Axiomatization in the Meaning Sciences

Wesley H. Holliday* and Thomas F. Icard, III†

∗ Department of Philosophy and Group in Logic and the Methodology of Science, UC Berkeley

† Department of Philosophy and Symbolic Systems Program, Stanford University

Preprint of April 2017

Forthcoming in *The Science of Meaning*, eds. D. Ball and B. Rabern, Oxford University Press.

Abstract

While much of semantic theorizing is based on intuitions about logical phenomena associated with linguistic constructions—phenomena such as consistency and entailment—it is rare to see axiomatic treatments of linguistic fragments. Given a fragment interpreted in some class of formally specified models, it is often possible to ask for a characterization of the reasoning patterns validated by the class of models. Axiomatizations provide such a characterization, often in a perspicuous and efficient manner. In this paper, we highlight some of the benefits of providing axiomatizations for the purpose of semantic theorizing. We illustrate some of these benefits using three examples from the study of modality.

Keywords: Logic in Language, Axiomatization, Entailment, Modality

1 Introduction

One prominent way of approaching the study of meaning in natural language is to consider logical properties of and relations between sentences. For instance, if it is possible for one sentence to be true while another is false, this shows that the sentences have different meanings [Lewis, 1970, Cresswell, 1982]. More generally, a theory of meaning should at a minimum capture relations of entailment and non-entailment among the possible sentences in the language. This is particularly central for the tradition of model-theoretic semantics, whose goal is to interpret linguistic expressions in terms of independent mathematical structures, over which logical notions such as validity, satisfiability, and entailment can be rigorously defined (see any textbook on formal semantics, e.g., Chierchia and McConnell-Ginet 2001). A formal semantics for some fragment of language can then be tested against entailment intuitions of ordinary speakers, either those of the theorist if the judgments are sufficiently obvious, or, as is increasingly common, those of subjects in a controlled experiment. To the extent that these intuitions match the account of entailment defined by the interpretation, this weighs in favor of the account. Given this empirical interest in entailment, it is perhaps surprising that contemporary work in model-theoretic semantics rarely considers questions of *axiomatization*.

Axiomatizing a fragment of language over some class of models requires proving a *completeness theorem*, showing that some set of basic inference patterns suffices to derive all and only the entailments predicted by the interpretation. Axiomatization can thus be seen as a way of systematically and perspicuously revealing what the entailment predictions of a given formal semantics actually are. In this paper, we would like to argue that such an activity can indeed be valuable in the search for adequate accounts of natural language meaning. After a general discussion of the history and use of completeness theorems in the study of natural language (§2), we will give several examples (§3), each involving a different *modal* phenomenon in English—counterfactual conditionals, epistemic comparatives, and indicative conditionals—in order to draw some general morals about the beneficial role that axiomatization can play in the meaning sciences.

2 Completeness in Semantics

Early work in model-theoretic semantics grew out of the model-theoretic tradition in logic [Tarski, 1936], where questions of axiomatization were front and center. In part this was due to connections between axiomatization and other metalogical properties of interest, such as (semi)decidability. However, it was also taken to be of inherent interest to encapsulate potentially unwieldy notions of *consequence* in terms of a few, ideally simple principles. This led to axiomatic study not just of logical and mathematical systems, but also of empirical theories as diverse as thermodynamics and psychological theories of learning (see Suppes 1974 for a review of some of this work, including a short discussion of axiomatization in semantics).

Against this backdrop, in many of Montague's foundational papers on natural language semantics, questions of axiomatization are raised as a matter of course. For instance, in "Pragmatics and Intensional Logic," Montague notes that "it would be desirable . . . to find natural axiomatizations of" the intensional logics studied in the paper, citing completeness results on related logics in the dissertation of David Kaplan [Montague, 1970, p. 90]. The problem of axiomatizing Montague's full system was left as an open question, and the solution to the problem would later be the main contribution of Daniel Gallin's seminal dissertation on intensional higher-order logic for natural language semantics [Gallin, 1975]. At the same time, work on specific constructions, e.g., counterfactuals [Lewis, 1973a], would also be accompanied by completeness results when possible (see §3.3 for more on counterfactuals).

Despite this distinguished history of axiomatics within semantics, completeness theorems have been all but absent in work over the past decades. Exceptions include projects that are directly responding to these earlier proposals: e.g., there have been a number of axiomatizations of alternatives to Montague-Gallin intensional type theory by Fitting, Muskens, Zalta, and others (see Muskens 2007 for a review); and there is work on completeness theorems for fragments of counterfactual conditional logics that differ from the Lewis-Stalnaker style account (e.g., Galles and Pearl 1998, Briggs 2012, Halpern 2013). To be sure, questions of completeness can be raised any time one has specified a model-theoretic interpretation of a well-defined class of expressions. Thus, many semantic analyses of language fragments would in principle be amenable to axiomatic treatment.[1] Yet it is rare to see any discussion of axiomatics in contemporary linguistic work.

Recently, Moss [2010] has issued a forceful call encouraging interest in completeness theorems. He writes:

> One motivation for semantics found in textbooks is that it should be the study of *inference in language*: just as syntax has grammaticality judgments to account for, semantics has inference judgments. Now I happen to be mainly a logician, and this point resonates with me as a motivation for semantics. But the semantics literature, it almost never gives a full account of any inferences whatsoever. It is seriously concerned with truth conditions and figuring out how semantics should work in a general way. But it rarely goes back and figures out, for various fragments, what the overall complete stock of inferences should be. (pp. 84-85)

While semanticists do routinely provide well-defined model-theoretic interpretations of linguistic fragments, and then test judgments about some of the entailment claims predicted by these interpretations, no attempt

[1] Of course, as is well known, there is no guarantee that a semantically defined logic will be (recursively) axiomatizable. Our points in this paper are about the value of axiomatizations when they are available.

is made to characterize the full range of entailment predictions, except implicitly by the usual model-theoretic definition. This definition, which goes back to Tarski [1936] (and arguably back even further to Bolzano), says that a sentence φ is a consequence of a set Γ of premises iff every model (in some relevant class of models) making all the premises in Γ true also makes φ true. Some authors have suggested that this semantic notion of consequence renders deductive, or axiomatic, treatments of consequence "superfluous" to the study of linguistic meaning (Dowty et al. 1981, p. 53). Even Lewis, who proved several important completeness theorems himself, was unsure of their usefulness [Lewis, 1973a]:

> I am not sure how much completeness proofs really add to our understanding, but I here provide them for those readers who do find them helpful and for those—like myself—who find them interesting in their own right. Readers of other persuasions have no reason to read on. . . . (p. 118)

Exactly what advantages could be gained by explicit axiomatization of a consequence relation? While our focus here will be on one specific use—number (IV) below—we first list several additional potential benefits.

