. "On the Truth Conditions of Certain 'If'-Sentences." *Philosophical Review* 105 (1): 1–37.

Menéndez-Benito, Paula. 2010. "On Universal Free Choice Items." *Natural Language Semantics* 18 (1): 33–64.

———. 2005. *The Grammar of Choice.* Ph.D. thesis, University of Massachusetts, Amherst.

Meyer, Marie-Christine. 2014. "Deriving Hurford's constraint." *Proceedings of SALT* 24:577–596.

———. 2013. *Ignorance and Grammar.* PhD Thesis, MIT.

Milne, Peter. 1997. "Bruno de Finetti and the Logic of Conditional Events." *British Journal for the Philosophy of Science* 48 (2): 195–232.

Moltmann, Friederike. 2004. "The Semantics of Together." *Natural Language Semantics* 12 (4): 289–318.

Nickel, Bernhard. 2010. "Generically Free Choice." *Linguistics and Philosophy* 33 (6): 479–512.

Nygren, Karl. 2019. "Supercover Semantics for Deontic Action Logic." *Journal of Logic, Language and Information* 28 (3): 427–458.

Pacuit, Eric. 2017. *Neighborhood Semantics for Modal Logic.* Springer International Publishing.

Partee, Barbara, and Mats Rooth. 1983. "Generalized Conjunction and Type Ambiguity." In *Meaning, Use, and Interpretation of Language,* edited by Christoph Schwarze Rainer Bauerle and Arnim von Stechow, 361–383.

Peters, Stanley. 1979. "A Truth-Conditional Formulation of Karttunen's Account of Presupposition." *Synthese* 40 (2): 301–316.

Phillips-Brown, Milo. 2021. "What Does Decision Theory Have to Do with Wanting?" *Mind* 130 (518): 413–437.

Portner, Paul. 2009. *Modality.* Oxford University Press.

———. 2018. *Mood.* Oxford University Press.

Prior, Arthur. 1958. "Escapism: The Logical Basis of Ethics." In *Essays in Moral Philosophy,* edited by A.I. Melden, 135–146.

Quine, W. V. O. 1959. *Methods of Logic.* 2nd ed. Harvard University Press.

Recanati, François. 2010. *Truth-Conditional Pragmatics.* Oxford University Press.

Roberts, Craige. 1989. "Modal Subordination and Pronominal Anaphora in Discourse." *Linguistics and Philosophy* 12 (6): 683–721.

Roelofsen, Floris. 2013. "Algebraic Foundations for the Semantic Treatment of Inquisitive Content." *Synthese* 190 (S1): 1–24.

Romoli, Jacopo, and Paolo Santorio. 2019. "Filtering free choice." *Semantics and Pragmatics* 12 (12): 1–27.

———. 2017. "Probability and implicatures: A unified account of the scalar effects of disjunction under modals." *Semantics and Pragmatics* 10 (3): 1–61.

Ross, Alf. 1941. "Imperatives and Logic." *Philosophy of Science* 11 (1): 30–46.

Rothschild, Daniel. 2014. "Capturing the Relationship Between Conditionals and Conditional Probability with a Trivalent Semantics." *Journal of Applied Non-Classical Logics* 24 (1-2): 144–152.

Santorio, Paolo. 2018. "Alternatives and Truthmakers in Conditional Semantics." *Journal of Philosophy* 115 (10): 513–549.

Sayre-McCord, Geoffrey. 1986. "Deontic Logic and the Priority of Moral Theory." *Noûs* 20 (2): 179–197.

Schroeder, Mark. 2007. *Slaves of the Passions.* Oxford University Press.

Schroeder, Tim. 2017. "Desire." In *The Stanford Encyclopedia of Philosophy,* Summer 2017, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Schroeder, Timothy. 2004. *Three Faces of Desire.* Oxford University Press.

Schwarzschild, Roger. 1993. "Plurals, Presuppositions and the Sources of Distributivity." *Natural Language Semantics* 2 (3): 201–248.

Searle, John R. 1983. *Intentionality: An Essay in the Philosophy of Mind.* Cambridge University Press.

Sergot, Marek. 2021. "Some Forms of Collectively Bringing About or 'Seeing to It That'." *Journal of Philosophical Logic* 50 (2): 249–283.

Sharvit, Yael. 2017. "A Note on (Strawson) Entailment." *Semantics and Pragmatics* 10.

Simons, Mandy. 2001. "Disjunction and Alternativeness." *Linguistics and Philosophy* 24 (5): 597–619.

———. 2005a. "Dividing Things Up: The Semantics of or and the Modal/or Interaction." *Natural Language Semantics* 13 (3): 271–316.

———. 2005b. "Semantics and Pragmatics in the Interpretation of Or." *Proceedings of SALT* 15:205–222.

Sinhababu, Neil. 2009. "The Humean Theory of Motivation Reformulated and Defended." *Philosophical Review* 118 (4): 465–500.

Smith, Michael. 1987. "The Humean Theory of Motivation." *Mind* 96 (381): 36–61.

———. 1994. *The Moral Problem.* Blackwell.

Soames, Scott. 1989. “Semantics and Semantic Competence.” *Philosophical Perspectives* 3 (n/a): 575–596.

Stalnaker, Robert. 1988. “Belief Attribution and Context.” In *Contents of Thought,* edited by Robert H. Grimm and D. D. Merrill, 140–156. University of Arizona Press.

———. 2011. “Conditional Propositions and Conditional Assertions.” In *Epistemic Modality,* edited by Andy Egan and B. Weatherson. Oxford University Press.

———. 2014. *Context.* Oxford University Press.

———. 1975. “Indicative Conditionals.” *Philosophia* 5 (3): 269–286.

———. 1984. *Inquiry.* Cambridge University Press.

———. 1998. “On the Representation of Context.” *Journal of Logic, Language and Information* 7 (1): 3–19.

———. 1973. “Presuppositions.” *Journal of Philosophical Logic* 2 (4): 447–457.

Stalnaker, Robert C. 1981. “A Defense of Conditional Excluded Middle.” In *Ifs,* edited by William Harper, Robert C. Stalnaker, and Glenn Pearce, 87–104. Reidel.

———. 1999. “Assertion.” In *Context and Content: Essays on Intentionality in Speech and Thought,* 78–95. Oxford University Press UK.

Stampe, Dennis W. 1986. “Defining Desire.” In *The Ways of Desire,* edited by J. Marks.

Stanley, Jason. 2002. “Nominal Restriction.” In *Logical Form and Language,* edited by Georg Peter and Gerhard Preyer, 365–390. Oxford University Press.

Stanley, Jason, and Zoltán Gendler Szabó. 2000. “On Quantifier Domain Restriction.” *Mind and Language* 15 (2-3): 219–61.

Starr, William B. 2016. “Expressing Permission.” *Semantics and Linguistic Theory* 26:325–349.

———. 2014. “Indicative Conditionals, Strictly.”

Stojnić, Una. 2019. “Content in a Dynamic Context.” *Noûs* 53 (2): 394–432.

———. 2021. *Context and Coherence: The Logic and Grammar of Prominence.* Oxford, UK: Oxford University Press.

Stojnić, Una. 2017. "One's Modus Ponens: Modality, Coherence and Logic." *Philosophy and Phenomenological Research* 95 (1): 167–214.

Stone, Matthew. 1999. "Reference to Possible Worlds." *Technical Report 49, Rutgers University Center for Cognitive Science.*

Strawson, P. F. 1950. "On Referring." *Mind* 59 (235): 320–344.

Tenenbaum, Sergio. 2013. "Guise of the Good." In *The International Encyclopedia of Ethics,* edited by Hugh LaFollette. Wiley-Blackwell.

Travis, Charles. 1994. "On Constraints of Generality." *Proceedings of the Aristotelian Society* 94 (1): 165–188.

Van Ditmarsch, Hans, Wiebe van Der Hoek, and Barteld Kooi. 2008. *Dynamic Epistemic Logic.* Springer.