(I) In proposing a particular axiomatic system for a fragment of language, one might optimistically take the system to capture important patterns in how people in fact reason with language. There is much work in psychology of reasoning suggesting that people may reason according to psychologically basic deductive steps [Rips, 1994]. For instance, within linguistics, Dowty [1994] has claimed that deduction provides a more plausible account of how people make inferences in language:

> It is . . . a truism that humans do not carry anything analogous to infinite sets of possible worlds or situations around in their heads, so the study of deduction—inferential relations based on syntactic properties of some kind of "representations" of denotations—are potentially of relevance to the psychology of language and to computational processing of meaning in a way that model-theoretic semantics alone is not. (p. 114)

These claims about the psychological relevance, not to mention centrality, of deduction are of course highly controversial. At the same time, some approaches to the psychology of reasoning based on "mental models" have become so precise about algorithms defined over algebraically structured mental representations that the very distinction between applying proof rules and model-checking is not so clear. For instance, Koralus and Mascarenhas [2013] develop an algorithmic analysis of deductive reasoning based on mental models (the so called Erotetic Theory of Reasoning) and in fact establish a completeness result for a special case of their derivation system with respect to "classical" logical semantics.

As Dowty suggests in the quotation above, explicit deduction may also be relevant in computational applications, and indeed theorem provers play an important role in some of the most extensively developed approaches to automated natural language understanding [Blackburn and Bos, 2005]. In these applications, just as in the psychological applications, a completeness result would show that some class of reasoning patterns does in fact capture everything sanctioned by the model-theoretic analysis.

(II) Partly in connection with (I), it is often useful to study very small fragments of language, restricted in expressive power, even when we know these fragments sit inside of larger, well-understood systems. For example, Pratt-Hartmann and Moss have considered formal languages inspired by English that can be seen as fragments of first-order logic, as well as languages whose expressive power is incomparable with that of first-order logic (see Pratt-Hartmann and Moss 2009). Completeness theorems help identify how strong these fragments actually are in terms of their inferential capacity. Similarly, one can consider syntactically rich fragments of language, but for which the semantic interpretation is given in a more coarse-grained manner.

For instance, in the context of full higher-order logical syntax, one can restrict attention to distinctive reasoning patterns pertaining to identifiable semantic features such as monotonicity [van Benthem, 1986, Dowty, 1994]. Here, too, axiomatization can help isolate these reasoning patterns in a perspicuous way and show just how much can be derived from assumptions in such restricted settings [Icard and Moss, 2013, 2014]. In both of these cases, we already have well-understood axiom systems for the most general languages (those of first- and higher-order logic), and axiomatizing the weaker fragments allows direct comparison.

(III) Related to points (I) and (II), completeness theorems may help reveal information about the complexity of some fragment of language or of a specific construction or set of reasoning patterns. The primary aim of work by McAllester and Givan [1992] and Pratt-Hartmann and Moss [2009], for example, is to find logical fragments whose syntax more closely resembles that of natural language, but whose complexity is (in ideal cases, significantly) less than that of first-order logic. This may be particularly important if one has psychological or computational motivations in mind. In a more theoretical vein, the axiomatizations of intensional higher-order logic mentioned above effectively showed that these fragments, when interpreted in a certain way (with so-called general or Henkin models), would have no worse complexity than first-order logic, the set of validities being computably enumerable.

(IV) A fourth motivation for investigating completeness, arguably independent of (I)-(III), but perhaps closest to the actual practice of linguistic semantics, is to make clear and precise for the theorist what exactly the entailment predictions of a given model-theoretic account are. The point is that if among the primary semantic data are intuitions about logical properties such as entailment, then we may not fully understand the predictions of a semantic account until we have an intuitive, complete proof system. On the one hand, there may be predictions that were not initially obvious by simply looking at the model-theoretic interpretation, and completeness theorems can draw attention to them. This could either vindicate an account in cases where the prediction is distinctive but empirically accurate, or it could show the account is problematic in cases where the prediction is shown incorrect. (In one example we consider below, on epistemic comparatives, the problematic pattern emerges as the central axiom in a complete proof system.) On the other hand, proving completeness greatly reduces the chance that there are spurious entailment patterns that simply escaped the attention of the semanticist. In general, one may not be able to rule out the possibility that some combination of axioms and rules could lead to repugnant conclusions. But as long as the axioms and rules are themselves intuitively compelling, such doubt can at least be tempered. This fourth motivation is also seen in the broader theory of measurement, where, just as in semantics, a theory is intended to have empirical consequences (see, e.g., the discussion in Chapter 1 of Krantz et al. 1971). Axiomatizations can summarize those consequences in an intuitive, accessible manner. In some cases, the point of this exercise is to show simply that two semantic proposals give rise to the exact same entailment predictions. We will see several examples of these points in what follows.

3 Case Study: Modality

Modality is one of the most captivating topics in natural language semantics, in part because it concerns the distinctively human tendency to think and communicate about alternative ways things could have been or might be. This makes reasoning about modality particularly intriguing and subtle. In this section, we will present case studies of axiomatization involving three flavors of modality:

- counterfactuals (§3.3), e.g., *If Kangaroos had no tails, they would topple over.*

- epistemic comparatives (§3.4), e.g., *It's more likely that you'll be struck by lightning than win the lottery.*
- indicative conditionals and epistemic modals (§3.5), e.g., *If Miss Scarlet isn't the culprit, then Professor Plum is and might be in cahoots with Colonel Mustard.*

In each case, we will consider a formal semantics that has been proposed for such constructions and then present an axiomatization of the entailments predicted by that formal semantics for the relevant fragment of language. The point is not to argue in favor of or against any of these formal semantics, but rather to give clear examples of the methodology of axiomatization.

Considerations of entailment also figure prominently in semantic studies of other flavors of modality, such as deontic modality (see, e.g., Kolodny and MacFarlane 2010, Cariani 2013, 2017), where axiomatization can also be fruitfully applied (see, e.g., Van De Putte 2016). In the interest of space, however, we hope the bulleted examples above are enough to illustrate our methodological points in this paper.