Velleman, J. David. 2014. "The Guise of the Good." In *The Possibility of Practical Reason.* Michigan Publishing, University of Michigan Library.

Villalta, Elisabeth. 2008. "Mood and Gradability: An Investigation of the Subjunctive Mood in Spanish." *Linguistics and Philosophy* 31 (4): 467–522.

Von Fintel, Kai. 2011. "Conditionals." In *Semantics: an International Handbook of Natural Language Meaning,* edited by Klaus von Heusinger, Claudia Maienborn, and Paul Portner, 2:1515–1537. Berlin : De Gruyter Mouton.

———. 2007. "If: The biggest little word. Slides from talk at Georgetown University Roundtable."

———. 1999. "NPI Licensing, Strawson Entailment, and Context Dependency." *Journal of Semantics* 16 (2): 97–148.

———. 2012. "The best we can (expect to) get? Challenges to the classic semantics for deontic modals." *Paper presented in a session on Deontic Modals at the Central APA, February 17, 2012.*

———. 1997. "The Presupposition of Subjunctive Conditionals."

Von Fintel, Kai, and Anthony Gillies. 2011. "'Might' Made Right." In *Epistemic Modality,* edited by Andy Egan and Brian Weatherson, 108–130. Oxford University Press.

Von Fintel, Kai, and Anthony S. Gillies. 2007. "An Opinionated Guide to Epistemic Modality." In *Oxford Studies in Epistemology: Volume 2,* edited by Tamar Gendler and John Hawthorne, 32–62. Oxford University Press.

Von Fintel, Kai, and Irene Heim. 2011. *Intensional Semantics.* Manuscript.

Von Wright, G. H. 1968. *An Essay in Deontic Logic and the General Theory of Action.* Amsterdam: North-Holland Pub. Co.

Wedgwood, Ralph. 2006. "The Meaning of 'Ought'." In *Oxford Studies in Metaethics: Volume 1,* edited by Russ Shafer-Landau, 127–160. Clarendon Press.

Wheeler, Gregory. 2020. "Bounded Rationality." In *The Stanford Encyclopedia of Philosophy,* Fall 2020, edited by Edward N. Zalta. Metaphysics Research Lab, Stanford University.

Willer, Malte. 2018. "Simplifying with Free Choice." *Topoi* 37 (3): 379–392.

———. 2021. "Two Puzzles About Ability Can." *Linguistics and Philosophy* 44 (3): 551–586.

Yablo, Stephen. 2014. *Aboutness.* Princeton University Press.

Yalcin, Seth. 2012. "A Counterexample to Modus Tollens." *Journal of Philosophical Logic* 41 (6): 1001–1024.

Zimmermann, Thomas Ede. 2000. "Free Choice Disjunction and Epistemic Possibility." *Natural Language Semantics* 8 (4): 255–290.

———. 2006. "Monotonicity in Opaque Verbs." *Linguistics and Philosophy* 29 (6): 715–761.

Zvolenszky, Zsófia. 2002. "Is a Possible-Worlds Semantics of Modality Possible? A Problem for Kratzer's Semantics." *Proceedings of SALT* 12:339–358.

Zwarts, Frans. 1995. "Nonveridical contexts." *Linguistic Analysis,* 286–312.

Appendix A: Appendix to Chapter 3

A.1 Syntax and Semantics

Definition (Language)**.** From a countable set of atomic sentences At, our language $\mathcal{L}$ is built from the following grammar (where $p \in \mathsf{At}$ and $\phi, \psi \in \mathcal{L}$):

$$p \mid \neg\phi \mid \phi \wedge \psi \mid \phi \vee \psi \mid \phi \rightarrow \psi \mid \Box\phi \mid \Diamond\phi$$

Definition (Models)**.** A model $\mathcal{M}$ is a triple $\langle W, R, V \rangle$ such that:

- W is a set of worlds.
- R is a function from worlds to sets of relevant worlds $R : W \mapsto \wp(W)$.
- V is a valuation function from atomic sentences to sets of worlds $V : \mathsf{At} \mapsto \wp(W)$

Definition (Bilateral Inquisitive Propositions)**.** Given a model $\mathcal{M} = \langle W_{\mathcal{M}}, R_{\mathcal{M}}, V_{\mathcal{M}} \rangle$. Let $\mathcal{P}$ be the set of bilateral inquisitive propositions over $W_{\mathcal{M}}$ defined as follows. $P = \langle P^+, P^- \rangle \in \mathcal{P}_{\mathcal{M}}$ iff (where $P^\circ \in \{P^+, P^-\}$):

- $P^\circ \subseteq \wp(W)$ (set of sets of worlds)
- For any $s, t \subseteq W$, if $t \in P^\circ$ and $s \subseteq t$, then $s \in P^\circ$ (subset closed)
- $P^+ \cap P^- = \{\varnothing\}$ (no substantive overlap)

Definition (Downward Closure)**.** Given a set of sets of worlds $S \subseteq \wp(W)$:

$$\downarrow S := \{t \in \wp(W) \mid \exists s \in S : \ t \subseteq s\}$$

Definition (Informative Content)**.** Given a set of sets of worlds $S \subseteq \wp(W)$:

$$\mathsf{info}(S) := \bigcup S$$

Definition (Alternatives)**.** Given a bilateral inquisitive proposition $P = \langle P^+, P^- \rangle$:

$$\mathsf{alt}^+(P) = \{s \in P^+ \mid \forall t \in P^+, \text{ if } s \subseteq t, \text{ then } s = t\}$$
$$\mathsf{alt}^-(P) = \{s \in P^- \mid \forall t \in P^-, \text{ if } s \subseteq t, \text{ then } s = t\}$$

Definition (Semantics)**.** Given a model $\mathcal{M}$, the assignment of bilateral inquisitive propositions $[\![\phi]\!]_{\mathcal{M}} = \langle [\![\phi]\!]^+_{\mathcal{M}}, [\![\phi]\!]^-_{\mathcal{M}} \rangle$ to $\phi \in \mathcal{L}$ goes as follows (with reference to the model suppressed when uninteresting):

$$
\begin{aligned}
[\![p]\!]^+ &= \downarrow \{V(p)\} \\
[\![p]\!]^- &= \downarrow \{W \backslash V(p)\} \\
[\![\neg\phi]\!]^+ &= [\![\neg\phi]\!]^- \\
[\![\neg\phi]\!]^- &= [\![\neg\phi]\!]^+ \\
[\![\phi \vee \psi]\!]^+ &= [\![\phi]\!]^+ \cup [\![\psi]\!]^+ \\
[\![\phi \vee \psi]\!]^- &= [\![\phi]\!]^- \cap [\![\psi]\!]^- \\
[\![\Box\phi]\!]^+ &= \downarrow \{\{w \in W \mid R(w) \neq \varnothing \text{ and } \mathsf{alt}^+([\![\phi]\!]) \text{ m-covers } R(w)\}\} \\
[\![\Box\phi]\!]^- &= \downarrow \{\{w \in W \mid \exists R' \subseteq R(w) : \ R' \neq \varnothing \text{ and } \mathsf{alt}^-([\![\phi]\!]) \text{ m-covers } R'\}\}
\end{aligned}
$$

Let:

$$[\![\phi \wedge \psi]\!] = [\![\neg(\neg\phi \vee \neg\psi)]\!]$$
$$[\![\Diamond\phi]\!] = [\![\neg\Box\neg\phi]\!]$$

The semantics of conditionals will depend on the notion of a model update:[1]

Definition (Accessibility Update)**.** Given a model $\mathcal{M} = \langle W_{\mathcal{M}}, R_{\mathcal{M}}, V_{\mathcal{M}} \rangle$, and a sentence ϕ, we define the update of $R_{\mathcal{M}}$ with ϕ ($R^{\phi}_{\mathcal{M}}$) as follows:

$$R^{\phi}_{\mathcal{M}} := \{\langle w, S \rangle \mid w \in W_{\mathcal{M}} \text{ and } S = R_{\mathcal{M}}(w) \cap \mathsf{info}([\![\phi]\!]^{+}_{\mathcal{M}})\}$$

So $R^{\phi}_{\mathcal{M}}$ applied to w delivers $R(w) \cap \mathsf{info}([\![\phi]\!]^{+}_{\mathcal{M}})$. With this notion in hand, we define model update as follows:

Definition (Model Update)**.** Given a model $\mathcal{M} = \langle W_{\mathcal{M}}, R_{\mathcal{M}}, V_{\mathcal{M}} \rangle$, and a sentence ϕ, we define the update of $\mathcal{M}$ with ϕ, ($\mathcal{M}^{\phi}$) as follows:

$$\mathcal{M}^{\phi} := \langle W_{\mathcal{M}}, R^{\phi}_{\mathcal{M}}, V_{\mathcal{M}} \rangle$$

Fact 1 (Adequacy of Model Update)*.* Let $\mathcal{M}$ be a model, and $[\![\phi]\!]_{\mathcal{M}}$ a bilateral inquisitive proposition. Then, $\mathcal{M}^{\phi}$ is a model.

Proof. Since $[\![\phi]\!]_{\mathcal{M}}$ a bilateral inquisitive proposition, $\mathsf{info}([\![\phi]\!]^{+}_{\mathcal{M}}) \subseteq W_{\mathcal{M}}$. Thus, for each $w \in W_{\mathcal{M}}$, $R_{\mathcal{M}}(w) \cap \mathsf{info}([\![\phi]\!]^{+}_{\mathcal{M}}) \in \wp(W_{\mathcal{M}})$. So $R^{\phi}_{\mathcal{M}}$ is an accessibility relation over $W_{\mathcal{M}}$, so $\mathcal{M}^{\phi} = \langle W_{\mathcal{M}}, R^{\phi}_{\mathcal{M}}, V_{\mathcal{M}} \rangle$ is a model. □

Finally, we can give a semantics for our restrictor conditional as follows:

$$[\![\phi \to \psi]\!]_{\mathcal{M}} = [\![\psi]\!]_{\mathcal{M}^{\phi}}$$

A.2 Important Characteristics

Note that the present semantics retains duality for necessity and possibility modals (by definition of $\Diamond$), and in contrast to standard inquisitive semantics (see, e.g., Ciardelli,

1. See van Ditmarsch, van Der Hoek, and Kooi 2008 and other research in dynamic epistemic logic for similar operators.

Groenendijk, and Roelofsen 2018) it also generates the equivalence of $[\![\phi]\!]$ and $[\![\neg\neg\phi]\!]$:

Fact 2 (Double Negation). For any $\phi \in \mathcal{L}$ and model $\mathcal{M}$:

$$[\![\phi]\!]^+ = [\![\neg\phi]\!]^- = [\![\neg\neg\phi]\!]^+$$
$$[\![\phi]\!]^- = [\![\neg\phi]\!]^+ = [\![\neg\neg\phi]\!]^-$$

It is simple to prove that the semantic clauses given above do assign bilateral inquisitive propositions for every sentence ϕ of our language:

Fact 3 (Adequacy). For any sentence $\phi \in \mathcal{L}$, if $\mathcal{M}$ is a model, $[\![\phi]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition.

Adequacy:

Proof. Atoms. Suppose $p \in \mathsf{At}$ and let $\mathcal{M}$ be a model. Clearly, $\downarrow \{V(p)\}$ and $\downarrow \{W\backslash V(p)\}$ are subset-closed sets of sets of worlds, and $\varnothing \in\downarrow \{V(p)\}\cap \downarrow \{W\backslash V(p)\}$. Suppose there is some non-empty set $s \in\downarrow \{V(p)\}\cap \downarrow \{W\backslash V(p)\}$. Thus, $s \subseteq V(p)$ and $s \subseteq W\backslash V(p)$. Since s is non-empty, there is an $x \in s$ such that $x \in V(p)$ and $x \in W\backslash V(p)$, so $s \notin V(p)$, a contradiction. Thus, $\downarrow \{V(p)\}\cap \downarrow \{W\backslash V(p)\} = \{\varnothing\}$. Thus, $[\![p]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition.

Negation. Suppose the property holds for ϕ, and let $\mathcal{M}$ be a model.. Then $[\![\phi]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition. Clearly, $[\![\neg\phi]\!]^+ = [\![\phi]\!]^-$ and $[\![\neg\phi]\!]^- = [\![\phi]\!]^+$ are therefore a subset-closed set of sets of worlds. And clearly, $[\![\neg\phi]\!]^+ \cap [\![\neg\phi]\!]^- = [\![\phi]\!]^+ \cap [\![\phi]\!]^- = \{\varnothing\}$. So $[\![\neg\phi]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition.

Disjunctions. Suppose the property holds for ϕ, ψ, and let $\mathcal{M}$ be a model. Clearly, $[\![\phi]\!]^+ \cup [\![\psi]\!]^+$ and $[\![\phi]\!]^- \cap [\![\psi]\!]^-$ are subset-closed sets of sets of worlds, and therefore, $\varnothing \in ([\![\phi]\!]^+ \cup [\![\psi]\!]^+) \cap ([\![\phi]\!]^- \cap [\![\psi]\!]^-)$. Suppose there is a $w \in W$ such that $\{w\} \in [\![\phi \vee$

$\psi]\!]^+ \cap [\![\phi \vee \psi]\!]^-$. Then $\{w\} \in [\![\phi]\!]^-$ and $\{w\} \in [\![\psi]\!]^-$. Furthermore, either $\{w\} \in [\![\phi]\!]^+$ or $w \in [\![\psi]\!]^+$. If the former, then $[\![\phi]\!]$ is not a bilateral inquisitive proposition. If the latter, then $[\![\psi]\!]$ is not. Either way, we reach a contradiction. So there is no $w \in W$ such that $\{w\} \in [\![\phi \vee \psi]\!]^+ \cap [\![\phi \vee \psi]\!]^-$. Thus, $[\![\phi \vee \psi]\!]^+ \cap [\![\phi \vee \psi]\!]^- = \{\varnothing\}$, so $[\![\phi \vee \psi]\!]$ is a bilateral inquisitive proposition.

Modals. Suppose the property holds for ϕ and let $\mathcal{M}$ be a model. Then $[\![\phi]\!]_{\mathcal{M}}$ is a bilateral proposition. Clearly the following two sets are subset-closed sets of sets of worlds:

$$[\![\Box\phi]\!]^+ = \downarrow \{\{w \in W \mid R(w) \neq \varnothing \text{ and } \mathsf{alt}^+([\![\phi]\!]) \text{ m-cover } R(w)\}\}$$
$$[\![\Box\phi]\!]^- = \downarrow \{\{w \in W \mid \exists R' \subseteq R(w) : \; R' \neq \varnothing \text{ and } \mathsf{alt}^-([\![\phi]\!]) \text{ m-cover } R'\}\}$$

Since they are subset-closed, $\varnothing \in [\![\Box\phi]\!]^+ \cap [\![\Box\phi]\!]^-$. Suppose there is a non-empty set $s \in [\![\Box\phi]\!]^+ \cap [\![\Box\phi]\!]^-$. So for some $w \in s$:

$$R(w) \neq \varnothing \text{ and } \mathsf{alt}^+([\![\phi]\!]) \text{ m-cover } R(w)$$
$$\exists R' \subseteq R(w) : \; R' \neq \varnothing \text{ and } \mathsf{alt}^-([\![\phi]\!]) \text{ m-cover } R'$$

Let R' be a non-empty subset of $R(w)$ such that $\mathsf{alt}^-([\![\phi]\!])$ m-covers R'. Then there is a $v \in R'$ such that for some $a \in \mathsf{alt}^-([\![\phi]\!])$, $\{v\} \subseteq a$. Since $\mathsf{alt}^+([\![\phi]\!])$ m-cover $R(w)$, there is a $b \in \mathsf{alt}^+([\![\phi]\!])$ such that $\{v\} \subseteq b$. Since $[\![\phi]\!]$ is a bilateral inquisitive proposition, it is closed under the subset relation, so $\{v\} \in [\![\phi]\!]^-$ and $\{v\} \in [\![\phi]\!]^+$. But then $[\![\phi]\!]^+ \cap [\![\phi]\!]^- \neq \varnothing$, a contradiction. Thus, there is no non-empty set $s \in [\![\Box\phi]\!]^+ \cap [\![\Box\phi]\!]^-$. Thus, $[\![\Box\phi]\!]^+ \cap [\![\Box\phi]\!]^- = \{\varnothing\}$, so $[\![\Box\phi]\!]$ is a bilateral inquisitive proposition.