3.1 Logical Preliminaries

For the purposes of what follows, it will be convenient to stipulate an abstract definition of a "logic" for a formal language. Given a formal language $\mathcal{L}$ extending the language of propositional logic, by a *logic for* $\mathcal{L}$ we mean a set L of formulas of $\mathcal{L}$ that meets at least the following conditions: L contains all formulas in the language $\mathcal{L}$ that have the form of *classical propositional tautologies*; L is closed under *modus ponens*, i.e., if $\varphi \in \mathsf{L}$ and $\varphi \to \psi \in \mathsf{L}$, then $\psi \in \mathsf{L}$; and L is closed under *replacement of equivalents*, i.e., if $\varphi \leftrightarrow \psi \in \mathsf{L}$ and $\alpha \in \mathsf{L}$, then $\alpha' \in \mathsf{L}$ where α' is the result of replacing some occurrences of φ in α by ψ. From Montague onward, formalizations of fragments of English have typically given rise to logics in this sense. One might be motivated to relax these conditions to accommodate non-classicality or hyperintensionality—and one could still seek completeness theorems in such a context—but all of our case studies in this paper will fit the classical template. By a *normal* logic we mean a logic that also satisfies the following further conditions: L is closed under *uniform substitution*, i.e., if $\varphi \in \mathsf{L}$ and ψ is obtained from φ by uniformly substituting formulas for proposition letters, then $\psi \in \mathsf{L}$; and if the language $\mathcal{L}$ contains a modal operator $\Box$, then L contains the formula $\Box(p \wedge q) \leftrightarrow (\Box p \wedge \Box q)$ and is such that if $\varphi \in \mathsf{L}$, then $\Box\varphi \in \mathsf{L}$, and if $\varphi \leftrightarrow \psi \in \mathsf{L}$, then $\Box\varphi \leftrightarrow \Box\psi \in \mathsf{L}$. Uniform substitution is often built in to the definition of a logic in the first place. Finally, while the conditions on $\Box$ considerably restrict the range of modals that could be formalized in this way—for example, the modal *probably* cannot be formalized as such a $\Box$, because the right-to-left direction of distribution over conjunction fails for *probably*—again our case studies will fit the normal template.

3.2 Over-generation and Under-generation

When it comes to comparing the entailment predictions of a formal semantics with intuitive judgments about entailments, there are two questions to consider:

- Over-generation question: does the formal semantics count as valid what are intuitively non-entailments?
- Under-generation question: does the formal semantics count as invalid what are intuitively entailments?

As for the over-generation question, failure on the part of the theorist to find examples provides some evidence that the semantics does not over-generate. However, much stronger evidence against over-generation in the relevant fragment of language would come from a complete axiomatization in that language that draws only upon intuitively valid patterns. Given such a complete axiomatization, the only way the semantics could over-generate entailments in the relevant fragment is if intuitively valid entailments could be strung together to arrive at a conclusion that intuitively does not follow from the original premises. If we understand the conclusion and judge that it can be false while all the premises are true, despite the string of intuitively valid entailments leading from premises to conclusion, we would have a paradox.[2] Of course, such entailments could be strung together to arrive a conclusion so complicated that no one can understand it, but this is a relatively benign form of over-generation common to all formal semantics.

Thus, a complete axiomatization using only intuitively valid patterns provides strong evidence against the kind of over-generation problems for which a formal semantics can be rejected in favor of another formal semantics. In addition, we may easily catch cases of over-generation by trying to find a complete axiomatization for a given semantics. We will give examples of these points in the case studies to follow.

As for the under-generation question, completeness theorems by themselves provide no guarantee that a semantics is not missing some intuitive entailment predictions. In some special cases, it may be possible to have assurance against under-generation by showing that the logic associated with one's semantics is such that any *stronger* logic would have to contain some intuitively invalid patterns. To take one famous example, classical propositional logic has the special property ("Post-completeness") that the only stronger normal propositional logic as in §3.1 is the *inconsistent* logic containing $p \wedge \neg p$. For another example from a philosophical rather than linguistic context, many philosophers have held that the modal logic **S5** captures the correct principles for reasoning about what *could have been the case* and what *must have been the case.* It can be proved that any normal logic extending **S5** contains a principle of a very specific form, which is not an intuitively valid principle for reasoning about what could have been the case (see Williamson 2013, p. 111). It is a curious question, which we will not investigate here, whether results of this kind about stronger logics can be used to address the under-generation question in a case from natural language semantics. Another curious question is whether a complete axiomatization of the *invalid* inferences according to a semantics, by means of "refutation axioms" and "refutation rules" as in Łukasiewicz 1957 and Goranko 1994, could shed light on whether a semantics over-generates invalidities, or equivalently, under-generates validities.[3]

It is not clear how problematic it is for a semantics to under-generate entailments relative to what speakers would accept. Other things equal, perhaps a semantics that makes those additional predictions should be preferred. Yet even if a semantics is shown to under-generate in this sense, one might argue that it is still on the right track—that it has the "right form" for the truth conditions, though further assumptions about models are needed to issue in further intuitive entailment predictions. Moreover, how much of what speakers judge to be true should be built into the semantics of the language has been a matter of some controversy. One particularly austere view has it that semantics is only concerned with a very strict notion of logical form, a view expressed early on by Wheeler [1972]: "It is certainly a worthwhile project, when semantics is done, to state some truths using the predicates the semantics has arrived at, but this is to do science, not semantics. . . . The tendency we oppose is the tendency to turn high-level truths into *analytic* truths" (p. 319). We are not endorsing this view, but we mention it to suggest that purported under-generation problems might be more controversial than over-generation problems. (For a detailed discussion of the relation between logical form and lexical semantics, see Glanzberg 2017 in this volume.)

An additional complication concerning under-generation is the fact that many of our intuitions about

[2]Perhaps there are such paradoxes (e.g., the Sorites?), so we do not claim that a complete axiomatization drawing only upon intuitively valid patterns is *decisive* evidence against over-generation. Moreover, the intuitive validity of the inference patterns may be sensitive to what we know about what can be derived by these patterns. Discovering a counter-intuitive consequence of the patterns may disturb our intuitive acceptance of the patterns themselves.

[3]If both the set of validities and its complement, the set of invalidities, are axiomatizable, then of course the logic is decidable. It is noteworthy that all of the logics discussed in our case studies in §§3.3-3.5 are decidable.

entailment presumably derive neither from semantic knowledge nor from general world knowledge, but rather from situational pragmatic reasoning. Axiomatizing aspects of pragmatic reasoning could also be of interest, but one might not want to criticize a semantic theory for failing to capture entailment intuitions that evidently fall outside the scope of semantics. Of course, negotiating this demarcation between semantics and pragmatics is notoriously difficult (see, e.g., the essays collected in Szabó 2005).

3.3 Counterfactuals

The first of our cases studies is the most famous of the three: the semantics of counterfactual conditionals. To fix a simple formal language with which to present a semantics, let the *counterfactual language* be the language of propositional logic plus a new binary sentential connective $\Box\!\!\rightarrow$.[4] The intended reading of $\varphi \Box\!\!\rightarrow \psi$ is "if it were the case that φ, then it would be the case that ψ."