Conditionals. Suppose the property holds for ϕ, ψ. Let $\mathcal{M}$ be a model. Then $[\![\phi]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition. Thus, by Fact (1), $\mathcal{M}^{\phi}$ is a model. Since the property holds of ψ and $\mathcal{M}^{\phi}$ is a model, $[\![\psi]\!]_{\mathcal{M}^{\phi}} = [\![\phi \to \psi]\!]_{\mathcal{M}}$ is a bilateral inquisitive proposition. □

Fact 4 (Compactness of Alternatives)*.* For every sentence ϕ of $\mathcal{L}$, if $\mathcal{M}$ is a model:

$$\textsf{alt}^+(\llbracket\phi\rrbracket) \text{ is finite, and } \llbracket\phi\rrbracket^+ =\downarrow \textsf{alt}^+(\llbracket\phi\rrbracket)$$

$$\textsf{alt}^-(\llbracket\phi\rrbracket) \text{ is finite, and } \llbracket\phi\rrbracket^- =\downarrow \textsf{alt}^-(\llbracket\phi\rrbracket)$$

Proof. By definition, $\llbracket\phi\rrbracket^+ \supseteq\downarrow \textsf{alt}^+(\llbracket\phi\rrbracket)$ and $\llbracket\phi\rrbracket^- \supseteq\downarrow \textsf{alt}^-(\llbracket\phi\rrbracket)$. So we prove the simpler:

$$\textsf{alt}^+(\llbracket\phi\rrbracket) \text{ is finite, and } \llbracket\phi\rrbracket^+ \subseteq\downarrow \textsf{alt}^+(\llbracket\phi\rrbracket)$$

$$\textsf{alt}^-(\llbracket\phi\rrbracket) \text{ is finite, and } \llbracket\phi\rrbracket^- \subseteq\downarrow \textsf{alt}^-(\llbracket\phi\rrbracket)$$

By induction. For atomic sentences, modals, and negations, the proof is trivial. I include here only the case of conditionals and disjunctions.

Conditionals. Suppose the property holds for ϕ, ψ, and let $\mathcal{M}$ be a model. By Fact (3), $\llbracket\phi\rrbracket_{\mathcal{M}}$ is a bilateral inquisitive proposition. Thus, by Fact (1), $\mathcal{M}^{\phi}$ is a model. Thus, since the property holds of ψ and $\mathcal{M}^{\phi}$ is a model:

$$\textsf{alt}^+(\llbracket\psi\rrbracket_{\mathcal{M}^\phi}) \text{ is finite, and } \llbracket\psi\rrbracket^+_{\mathcal{M}^\phi} \subseteq\downarrow \textsf{alt}^+(\llbracket\psi\rrbracket_{\mathcal{M}^\phi})$$

$$\textsf{alt}^-(\llbracket\psi\rrbracket_{\mathcal{M}^\phi}) \text{ is finite, and } \llbracket\psi\rrbracket^+_{\mathcal{M}^\phi} \subseteq\downarrow \textsf{alt}^-(\llbracket\psi\rrbracket_{\mathcal{M}^\phi})$$

By our semantics, $\llbracket\phi \to \psi\rrbracket^+_{\mathcal{M}} = \llbracket\psi\rrbracket^+_{\mathcal{M}^\phi}$ and $\llbracket\phi \to \psi\rrbracket^-_{\mathcal{M}} = \llbracket\psi\rrbracket^-_{\mathcal{M}^\phi}$. Thus, substituting equivalents:

$$\textsf{alt}^+(\llbracket\phi \to \psi\rrbracket_{\mathcal{M}}) \text{ is finite, and } \llbracket\phi \to \psi\rrbracket^+_{\mathcal{M}} \subseteq\downarrow \textsf{alt}^+(\llbracket\phi \to \psi\rrbracket_{\mathcal{M}})$$

$$\textsf{alt}^-(\llbracket\phi \to \psi\rrbracket_{\mathcal{M}}) \text{ is finite, and } \llbracket\phi \to \psi\rrbracket^+_{\mathcal{M}} \subseteq\downarrow \textsf{alt}^-(\llbracket\phi \to \psi\rrbracket_{\mathcal{M}})$$

Disjunctions: Positive. Suppose the property holds of ϕ, ψ, and let $\mathcal{M}$ be a model. Let $s \in \textsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$. Then either (a) $s \in \llbracket\phi\rrbracket^+$ or (b) $s \in \llbracket\psi\rrbracket^+$. If (a), then if $s \notin \textsf{alt}^+(\llbracket\phi\rrbracket)$, there is a $t \supset s$ such that $t \in \llbracket\phi\rrbracket^+$. But then $t \in \llbracket\phi \vee \psi\rrbracket^+$, so $s \notin \textsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$, a contradiction.

So if (a), $s \in \mathsf{alt}^+(\llbracket\phi\rrbracket)$, so $s \in \mathsf{alt}^+(\llbracket\phi\rrbracket) \cup \mathsf{alt}^+(\llbracket\psi\rrbracket)$. Parallel reasoning shows that if (b), $s \in \mathsf{alt}^+(\llbracket\psi\rrbracket)$, so $s \in \mathsf{alt}^+(\llbracket\phi\rrbracket) \cup \mathsf{alt}^+(\llbracket\psi\rrbracket)$. Thus, $\mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket) \subseteq \mathsf{alt}^+(\llbracket\phi\rrbracket) \cup \mathsf{alt}^+(\llbracket\psi\rrbracket)$. Since $\mathsf{alt}^+(\llbracket\phi\rrbracket) \cup \mathsf{alt}^+(\llbracket\psi\rrbracket)$ is finite, $\mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$ is finite.

Let $s \in \llbracket\phi \vee \psi\rrbracket^+$, and suppose $s \notin\downarrow \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$. Thus, there is no $t \in \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$ such that $s \subseteq t$. Thus, there is an infinite chain $C \subseteq \llbracket\phi \vee \psi\rrbracket^+$ of elements $\{s_1, s_2, ...\}$ such that $s \subset s_1 \subset s_2, ...$

Clearly, either (a) $s_1 \in \llbracket\phi\rrbracket^+$ or (b) $s_1 \in \llbracket\psi\rrbracket^+$. Suppose (a). then since the property holds of ϕ, there is an $a \in \mathsf{alt}^+(\llbracket\phi\rrbracket)$ such that $s_1 \subseteq a$. If $a \in \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$, then $s \subseteq a$, so $s \in\downarrow \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$, a contradiction. If $a \notin \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$, then there must be a $b \in \llbracket\psi\rrbracket^+$ such that $a \subset b$. Since the property holds for ψ, there is a $c \in \mathsf{alt}^+(\llbracket\psi\rrbracket)$ such that $b \subset c$. Clearly, $c \in \llbracket\phi \vee \psi\rrbracket^+$. Suppose for reductio that $c \notin \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$. Then there is a $d \in \llbracket\phi \vee \psi\rrbracket^+$ such that $c \subset d$. If $d \in \llbracket\phi\rrbracket^+$, then since $a \subset d$, $a \notin \mathsf{alt}^+(\llbracket\phi\rrbracket)$, a contradiction. If $d \in \llbracket\psi\rrbracket^+$, then since $c \subset d$, $c \notin \mathsf{alt}^+(\llbracket\psi\rrbracket)$, a contradiction. Thus, there can be no such d, and thus, $c \in \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$. Thus, since $s \subset s_1 \subseteq a \subset b \subset c$, $s \in\downarrow \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$, a contradiction.

Parallel reasoning shows that assuming (b) leads to a contradiction. Thus, we reject our supposition. It follows that:

$$\llbracket\phi \vee \psi\rrbracket^+ \subseteq\downarrow \mathsf{alt}^+(\llbracket\phi \vee \psi\rrbracket)$$

Disjunctions: Negative. Assume the property holds for ϕ, ψ. Allow me the following definitions:

$$X \sqcap Y := \{x \cap y \mid x \in X, y \in Y\}$$

$$max(S) := \{s \in S \mid \neg\exists t \in S : \ s \subset t\}$$

I will show that:

$$\mathsf{alt}^-(\llbracket\phi \vee \psi\rrbracket) = max(\mathsf{alt}^-(\llbracket\phi\rrbracket) \sqcap \mathsf{alt}^-(\llbracket\psi\rrbracket))$$

Allow the following abbreviations for concision:

$$A := \mathsf{alt}^-(\llbracket \phi \vee \psi \rrbracket) \qquad\qquad B := \mathsf{alt}^-(\llbracket \phi \rrbracket) \sqcap \mathsf{alt}^-(\llbracket \psi \rrbracket)$$

If $s \in A$, then $s \in \llbracket \phi \rrbracket^-$ and $s \in \llbracket \psi \rrbracket^-$. Since the property holds of ϕ, ψ, there is an $a \in \mathsf{alt}^-(\llbracket \phi \rrbracket)$ and a $b \in \mathsf{alt}^-(\llbracket \psi \rrbracket)$ such that $s \subseteq a$ and $s \subseteq b$. Thus, $s \subseteq a \cap b$. By downward closure, $a \cap b \in \llbracket \phi \vee \psi \rrbracket^-$. Thus, if $s \subset a$, $s \notin A$, a contradiction. So $s = a$. Thus, $s \in B$. Suppose for reductio that there is a $t \in B$ such that $s \subset t$. Then there is a $c \in \mathsf{alt}^-(\llbracket \phi \rrbracket)$ and a $d \in \mathsf{alt}^-(\llbracket \psi \rrbracket)$ such that $t = c \cap d$. By subset-closure, $c \cap d \in \llbracket \phi \rrbracket^-$ and $c \cap d \in \llbracket \psi \rrbracket^-$, so $c \cap d \in \llbracket \phi \vee \psi \rrbracket^-$. But, since $s \subset c \cap d$, it follows that $s \notin A$, a contradiction. Thus, there is no $t \in B$ such that $s \subset t$. So, $s \in max(B)$. Thus, $A \subseteq max(B)$.

Suppose $s \in max(B)$. Then $s \in B$, so there is an $a \in \mathsf{alt}^-(\llbracket \phi \rrbracket)$ and a $b \in \mathsf{alt}^-(\llbracket \psi \rrbracket)$ such that $s = a \cap b$. By subset-closure, $a \cap b \in \llbracket \phi \rrbracket^-$ and $a \cap b \in \llbracket \psi \rrbracket^-$, so $a \cap b \in \llbracket \phi \vee \psi \rrbracket^-$. Suppose for reductio that there is a $t \in \llbracket \phi \vee \psi \rrbracket^-$ such that $a \cap b \subset t$. Then $t \in \llbracket \phi \rrbracket^-$ and $t \in \llbracket \psi \rrbracket^-$. Since the property holds of ϕ, ψ, there is a $c \in \mathsf{alt}^-(\llbracket \phi \rrbracket)$ and a $d \in \mathsf{alt}^-(\llbracket \psi \rrbracket)$ such that $t \subset c$ and $t \subset d$. So $t \subseteq c \cap d$. Clearly $c \cap d$ in B. Since $s \subset t \subseteq c \cap d$, $s \subset c \cap d$, so $s \notin max(B)$, a contradiction. Thus, there is no $t \in \llbracket \phi \vee \psi \rrbracket^-$ such that $a \cap b \subset t$. So $s \in A$.

Since the property holds of ϕ, ψ, $\mathsf{alt}^-(\llbracket \phi \rrbracket)$ and $\mathsf{alt}^-(\llbracket \psi \rrbracket)$ are finite. Thus, B is finite. Since $max(B)$ is a subset of B, it is finite. Therefore, $\mathsf{alt}^-(\llbracket \phi \vee \psi \rrbracket)$ is finite.

Finally, let $s \in \llbracket \phi \vee \psi \rrbracket^-$. Then $s \in \llbracket \phi \rrbracket^-$ and $s \in \llbracket \psi \rrbracket^-$. Since the property holds of ϕ, ψ, there is an $a \in \mathsf{alt}^-(\llbracket \phi \rrbracket)$ and a $b \in \mathsf{alt}^-(\llbracket \psi \rrbracket)$ such that $s \subseteq a \cap b$. Clearly, $a \cap b \in B$. Since B is finite, the subset of B of elements that are proper supersets of $a \cap b$ is finite. Therefore, there must be some element x of B such that $a \cap b \subseteq x$ and there is no element $y \in B$ such that $x \subset y$. Thus, there is some element x in $max(B)$ such that $ss \subseteq x$. Since $A = max(B)$, it follows that $s \in\downarrow (A)$. So $\llbracket \phi \vee \psi \rrbracket^- \subseteq\downarrow \mathsf{alt}^-(\llbracket \phi \vee \psi \rrbracket)$. □

A.3 Predictions

Definition (Truth and Falsity)**.** Given a model $\mathcal{M}$ and world $w \in W_{\mathcal{M}}$, we define $\mathcal{M}, w \vDash \phi$ (ϕ is true) and $\mathcal{M}, w \Dashv \phi$ (ϕ is false) as follows:

$$\mathcal{M}, w \vDash \phi := w \in [\![\phi]\!]^{+}_{\mathcal{M}}$$
$$\mathcal{M}, w \Dashv \phi := w \in [\![\phi]\!]^{-}_{\mathcal{M}}$$

Definition (Truth-Preservational Entailment)**.** Given a model $\mathcal{M}$, we say that $\phi_0, ..., \phi_n \vDash^{+}_{\mathcal{M}} \psi$ iff for every $w \in W_{\mathcal{M}}$:

$$\text{if } \mathcal{M}, w \vDash \phi_0 \text{ and } ..., \text{ and } \mathcal{M}, w \vDash \phi_n, \text{ then } \mathcal{M}, w \vDash \psi$$

Definition (Validity)**.** Given a class of models $\mathfrak{M}$, we say that $\phi_0, ..., \phi_n \vDash^{+}_{\mathfrak{M}} \psi$ (the argument $\phi_0, ..., \phi_n \therefore \psi$ is valid over $\mathfrak{M}$) iff for every $\mathcal{M} \in \mathfrak{M}$:

$$\phi_0, ..., \phi_n \vDash^{+}_{\mathcal{M}} \psi$$

Definition (Invalidity)**.** Given a class of models $\mathfrak{M}$, we say that $\phi \vDash^{+}_{\mathfrak{M}} \psi$ (the argument $\phi \therefore \psi$ is valid over $\mathfrak{M}$) iff for some $\mathcal{M} \in \mathfrak{M}$, there is a $w \in W_{\mathcal{M}}$ such that:

$$\mathcal{M}, w \vDash \phi \text{ and } \mathcal{M}, w \nvDash \psi$$

Definition (Strong Invalidity)**.** Given a class of models $\mathfrak{M}$, we say that $\phi \vDash^{+}_{\mathfrak{M}} \psi$ (the argument $\phi \therefore \psi$ is valid over $\mathfrak{M}$) iff there is no $\mathcal{M} \in \mathfrak{M}$, such that for some world $w \in W_{\mathcal{M}}$:

$$\mathcal{M}, w \vDash \phi \text{ and } \mathcal{M}, w \vDash \psi$$

Definition (Non-Hurford Disjunctions)**.** Given a model $\mathcal{M}$ we say that a disjunction