We will consider a semantics for the counterfactual language that is essentially due to Lewis [1971, 1973a, 1973b] and Stalnaker [1968] (also see Stalnaker and Thomason 1970). An *ordering model* for the counterfactual language consists of a set W of worlds, an assignment of subsets of W to proposition letters, and for each world $w \in W$, a reflexive and transitive binary relation $\succeq_w$ on W.[5] In these models, the Boolean connectives are interpreted in the standard way, and the counterfactual conditional is interpreted as follows:

$$[\![\varphi \Box\!\!\rightarrow \psi]\!]^w = 1 \quad \text{iff} \quad \forall x \in [\![\varphi]\!]\; \exists y \in [\![\varphi]\!] \colon y \succeq_w x \text{ and } \forall z \succeq_w y \colon z \in [\![\varphi]\!] \text{ implies } z \in [\![\psi]\!],$$

where $[\![\varphi]\!] = \{v \in W \mid [\![\varphi]\!]^v = 1\}$. As usual, a formula φ is a semantic *consequence* of a set Γ of formulas over a class of models iff for every model in the class and world in the model, if the world makes all the formulas in Γ true, then it makes φ true as well.

As is well known, one way of explaining the above semantics is that the relation $\succeq_w$ is a relation of *comparative similarity*: $x \succeq_w y$ indicates that world x is at least as similar to world w as world y is. The basic idea behind the truth condition is that a counterfactual $\varphi \Box\!\!\rightarrow \psi$ is true at a world w iff among the worlds that make φ true, those that are most similar to w all make ψ true. But in case there does not exist among the φ-worlds any that are *most similar* to w—either because there is an infinite sequence of more and more similar-to-w worlds that make φ true, or because there are worlds that make φ true but are incomparable in similarity to w—the more careful statement is that $\varphi \Box\!\!\rightarrow \psi$ is true at w iff for any φ-world x there is a φ-world y that is at least as similar to w and such that all at least as similar φ-worlds are ψ-worlds.

As we discussed in §2 and §3.2, one way to evaluate such a semantics is according to how well it predicts entailments between sentences of the relevant language fragment. Consider the over-generation question: are there intuitively invalid inference patterns between counterfactual sentences that are counted as valid consequences according to the comparative similarity semantics? Fortunately, a complete finite axiomatization is available for this semantics, which provides information of great value concerning the over-generation question. Let the *basic counterfactual logic* be the smallest normal logic for the counterfactual language that contains the axioms of Figure 1. We say that φ *is derivable from* Γ in the minimal counterfactual logic iff there are $\gamma_1, \dots, \gamma_n \in \Gamma$ such that $(\gamma_1 \wedge \cdots \wedge \gamma_n) \to \varphi$ belongs to this logic.

After the rather trivial axioms **A0** and **A1**, one can think of the rest as corresponding to interesting

[4] This is not the only way of treating the syntax of counterfactuals. For an alternative "restrictor view" of conditionals, see Kratzer 1986.

[5] A more general definition takes $\succeq_w$ to be a reflexive and transitive binary relation on some subset W_w of W, but for simplicity we do not consider this more general definition here. Also note that we have decided to flip Lewis's notation, writing '$x \succeq_w y$' where Lewis would write '$x \preceq_w y$'.

$$\textbf{A0} \quad \neg(\top \mathrel{\Box\!\!\rightarrow} \bot)$$

$$\textbf{A1} \quad p \mathrel{\Box\!\!\rightarrow} p$$

$$\textbf{A2} \quad \big(p \mathrel{\Box\!\!\rightarrow} (q \wedge r)\big) \rightarrow (p \mathrel{\Box\!\!\rightarrow} q)$$

$$\textbf{A3} \quad \big((p \mathrel{\Box\!\!\rightarrow} q) \wedge (p \mathrel{\Box\!\!\rightarrow} r)\big) \rightarrow \big(p \mathrel{\Box\!\!\rightarrow} (q \wedge r)\big)$$

$$\textbf{A4} \quad \big((p \mathrel{\Box\!\!\rightarrow} r) \wedge (q \mathrel{\Box\!\!\rightarrow} r)\big) \rightarrow \big((p \vee q) \mathrel{\Box\!\!\rightarrow} r\big)$$

$$\textbf{A5} \quad \big((p \mathrel{\Box\!\!\rightarrow} q) \wedge (p \mathrel{\Box\!\!\rightarrow} r)\big) \rightarrow \big((p \wedge q) \mathrel{\Box\!\!\rightarrow} r\big)$$

Figure 1: Axioms of the basic counterfactual logic.

inference patterns. Think of each of **A2-A5** as saying that the antecedent of the material conditional entails the consequent. For example, according to **A2**, something of the form *If it were that p, then it would be that q and r* entails *If it were that p, then it would be that q*. According to **A3**, *If it were that p, it would be that q* and *If it were that p, it would be that r* together entail *If it were that p, it would be that q and r*. If one thinks that these patterns are intuitively valid and that the rules of §3.1 preserve intuitive validity, then the following completeness theorem provides strong evidence that the semantics does not over-generate.

Theorem 3.1 (Burgess 1981)**.** For any finite set Γ of formulas and formula φ, the following are equivalent:

1. φ is a consequence of Γ over the class of all ordering models;
2. φ is derivable from Γ in the basic counterfactual logic.[6]

As noted in §3.2, a completeness theorem such as Theorem 3.1 does not resolve the under-generation question. In fact, according to Lewis [1971, 1973a] and Stalnaker [1968], the above semantics does under-generate. According to Lewis, the following should be valid:

$$(p \wedge q) \rightarrow (p \mathrel{\Box\!\!\rightarrow} q);$$
$$(p \mathrel{\Box\!\!\rightarrow} q) \rightarrow (\neg p \vee q).$$

According to Stalnaker, the following should also be valid:

$$(p \mathrel{\Box\!\!\rightarrow} q) \vee (p \mathrel{\Box\!\!\rightarrow} \neg q).$$

None of these three principles is valid according to the basic comparative similarity semantics above, as one can check by constructing simple counter-models. However, the semantics has two parts: the class of models and the truth condition for $\mathrel{\Box\!\!\rightarrow}$. Simply by restricting the class of models, while keeping the truth condition for $\mathrel{\Box\!\!\rightarrow}$ the same, one can obtain a semantics for which the basic counterfactual logic plus Lewis's two axioms above is sound and complete, or for which the basic logic plus Stalnaker's axiom is sound and complete. Let a *centered ordering model* be one in which for every distinct $w, v \in W$, we have $w \succeq_w v$ but $v \not\succeq_w w$, i.e., every world is more similar to itself than any other world is. Let a *linear ordering model* be one in which for every $w, x, y \in W$, we have $x \succeq_w y$ or $y \succeq_w x$, and if $x \succeq_w y$ and $y \succeq_w x$, then $x = y$.

[6] A similar result is proved in the dissertation of Veltman [1985, Theorem II.82, p. 132], which contains a wealth of further information about conditional logic and semantics, motivated by natural language (non-)inferences.

Theorem 3.2 (Burgess 1981)**.** For any finite set Γ of formulas and formula φ:

1. φ is a consequence of Γ over the class of *centered* ordering models iff φ is derivable from Γ in the counterfactual logic extending the basic logic with Lewis's two axioms above.

2. φ is a consequence of Γ over the class of *linear* ordering models iff φ is derivable from Γ in the counterfactual logic extending the basic logic with Stalnaker's axiom above.

3. φ is a consequence of Γ over the class of *centered linear* ordering models iff φ is derivable from Γ in the counterfactual logic extending the basic logic with the three axioms above.[7]

Theorem 3.2 can be viewed in another way. If one were to start with the idea that the *intended models*—that is, the models capturing some notion of what counterfactual statements are *about*—involve an ordering over worlds in which every world is more similar to itself than any other world is, then a result like Theorem 3.2.1 summarizes exactly the additional entailment predictions one obtains by restricting to such models. Analogous points apply to parts 2 and 3 of Theorem 3.2.

The semantics of counterfactuals continues to be a topic of lively discussion, in which alternatives to the Lewis-Stalnaker semantics have been motivated in large part by claims of better fit with entailment data (see, e.g., Santorio 2014, 2017a,b, Willer 2015 and references therein). Axiomatization could be especially helpful in a new phase of counterfactual semantics in showing exactly how the new proposals compare to each other and to the classic similarity semantics in their predictions of entailment.

3.4 Epistemic Comparatives

Our second case study takes us from language for expressing how the world could have been different to language for expressing how we think the world is likely to be. To fix a simple formal language, let the *comparative language* be the language of propositional logic plus a new binary sentential connective $\geqslant$. We read $\varphi \geqslant \psi$ is "φ is at least as likely as ψ." As an abbreviation, we define $\Box\varphi := (p \wedge \neg p) \geqslant \neg\varphi$.

We will consider a semantics for the comparative language that uses the same ordering models used in §3.3 to interpret the counterfactual language—but thought of in a different way, discussed below. To interpret comparatives, the semantics lifts the ordering $\succeq_w$ on *worlds* to an ordering $\succeq_w^\ell$ on *propositions*. Lewis [1973a] and Kratzer [1991] do so in the following way for $A, B \subseteq W$:

$$A \succeq_w^\ell B \quad \text{iff} \quad \forall b \in B\ \exists a \in A \colon a \succeq_w b. \tag{ℓ}$$

With this lifting, comparatives can be interpreted as follows:

$$[\![\varphi \geqslant \psi]\!]^w = 1 \quad \text{iff} \quad [\![\varphi]\!] \succeq_w^\ell [\![\psi]\!].$$

As before, a formula φ is a *consequence* of a set Γ of formulas iff for every ordering model and world therein, if the world makes all of the formulas in Γ true, then it makes φ true.

Kratzer [2012] conceives of the relation $\succeq_w$ as a ranking of "worlds according to how close they come to the normal course of events in the world of evaluation, given a suitable normalcy standard" (p. 39). The normalcy standard is given by a set of propositions—the *ordering source*—representing some conversational

[7]Cf. Theorems II.85 (p. 133) and II.90 (p. 138) of Veltman 1985. There are further logically-relevant distinctions between classes of models, for example between *weakly centered* (which Veltman calls *faithful*) ordering models vs. *centered* ordering models, but here we only hope to give a flavor of the results on counterfactual logics, not an exhaustive summary.

background. Technically, the ordering source O_w is a set of subsets of W, and it induces the relative normality relation $\succeq_w$ as follows: $y \succeq_w x$ iff for every $P \in O_w$, if $x \in P$ then $y \in P$. Thus, y is at least as normal as x relative to w iff y makes true every proposition from O_w that x does and possibly more. Note that any reflexive and transitive relation $\succeq_w$ on W can be generated in this way, by taking O_w to be the set of all subsets $U \subseteq W$ such that for all $v, v' \in W$, if $v \in U$ and $v' \succeq_w v$, then $v' \in U$. Finally, the truth condition for $\geqslant$ given above tells us that φ *is at least as likely as* ψ is true iff for every ψ-world there is a φ-world that is at least as normal according to the appropriate epistemic ordering source.

For a discussion of the advantages of this semantics, we refer the reader to Kratzer 1981, 1991, as well as Yalcin 2010, §3.2. The point we wish to make here arises from a disadvantage of the semantics. Yalcin [2010] and Lassiter [2010, 2015] observed that the above semantics for comparatives validates the principle

$$((\varphi \geqslant \psi) \wedge (\varphi \geqslant \chi)) \to (\varphi \geqslant (\psi \vee \chi)),$$

of which the principle

$$(\varphi \geqslant \chi) \to (\varphi \geqslant (\varphi \vee \chi))$$

is a special case, obtained from the first by taking φ and ψ to be the same. To see that this principle is problematic in the case of comparative likelihood, let φ stand for *the coin lands heads*, and let χ stand for *the coin lands tails*. Then the principle above implies the following:

(1) If it's at least as likely that the coin lands heads as it is that the coin lands tails, then it's at least as likely that the coin lands heads as it is that the coin lands heads or the coin lands tails.

Yet any fair coin shows how the antecedent of the above conditional can be true while the consequent is false. Thus, the principles above are not intuitively valid principles for "at least as likely as." This problem has come to be called *the disjunction problem* for this lifting semantics for epistemic comparatives.

Though it took two or three decades since the appearance of Kratzer's papers for the above disjunction problem to surface, the problem would have been apparent if someone had provided a completeness theorem for an ordering semantics for comparatives like the semantics above. In fact, someone did provide such a completeness theorem in 1973, namely Lewis [1973a]. Lewis's class of models was somewhat different than Kratzer's, for instance in that Lewis assumed that $\succeq_w$ is total,[8] but these differences are orthogonal to the disjunction problem. As suggested above, the disjunction problem can easily be spotted from an inspection of Lewis's completeness result. Unfortunately, Lewis relegated his completeness theorems to the appendix of his book. Moreover, Lewis's [1973a, p. 123] axiomatization for the comparative language contains a rule that may be difficult to digest. Only in a footnote does Lewis [1973a, p. 124] reformulate that rule with the help of an axiom that makes the disjunction problem quite clear:

$$(\varphi \geqslant (\varphi \vee \psi)) \vee (\psi \geqslant (\varphi \vee \psi)).$$

Applied to comparative likelihood, this principle implies the following:

(2) Either it's at least as likely that the coin lands heads as it is that the coin lands heads or tails, or it's at least as likely that the coin lands tails as it is that the coin lands heads or tails.