$\phi \vee \psi$ is non-Hurford iff:

$$[\![\phi]\!]^+_{\mathcal{M}} \nsubseteq [\![\psi]\!]^+_{\mathcal{M}}$$
$$[\![\psi]\!]^+_{\mathcal{M}} \nsubseteq [\![\phi]\!]^+_{\mathcal{M}}$$

Fact 5 (Independence Inferences: Meta-Language). Let $\mathfrak{M}$ be the set of models such that $\phi \vee \psi$ is non-Hurford. Then for each $\mathcal{M} \in \mathfrak{M}$:

$$\text{if } \mathcal{M}, w \vDash \Box(\phi \vee \psi) \text{ then } R(w) \cap (\mathsf{info}([\![\phi]\!]^+_{\mathcal{M}}) \backslash \mathsf{info}([\![\psi]\!]^+_{\mathcal{M}})) \neq \varnothing \text{ and}$$
$$R(w) \cap (\mathsf{info}([\![\psi]\!]^+_{\mathcal{M}}) \backslash \mathsf{info}([\![\phi]\!]^+_{\mathcal{M}})) \neq \varnothing$$

Proof. Suppose that in $\mathcal{M}$, $\phi \vee \psi$ is non-Hurford, and that for some $w \in W_{\mathcal{M}}$, $\mathcal{M}, w \vDash \Box(\phi \vee \psi)$.

Suppose for every $a \in \mathsf{alt}^+([\![\phi]\!]^+)$, there is a $b \in \mathsf{alt}^+([\![\psi]\!]^+)$ such that $a \subseteq b$. Then, by downward closure, $[\![\phi]\!]^+ \subseteq [\![\psi]\!]^+$, contradicting the fact that $\phi \vee \psi$ is non-Hurford. So there is an $a \in \mathsf{alt}^+([\![\phi]\!]^+)$ such that for no $a' \in \mathsf{alt}^+([\![\psi]\!]^+)$, $a \subseteq a'$. By compactness, $a \notin [\![\psi]\!]^+$. By the semantics of disjunction, $a \in [\![\phi \vee \psi]\!]^+$, and by compactness, $a \in \mathsf{alt}^+([\![\phi \vee \psi]\!])$.

Since $\mathcal{M}, w \vDash \Box(\phi \vee \psi)$, $\mathsf{alt}^+([\![\phi \vee \psi]\!]^+)$ m-covers $R(w)$. Thus there is a world $v \in R(w)$ such that $v \in a$ and for all $b \in \mathsf{alt}^+([\![\phi \vee \psi]\!]^+) - a$ $v \notin b$. Thus, $v \notin \bigcup(\mathsf{alt}^+([\![\phi \vee \psi]\!]) - a)$ Since $v \in a$, $v \in \mathsf{info}([\![\phi]\!]^+)$. Since $(\mathsf{alt}^+([\![\phi \vee \psi]\!]) - a) \supseteq \mathsf{alt}^+([\![\psi]\!])$, by compactness, $\bigcup(\mathsf{alt}^+([\![\phi \vee \psi]\!]) - a) \supseteq \mathsf{info}([\![\psi]\!]^+)$. Thus, it follows that $v \notin \mathsf{info}([\![\psi]\!]^+)$. Thus:

$$R(w) \cap (\mathsf{info}([\![\phi]\!]^+_{\mathcal{M}}) \backslash \mathsf{info}([\![\psi]\!]^+_{\mathcal{M}})) \neq \varnothing$$

Parallel reasoning with ψ traded for ϕ shows that:

$$R(w) \cap (\mathsf{info}([\![\psi]\!]^+_{\mathcal{M}}) \backslash \mathsf{info}([\![\phi]\!]^+_{\mathcal{M}})) \neq \varnothing$$

□

Fact 6 (Ross Inference Strongly Invalid). Let $\mathfrak{M}$ be the set of models such that $\phi \vee \psi$ is non-Hurford. Then there is no $\mathcal{M} \in \mathfrak{M}$ such that for some $w \in W_{\mathcal{M}}$,

$$\mathcal{M}, w \models \Box\phi \text{ and } \mathcal{M}, w \models \Box(\phi \vee \psi)$$

Proof. Take a model $\mathcal{M} \in \mathfrak{M}$ and world w such that $\mathcal{M}, w \models \Box\phi$. So $\mathsf{alt}^{+}([\![\phi]\!])$ m-covers $R(w)$. Since $\phi \vee \psi$ is non-Hurford, by compactness there is an alternative, let it be $a \in \mathsf{alt}^{+}([\![\psi]\!])$, such that for no alternative $b \in \mathsf{alt}^{+}([\![\phi]\!])$, $b \subseteq a$. Clearly, $a \in \mathsf{alt}^{+}([\![\phi \vee \psi]\!])$.

By the semantics of disjunction, $\mathsf{alt}^{+}([\![\phi]\!]) \subseteq [\![\phi \vee \psi]\!]^{+}$ and by compactness, for every $c \in \mathsf{alt}^{+}([\![\phi]\!])$, there is a $d \in \mathsf{alt}^{+}([\![\phi \vee \psi]\!])$ such that $c \subseteq d$. Clearly, for each such c, there is more specifically a $d \in (\mathsf{alt}^{+}([\![\phi \vee \psi]\!]) - a)$ such that $c \subseteq d$. Thus, $\mathsf{alt}^{+}([\![\phi \vee \psi]\!]) - a$ is a cover of $R(w)$, and since $\mathsf{alt}^{+}([\![\phi \vee \psi]\!]) - a \subset \mathsf{alt}^{+}([\![\phi \vee \psi]\!])$, $\mathsf{alt}^{+}([\![\phi \vee \psi]\!])$ is not an m-cover of $R(w)$. Thus, $\mathcal{M}, w \not\models \Box(\phi \vee \psi)$.

□

Fact 7 (Extended Ross Inference Strongly Invalid). Let $\mathfrak{M}$ be the set of models such that $\phi \vee \psi$ is non-Hurford. Then there is no $\mathcal{M} \in \mathfrak{M}$ such that for some $w \in W_{\mathcal{M}}$,

$$\mathcal{M}, w \models \Box\phi, \mathcal{M}, w \models \Diamond\psi, \text{ and } \mathcal{M}, w \models \Box(\phi \vee \psi)$$

Proof. By the previous fact, if $\mathcal{M}, w \models \Box\phi$, $\mathcal{M}, w \not\models \Box(\phi \vee \psi)$. □

The following properties concern the set of models $\mathfrak{N}$ such that for $p, q \in \mathsf{At}$, '$p \vee q$' is non-Hurford. They can be generalized in interesting ways, but I will not explore such generalizations further in this dissertation. Relative to a model $\mathcal{N} \in \mathfrak{N}$, since $p \vee q$ is non-Hurford:

$$\mathsf{alt}^+(\llbracket p \vee q \rrbracket) = \{V(p), V(q)\} \qquad \mathsf{alt}^-(\llbracket p \vee q \rrbracket) = \{W \backslash (V(p) \cup V(q))\}$$
$$\mathsf{alt}^+(\llbracket p \rrbracket) = \{V(p)\} \qquad \mathsf{alt}^-(\llbracket p \rrbracket) = \{W \backslash V(p)\}$$
$$\mathsf{alt}^+(\llbracket q \rrbracket) = \{V(q)\} \qquad \mathsf{alt}^-(\llbracket q \rrbracket) = \{W \backslash V(q)\}$$