Thus, if one considers the axiom in Lewis's footnote, it is quite clear that the above semantics will not work for comparative likelihood.[9] It is an interesting counterfactual question whether, if Lewis had not relegated

[8] That is, $x \succeq_w y$ or $y \succeq_w x$ for all worlds x and y in the field of $\succeq_w$.

[9] Lewis [1973a, §2.5] was interested in comparative *possibility* rather than likelihood.

his completeness results to his appendix, and if he had put more emphasis on understandable axioms, then the discovery of the disjunction problem might not have been delayed by several decades.

A completeness theorem for exactly the semantics for comparatives above (without Lewis's totality assumption) was given by Halpern [1997, 2003]. Strikingly, the disjunction problem manifests itself in the key axiom in the axiomatization—providing more evidence for the thesis that over-generation problems like the disjunction problem can be brought to light by completeness theorems. For Halpern's completeness result, let the *simple comparative logic* be the smallest normal logic for the comparative language that contains the axioms of Figure 2. The disjunction problem is represented by axiom **C4**. Given a set Γ of formulas and a formula φ, we say that φ *is derivable from* Γ in the simple comparative logic iff there are $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that $(\gamma_1 \wedge \cdots \wedge \gamma_n) \to \varphi$ belongs to the simple comparative logic.

C1 $(p \geqslant p)$

C2 $\Box(p \to q) \to (q \geqslant p)$

C3 $\big((p \geqslant q) \wedge (q \geqslant r)\big) \to (p \geqslant r)$

C4 $\big((p \geqslant q) \wedge (p \geqslant r)\big) \to \big(p \geqslant (q \vee r)\big)$

Figure 2: Axioms of the simple comparative logic.

Theorem 3.3 (Halpern 2003)**.** For any finite set Γ of formulas and formula φ, the following are equivalent:

1. φ is a consequence of Γ according to the Lewis-Kratzer lifting semantics using (ℓ);
2. φ is derivable from Γ in the simple comparative logic.

Based in part on the problematic entailment predictions of the ordering semantics above, Yalcin [2010] and Lassiter [2010, 2015] have proposed semantics for epistemic comparatives based on probability models. Instead of using qualitative orderings $\succeq_w$ over possible worlds, such models use finitely additive probability measures μ_w assigning probability values in $[0,1]$ to propositions. The semantic clause for $\geqslant$ becomes

$$[\![\varphi \geqslant \psi]\!]^w = 1 \quad \text{iff} \quad \mu_w([\![\varphi]\!]) \geq \mu_w([\![\psi]\!]).$$

Several examples of differences in entailment predictions between the qualitative ordering semantics and this numerical probability semantics are given in Yalcin 2010.

In fact, the entailment predictions of semantics based on probability measures (as well as many other kinds of measures) can be, and indeed have been, completely axiomatized as well [Segerberg, 1971, Gärdenfors, 1975]. One noteworthy principle validated by probability models is the principle of comparability, stating that any two propositions are comparable in likelihood:

$$(\varphi \geqslant \psi) \vee (\psi \geqslant \varphi).$$

A number of authors in the literature on foundations of probability, famously including Keynes [1921, §3], have questioned this assumption (see Fine 1973, p. 18), and it is not obvious that ordinary speakers of English would judge that every two propositions ought to be comparable in "probability" or "likelihood" either.

Instead of considering models based on a single probability measure, one can consider models based on sets of probability measures (as suggested in Lassiter 2011, p. 81), with a truth condition of the form

$$[\![\varphi \geqslant \psi]\!]^w = 1 \quad \text{iff} \quad \text{for all } \mu \in P_w\colon \mu([\![\varphi]\!]) \geq \mu([\![\psi]\!]).$$

An axiomatization for this class of models can be obtained from a suitable axiomatization of single-probability-measure models by simply dropping the axiom of comparability [Alon and Heifetz, 2014].[10]

A natural question is whether there could be a qualitative semantics—for example, in the spirit of Kratzer's semantics—that validates the same principles as the probability measure or set-of-measures semantics. As Yalcin [2010] asks, "Is there some better way of extending a preorder over worlds to a preorder over propositions, one which will get the inference patterns right?" (p. 923). In fact, there is a way of modifying the Lewis-Kratzer lifting, introduced in Holliday and Icard 2013, so that it validates exactly the same inference patterns in the comparative language as the set-of-measures semantics. Observe that the Lewis-Kratzer lifting in (ℓ) can equivalently be defined by

$$A \succeq_w^\ell B \quad \text{iff} \quad \text{there is an inflationary function } f\colon B \to A,$$

where *inflationary* means that $f(x) \succeq_w x$ for all $x \in B$. The source of the disjunction problem is that even if there are many more ways that B could come true, as long as there is one way that A could come true that is at least as highly ranked by $\succeq_w$ as each of the ways B could come true, the above lifting says that A is at least as likely as B. A natural idea for fixing this problem is to require that for each distinct way that B could come true, there is a distinct way that A could come true that is at least as highly ranked by $\succeq_w$. That is, the function f should be *injective*: if $x \neq y$, then $f(x) \neq f(y)$. This leads to the following modified lifting:

$$A \succeq_w^m B \quad \text{iff} \quad \text{there is an inflationary injection } f\colon B \to A. \qquad (m)$$

Now the interpretation of $\varphi \geqslant \psi$ is exactly as in Kratzer's semantics, but using $\succeq_w^m$ instead of $\succeq_w^\ell$.[11] In addition, we now take as our models *Noetherian* ordering models in which there is no infinite sequence $w_1, w_2, w_3, \ldots$ of distinct worlds such that $w_1 \preceq_w w_2 \preceq_w w_3 \ldots$. These are models in which each world ordered by $\succeq_w$ is assumed to have some likelihood, so that if there were such an infinite sequence, then we would have an infinite sequence $\{w_1\}, \{w_1, w_2\}, \{w_1, w_2, w_3\}, \ldots$ of propositions each more likely than the previous one by some nondecreasing increment—whereas there should be some maximum likelihood that a proposition may have. Similarly, in probabilistic semantics, there cannot be an infinite sequence $w_1, w_2, w_3, \ldots$ of distinct worlds such that $0 \neq \mu(\{w_1\}) \leq \mu(\{w_2\}) \leq \mu(\{w_3\}) \ldots$, for then the sequence of probabilities $\mu(\{w_1\}), \mu(\{w_1, w_2\}), \mu(\{w_1, w_2, w_3\}), \ldots$ would grow without bound.