Fact 8 (Independence Inferences: Object-Language). For each $\mathcal{N} \in \mathfrak{N}$:

$$\Box(p \vee q) \models^+_{\mathfrak{N}} \Diamond(p \wedge \neg q)$$
$$\Box(p \vee q) \models^+_{\mathfrak{N}} \Diamond(q \wedge \neg p)$$

Proof. Assume $\mathcal{N} \in \mathfrak{N}$, and let $w \in W_{\mathcal{N}}$ be a world such that $\mathcal{N}, w \models \Box(p \vee q)$. Since $p \vee q$ is non-Hurford, $\mathsf{alt}^+(p \vee q) = \{V(p), V(q)\}$ and $V(p) \neq V(q)$. Since $\mathcal{N}, w \models \Box(p \vee q)$, $\{V(p), V(q)\}$ m-covers $R(w)$. Thus, there is a $v \in R(w)$ such that $v \in V(p) \cap (W \backslash V(q))$. Now, $\mathsf{alt}^+(p \wedge \neg q) = \{V(p) \cap (W \backslash V(q))\}$, and clearly $\mathsf{alt}^+(p \wedge \neg q)$ m-covers $\{v\}$. Since $\mathsf{alt}^+(p \wedge \neg q) = \mathsf{alt}^-(\neg(p \wedge \neg q))$, $\mathsf{alt}^-(\neg(p \wedge \neg q))$ m-covers $\{v\}$. Thus $\Box\neg(p \wedge \neg q)$ is false at w, so $\neg\Box\neg(p \wedge \neg q)$ is true at w, i.e. $\Diamond(p \wedge \neg q)$ is true. Parallel reasoning shows the same for $\Diamond(q \wedge \neg p)$. □

Fact 9 (Free Choice).

$$\Diamond(p \vee q) \models^+_{\mathfrak{N}} \Diamond p$$
$$\Diamond(p \vee q) \models^+_{\mathfrak{N}} \Diamond q$$

Proof. Assume $\mathcal{N} \in \mathfrak{N}$, and let $w \in W_{\mathcal{N}}$ be a world such that $\mathcal{N}, w \models \Diamond(p \vee q)$. Thus, there is a non-empty $R' \subseteq R(w)$ such that $\mathsf{alt}^-(\llbracket \neg(p \vee q) \rrbracket) = \mathsf{alt}^+(\llbracket p \vee q \rrbracket) = \{V(p), V(q)\}$ m-covers R'. Thus there is a world $v \in R' \cap (V(p) \backslash V(q))$. Clearly, $\{v\} \subseteq V(p)$, so there is a non-empty subset of $R(w)$, namely $\{v\}$, that is minimally covered by $\{V(p)\}$. Since

$\mathsf{alt}^-(\llbracket\neg p\rrbracket) = \{V(p)\}$, so $\mathcal{N}, w \vDash \neg\Box\neg p$. In other words, $\mathcal{N}, w \vDash \Diamond p$. Parallel reasoning shows the same for $\Diamond q$. □

Fact 10 (Independence Conditionals). Let $\mathfrak{N}$ be the set of models such that for $p, q \in \mathsf{At}$, '$p \vee q$' is non-Hurford. Then:

$$\Box(p \vee q) \vDash^+_{\mathfrak{N}} \neg p \rightarrow \Box q$$

$$\Box(p \vee q) \vDash^+_{\mathfrak{N}} \neg q \rightarrow \Box p$$

Proof. Let $\mathcal{N} \in \mathfrak{N}$. Suppose $\mathcal{N}, w \vDash \Box(p \vee q)$. Then $\{V(p), V(q)\}$ m-covers $R(w)$, so $R(w) \subseteq (V(p) \cup V(q))$ and there is a $w_q \in R(w)\backslash V(p)$, and $R(w)\backslash V(p) \subseteq V(q)$.

Clearly, $\mathsf{info}(\llbracket\neg p\rrbracket^+) = \mathsf{info}(\llbracket p\rrbracket^-) = W\backslash V(p)$. Thus, $R^{\neg p}(w) = R(w) \cap W\backslash V(p) = R(w)\backslash V(p)$. Thus, $R^{\neg p}(w)$ is non-empty, and $\mathsf{alt}^+(\llbracket q\rrbracket)$ m-covers $R^{\neg p}(w)$. Thus, $\langle W_{\mathcal{N}}, R^{\neg p}_{\mathcal{N}}, V_{\mathcal{N}}\rangle, w \vDash \Box q$, so $\mathcal{N}^{\neg p}, w \vDash \Box q$. Thus, $\mathcal{N}, w \vDash \neg p \rightarrow \Box q$. Parallel reasoning shows the same for $\neg q \rightarrow \Box p$. □

Although I did not discuss this data in the main section, the present *bilateral* version of our minimal covering semantics retains some desirable results of the Kratzerian semantics for modals. First, negated necessity modals seem to distribute over disjunctions:

Fact 11 (Unnecessity Distribution).

$$\neg\Box(p \vee q) \vDash^+_{\mathfrak{N}} \neg\Box p$$

$$\neg\Box(p \vee q) \vDash^+_{\mathfrak{N}} \neg\Box q$$

Proof. Let $\mathcal{N} \in \mathfrak{N}$ and suppose $\mathcal{N}, w \vDash \neg\Box(p \vee q)$. Then there is a non-empty subset $R' \subseteq R(w)$ such that $\mathsf{alt}^-(\llbracket p \vee q\rrbracket) = \{W\backslash(V(p) \cup V(q))\}$ m-covers R'. Since $W\backslash(V(p) \cup V(q)) \subseteq W\backslash V(p)$, $\{W\backslash V(p)\}$ also m-covers R'. Thus, there is a non-empty subset of $R(w)$, namely, R', such that $\mathsf{alt}^-(\llbracket p\rrbracket) = \{W\backslash V(p)\}$ m-covers it. So $\mathcal{N}, w \vDash \neg\Box p$. Parallel reasoning shows the same for $\Box q$. □

Similarly, negated possibility modals seem to distribute over disjunctions:

Fact 12 (Impossibility Distribution).

$$\neg\Diamond(p \vee q) \models^{+}_{\mathfrak{N}} \neg\Diamond p$$

$$\neg\Diamond(p \vee q) \models^{+}_{\mathfrak{N}} \neg\Diamond q$$

Proof. Let $\mathcal{N} \in \mathfrak{N}$ and suppose $\mathcal{N}, w \models \neg\Diamond(p \vee q)$. Then $\mathcal{N}, w \models \Box\neg(p \vee q)$, so $R(w)$ is non-empty and $R(w)$ is m-covered by $\{W \backslash (V(p) \cup V(q))\}$. Since $W \backslash (V(p) \cup V(q)) \subseteq W \backslash V(p)$, $R(w)$ is m-covered by $\{W \backslash V(p)\} = \mathsf{alt}^{+}(\llbracket \neg p \rrbracket)$. Thus, $\mathcal{N}, w \models \Box\neg p$. Since $\neg\Diamond p$ abbreviates $\neg\neg\Box\neg p$, which is clearly equivalent to $\Box\neg p$, $\mathcal{N}, w \models \neg\Diamond p$. Parallel reasoning shows the same for $\neg\Diamond q$. □

Appendix B: Appendix to Chapter 5

In this appendix, I prove that the system of §5.2 guarantees that sentences of the form $\phi \to \psi$ presuppose $\Diamond\phi$.

Remark 1. Updates are Monotonic, i.e., if $c + \phi$ is defined, then $c + \phi \subseteq c$.

Proof. By induction on the structure of ϕ.

- If $\phi \in \mathsf{At}$ or ϕ is of the form $\Diamond\psi$ or $\Box\psi$. If $c + \phi$ is defined, then $c + \phi = c \cap [\![\phi]\!]_{\mathcal{M},c}$. Thus, $c + \phi \subseteq c$.