It is shown in Harrison-Trainor et al. 2017 that the injective lifting semantics above has exactly the same logic as the set-of-measures semantics.[12]

[10] This is only true for a *suitable* axiomatization of single-probability-measure models. It is not the case that the axiomatization for sets-of-measure semantics can be obtained from that of single-measure semantics found in Segerberg 1971 and Gärdenfors 1975 simply by dropping the axiom of comparability. One must drop comparability and then strengthen the "cancellation" axiom used in Segerberg 1971 and Gärdenfors 1975 to the generalized cancellation axiom in Ríos Insua 1992, Alon and Lehrer 2014, and Alon and Heifetz 2014. For further explanation, see Harrison-Trainor et al. 2016.

[11] More precisely, in the models of Holliday and Icard 2013, $\succeq_w$ is a relation on a distinguished set $R(w)$ of epistemic possibilities, and the lifted $\succeq_w^m$ is a relation on $\wp(R(w))$, so the semantic clause is $[\![\varphi \geqslant \psi]\!]^w = 1$ iff $[\![\varphi]\!] \cap R(w) \succeq_w^m [\![\psi]\!] \cap R(w)$.

[12] In Holliday and Icard 2013, we stated that the logic obtained from that of Segerberg 1971 and Gärdenfors 1975 by dropping the axiom of comparability is *sound* with respect to ordering models with the injective lifting of (m), and we left completeness as an open question. In fact, as suggested in footnote 10, that logic is not complete with respect to the semantics, since when one drops the comparability axiom one must also strengthen the "cancellation" axiom.

Theorem 3.4. For any finite set Γ of formulas and formula φ, the following are equivalent:

1. φ is a consequence of Γ according to the set-of-measures semantics;
2. φ is a consequence of Γ according to the injective lifting semantics using (m).

At this point one could raise a number of methodological questions about the choice between probabilistic vs. qualitative semantics for "at least as likely as." For such a discussion, we refer to Holliday and Icard 2013 and Lassiter 2017. Our focus here is rather on general themes about the value of axiomatizations. Theorem 3.4 is not itself an axiomatization result, but its proof takes advantage of one. As noted above, the logic of set-of-measures models has been completely axiomatized [Alon and Heifetz, 2014]. The implication from 1 to 2 in Theorem 3.4 can therefore be established by proving that the complete logic of set-of-measures models is *sound* with respect to ordering models with the injective lifting (again see Harrison-Trainor et al. 2017).[13] This shows another way in which completeness theorems can be useful—this time in establishing that two semantics are equivalent in terms of their entailment predictions.

Finally, we would like to mention one other kind of semantics for the comparative language, in which the gap between inference patterns and semantics is much smaller. Each of the semantics discussed above generates an ordering on propositions from something considered to be more basic: an ordering on worlds or a (set of) probability measure(s). Another approach would be to take as a model for the comparative language a set W of worlds together with a collection of orderings $\succsim_w$ on the powerset of W, required to satisfy certain properties, with which comparatives are interpreted by

$$[\![\varphi \geqslant \psi]\!]^w = 1 \quad \text{iff} \quad [\![\varphi]\!] \succsim_w [\![\psi]\!].$$

Lewis [1973a, pp. 54-55] considers this approach to the semantics of comparative possibility (his "comparative possibility systems"), and in Holliday and Icard 2013, we considered this approach to the semantics of comparative likelihood (our "event ordering models"). On such an approach, the inference patterns that are valid for the comparative operator $\geqslant$ come directly from the properties assumed for the orderings $\succsim_w$ of propositions. For example, if we require that each $\succsim_w$ satisfies the transitivity property

$$\text{if } A \succsim_w B \text{ and } B \succsim_w C, \text{ then } A \succsim_w C,$$

then the semantics validates the principle

$$((\varphi \geqslant \psi) \wedge (\psi \geqslant \chi)) \rightarrow (\varphi \geqslant \chi).$$

If we require that each $\succsim_w$ satisfies the "qualitative additivity" property

$$A \succsim_w B \text{ iff } A \setminus B \succsim_w B \setminus A,$$

then the semantics validates the principle

$$(\varphi \geqslant \psi) \leftrightarrow ((\varphi \wedge \neg\psi) \geqslant (\psi \wedge \neg\varphi)).$$

[13] For the implication from 2 to 1, it can be shown that any set-of-measures model that serves as a counterexample to φ being a consequence of Γ can be transformed into an ordering model with the injective lifting that serves as a counterexample to φ being a consequence of Γ (see Harrison-Trainor et al. 2017, Lemma 5).

Assuming the above two properties (plus some obvious properties like $A \succsim_w \emptyset$ and $\emptyset \not\succsim_w W$) gives a semantics that makes entailment predictions for the fragment of language in question that may be very difficult to distinguish empirically from those of the probabilistic semantics discussed above (see Holliday and Icard 2013). However, one has the feeling that this "semantics" is more like a direct algebraic encoding of the inference patterns we want to be valid—that it is merely axiomatics in disguise (cf. van Benthem 2001, p. 358). As a result, completeness theorems for such semantics are not very surprising: the valid principles are just what we built in when defining the models. By contrast, the ordering models and probabilistic models discussed above are much farther from direct encodings of desired entailments. Thus, completeness theorems for these kinds of models can be much more illuminating.

3.5 Indicative Conditionals

Our third and final case study involves language for expressing uncertainty without explicit talk of likelihood. For our formal language, let the *modal-indicative language* be the language of propositional logic plus a new unary sentential operator $\Diamond$ and a new binary sentential connective $\Rightarrow$. The intended readings of $\Diamond\varphi$ and $\varphi \Rightarrow \psi$ are "It might be that φ" and "if φ, then ψ", respectively. We define $\Box\varphi := \neg\Diamond\neg\varphi$.

We will discuss a semantics for the modal-indicative language due to Yalcin [2012]. This is a simplified version of the semantics proposed in Yalcin 2007, which has a more subtle semantic clause for indicative conditionals, in order to deal with epistemic modals in the antecedents of such conditionals. For non-modal antecedents, the two semantics for the indicative conditional are equivalent. We refer the reader to Holliday and Icard 2017 for discussion of semantics designed to handle modal antecedents in indicative conditionals (especially one due to Kolodny and MacFarlane 2010).

A *domain model* for the modal-indicative language is simply a set W of worlds together with an assignment of subsets of W to proposition letters. But now formulas are evaluated with respect to both a world $w \in W$ and an *information state* $i \subseteq W$, with these key clauses:

$$\begin{aligned} [\![\Diamond\psi]\!]^{w,i} = 1 \quad &\text{iff} \quad i \cap [\![\psi]\!]^i \neq \varnothing \\ [\![\varphi \Rightarrow \psi]\!]^{w,i} = 1 \quad &\text{iff} \quad i + \varphi \subseteq [\![\psi]\!]^{i+\varphi}, \end{aligned}$$

where $[\![\psi]\!]^i = \{w \in W \mid [\![\psi]\!]^{w,i} = 1\}$ and $i + \varphi = i \cap [\![\varphi]\!]^i$. Intuitively, *it might be that* φ is true relative to an information state i just in case i contains some φ-possibility; and *if* φ*, then* ψ is true just in case updating i with the information φ yields an information state in which ψ is true throughout. For the dual $\Box$ of $\Diamond$, note that we have $[\![\Box\psi]\!]^{w,i} = 1$ iff $i \subseteq [\![\psi]\!]^i$, and $[\![\varphi \Rightarrow \psi]\!]^i = [\![\Box\psi]\!]^{i+\varphi}$.