- If ϕ is of the form $\psi \to \chi$. Assume that the proposition holds for ψ and χ. If $c + \psi \to \chi$ is defined, so is $(c + \psi) + \chi$, and thus, so is $c + \psi$. Now, $c + \psi \to \chi = [c \backslash (c + \psi)] \cup [(c + \psi) + \chi]$. Since $c + \psi$ is defined, the left-side of this union is a subset of c. Since $(c + \psi) + \chi$ is defined, and by the inductive hypothesis, the proposition holds for any $c + \chi$ that are defined, $(c + \psi) + \chi$ is also a subset of c. Since the union of two subsets of c is also a subset of c, the proposition holds.

□

Remark 2. Given an arbitrary formula ϕ, model $\mathcal{M}$, and a context c in that model, we always have the following:

$$\text{(1) Either } [\![\Box\phi]\!]^{+}_{\mathcal{M},c} = W_{\mathcal{M}} \text{ or } [\![\Box\phi]\!]^{+}_{\mathcal{M},c} = \varnothing;$$
$$\text{and (2) Either } [\![\Diamond\phi]\!]^{+}_{\mathcal{M},c} = W_{\mathcal{M}} \text{ or } [\![\Diamond\phi]\!]^{+}_{\mathcal{M},c} = \varnothing$$

Proof. In our modal truth conditions (Def. 5.6.2), the comprehension predicate on the right of '|' does not mention the v variable on the left of '|'. Thus, for either clause, if the comprehension predicate is satisfied, the set is W. If it is not satisfied, the set is $\varnothing$. □

Remark 3. If $c + \phi$ accepts $\Diamond\psi$, then c accepts $\phi \to \Diamond\psi$.

Proof. Let $c + \phi$ be defined, and suppose $c + \phi$ accepts $\Diamond\psi$, in other words, that $(c + \phi) + \Diamond\psi = (c + \phi)$. Note that $(c + \phi) + \Diamond\psi$ must therefore be defined.

Suppose for contradiction that c does not accept $\phi \to \Diamond\psi$. In other words, $c \neq c + \phi \to \Diamond\psi$. Thus:

$$c \neq [c \backslash (c + \phi)] \cup [(c + \phi) + \Diamond\psi]$$

By Remark 1, $c + \phi \subseteq c$, and $(c + \phi) + \Diamond\psi \subseteq (c + \phi)$. Thus, if $c \neq [c \backslash (c + \phi)] \cup [(c + \phi) + \Diamond\psi]$, this can only be because:

$$(c + \phi) \neq (c + \phi) + \Diamond\psi$$

By the rule for unconditional assertion of modals,

$$(c + \phi) + \Diamond\psi = (c + \phi) \cap [\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi}$$

By Remark 2, either $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = W$ or $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = \varnothing$. If $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = W$, then we contradict ourselves:

$$(c + \phi) + \Diamond\psi = (c + \phi) \cap W = c + \phi$$

On the other hand, if $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = \varnothing$, then $(c + \phi) \cap [\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = \varnothing$. But then, by Def. 5.6.2 above, $(c + \phi) + \Diamond\psi$ is undefined, a contradiction.

Thus, c must accept $\phi \to \Diamond\psi$. □

Remark 4. If c accepts $\phi \to \Diamond\psi$, then c accepts $\Diamond(\phi \to \psi)$.

Proof. Suppose c accepts $\phi \to \Diamond\psi$; and suppose for contradiction that c does not accept $\Diamond(\phi \to \psi)$. In other words:

$$c \cap [\![\Diamond(\phi \to \psi)]\!]^+_{\mathcal{M},c} \neq c$$

Thus, $[\![\Diamond(\phi \rightarrow \psi)]\!]^+_{\mathcal{M},c} \neq W$, and by Remark 2, $[\![\Diamond(\phi \rightarrow \psi)]\!]^+_{\mathcal{M},c} = \varnothing$. By Def. 5.6.2, then, there is no $w \in \textsc{domid}(c, \phi \rightarrow \psi)$ such that $w \in \textsc{propid}(c, \phi \rightarrow \psi)^+$. By the Def.s of DOMID and PROPID, it follows that there is no $w \in \textsc{domid}(c + \phi, \psi)$ such that $w \in \textsc{propid}(c + \phi, \psi)$.

Now, recall our initial assumption that c accepts $\phi \rightarrow \Diamond\psi$. Thus:

$$\begin{aligned} c &= c + \phi \rightarrow \Diamond\psi \\ &= [c \backslash (c + \phi)] \cup [(c + \phi) + \Diamond\psi] \end{aligned}$$

It follows that $c + \phi = (c + \phi) + \Diamond\psi$. By Def. 5.6.2, then:

$$\begin{aligned} c + \phi &= (c + \phi) + \Diamond\psi \\ &= (c + \phi) \cap [\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} \end{aligned}$$

Since $c + \phi$ is defined, then Def. 5.6.2, it follows that $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} \neq \varnothing$. By Remark 2, then, $[\![\Diamond\psi]\!]^+_{\mathcal{M},c+\phi} = W$. By Def. 5.6.2, it follows that for some $v' \in \textsc{domid}(c + \phi, \psi) : v' \in \textsc{propid}(c + \phi, \psi)^+$, a contradiction.

Therefore, c must accept $\Diamond(\phi \rightarrow \psi)$. □

Remark 5. If $c + \phi$ accepts $\Diamond\psi$, then c accepts $\Diamond(\phi \rightarrow \psi)$.

Proof. Suppose $c + \phi$ accepts $\Diamond\psi$. Then by Remark 3, c accepts $\phi \rightarrow \Diamond\psi$. Then by Remark 4, c accepts $\Diamond(\phi \rightarrow \psi)$. □

Proposition 5. *If $c + \phi$ is defined, then c accepts $\Diamond\phi$.*

Proof. By induction on the structure of ϕ.

- When $\phi \in \mathsf{At}$ or is of the form $\Box\psi$ or $\Diamond\psi$. Suppose $c + \phi$ is defined. By Def. 5.6.2, $c \cap [\![\phi]\!]^+_{\mathcal{M},c} \neq \varnothing$. Let $w \in c$ and $w \in [\![\phi]\!]^+_{\mathcal{M},c}$. By the Def. of DOMID, $\textsc{domid}(c, \phi) = c$ and by the Def. of PROPID, $\textsc{propid}(c, \phi) = [\![\phi]\!]_{\mathcal{M},c}$. Thus, by substitution, $w \in$

$\textsc{domid}(c, \phi)$ and $w \in \textsc{propid}(c, \phi)^+$. Thus, by Def. 5.6.2, $[\![\Diamond\phi]\!]^+_{\mathcal{M},c} = W$. By Def. 5.6.2, then, $c + \Diamond\phi = c$. So c accepts $\Diamond\phi$.

- When ϕ is of the form $\psi \to \chi$. Suppose as the inductive hypothesis (IH) that for any c', c'', if $c' + \psi$ is defined, then c' accepts $\Diamond\psi$, and if $c'' + \chi$ is defined, then c'' accepts $\Diamond\chi$. Suppose that $c + \psi \to \chi$ is defined. Then by Def. 5.6.2, $c + \psi$ is defined and $(c + \psi) + \chi$ is defined. By IH, $c + \psi$ accepts $\Diamond\chi$. By Remark 5, c accepts $\Diamond(\psi \to \chi)$. So the proposition holds: if $c + \psi \to \chi$ is defined, then c accepts $\Diamond(\psi \to \chi)$.

$\square$

Proposition 6. *If $c + \phi \to \psi$ is defined, then c accepts $\Diamond\phi$.*

Proof. If $c + \phi \to \psi$ is defined, then by Def. 5.6.2, $c + \phi$ is defined. By Proposition 1, c accepts $\Diamond\phi$. $\square$