Concerning the notion of semantic consequence for this language, we now deviate from the standard definition of consequence used in the previous two case studies. Following Yalcin [2012, p. 1019], we say that a formula φ is an *informational consequence* of a set Γ of formulas iff for every model W and information state $i \subseteq W$, if $i \subseteq [\![\gamma]\!]^i$ ("i accepts γ") for every $\gamma \in \Gamma$, then $i \subseteq [\![\varphi]\!]^i$ ("i accepts φ"). An argument from Yalcin 2007 in favor of this definition of consequence for a language with epistemic modals is that it predicts the defectiveness of sentences like "It is raining and it might not be raining", formalized as $p \wedge \Diamond\neg p$. According to the informational consequence relation, $p \wedge \Diamond\neg p$ is *contradictory*, in the sense that $p \wedge \neg p$ is an informational consequence of $p \wedge \Diamond\neg p$. For the only way that an information state i can be such that $i \subseteq [\![p \wedge \Diamond\neg p]\!]^i$ is for i to be empty, because if $\Diamond\neg p$ is true at w relative to i, then there is a $\neg p$-world in i, while $i \subseteq [\![p \wedge \Diamond\neg p]\!]^i$ requires that all worlds in i are p-worlds; but if i is empty, then $i \subseteq [\![p \wedge \neg p]\!]^i$.

The semantics above, with its information shifting clause for the conditional and its modified notion of consequence, looks quite different than traditional semantics for classical logic. Does this pose any obstacle to axiomatization? Fortunately, it does not. In the field of *dynamic epistemic logic* [van Ditmarsch et al., 2008, van Benthem, 2011], such information shifting semantic clauses are standard, and axiomatizations for such semantics abound. Moreover, the notion of informational consequence is familiar from the notion of *global consequence* in modal logic (Blackburn et al. 2001, §1.5, Kracht 1999, §3.1). Using these connections, we can completely axiomatize the informational consequence relation for the modal-indicative language, as in Theorem 3.5 below, using standard techniques from dynamic epistemic logic and modal logic.[14]

Let the *Yalcin logic* be the smallest logic (in the sense of §3.1) for the modal-indicative language that contains the axioms of Figure 3. Given a set Γ of formulas and a formula φ, we say that φ *is derivable from* Γ in the Yalcin logic iff there are $\gamma_1, \ldots, \gamma_n \in \Gamma$ such that $(\Box\gamma_1 \wedge \cdots \wedge \Box\gamma_n) \rightarrow \Box\varphi$ belongs to the Yalcin logic. Note that this differs from the definition of derivability in the previous sections, in a way that matches the difference between informational and classical consequence.

D $\Box\varphi \rightarrow \Diamond\varphi$ **4** $\Diamond\Diamond\varphi \rightarrow \Diamond\varphi$ **5** $\Diamond\Box\varphi \rightarrow \Box\varphi$

I1 $(\varphi \Rightarrow \pi) \leftrightarrow \Box(\varphi \rightarrow \pi)$ for π nonmodal

I2 $(\varphi \Rightarrow (\alpha \wedge \beta)) \leftrightarrow ((\varphi \Rightarrow \alpha) \wedge (\varphi \Rightarrow \beta))$

I3 $(\varphi \Rightarrow \alpha) \rightarrow (\varphi \Rightarrow (\alpha \vee \beta))$

I4 $(\varphi \Rightarrow \alpha) \rightarrow (\varphi \Rightarrow \Box\alpha)$

I5 $((\varphi \Rightarrow (\alpha \vee \Box\beta)) \wedge \neg(\varphi \Rightarrow \beta)) \rightarrow (\varphi \Rightarrow \alpha)$

I6 $((\varphi \Rightarrow (\alpha \vee \Diamond\beta)) \wedge (\varphi \Rightarrow \neg\beta)) \rightarrow (\varphi \Rightarrow \alpha)$

I7 $\neg(\varphi \Rightarrow \beta) \rightarrow (\varphi \Rightarrow \Diamond\neg\beta)$

Figure 3: Axioms of the Yalcin logic.

Now we have the following completeness theorem, proved in Holliday and Icard 2017.

Theorem 3.5. For any finite set Γ of formulas and formula φ, the following are equivalent:

1. φ is an informational consequence of Γ over domain models;
2. φ is derivable from Γ in the Yalcin logic.

This axiomatization result brings to light some key entailment predictions of the semantics. First, **I1** tells us that if the indicative conditional has no modals in the consequent (and recall we have been using the version of Yalcin's semantics suitable for non-modal antecedents), then the indicative conditional is equivalent to the strict conditional. Axioms **I2** and **I3** are not too surprising, but axiom **I4** corresponds to the key commitment of the semantics that a sentence like

(3) If Miss Scarlet didn't do it, then Colonel Mustard did it.

entails

[14] A quite different approach is pursued by Bledin [2014], motivated by considerations related to our point (I) from §2. Bledin starts with the same semantics for the modal-indicative language as above and then proposes a Fitch-style natural deduction system for an extension of the language that includes symbols for information states and information acceptance relations.

(4) If Miss Scarlet didn't do it, then it must be that Colonel Mustard did it.

Axioms **I5** and **I6** also correspond to interesting entailment predictions. Axiom **I5** suggests the prediction, for example, that the sentence

(5) If Miss Scarlet did it, then either Colonel Mustard was her accomplice or it must be that Professor Plum was involved.

together with

(6) It's not the case that if Miss Scarlet did it, then Professor Plum was involved.

entails

(7) If Miss Scarlet did it, then Colonel Mustard was her accomplice.

And axiom **I6** suggests the prediction, for example, that

(8) If Miss Scarlet did it, then either she used the pipe or she might have used the candlestick.

together with

(9) If Miss Scarlet did it, she didn't use the candlestick.

entails

(10) If Miss Scarlet did it, then she used the pipe.

Finally, axiom **I7** suggests the prediction that one who *rejects*

(11) If Miss Scarlet was in the ballroom, then Colonel Mustard is guilty

should accept

(12) If Miss Scarlet was in the ballroom, it might be that Colonel Mustard is not guilty.

Our point here is not about whether the above predictions are empirically accurate or not, but rather that the axioms that naturally appear in a complete axiomatization for the semantics can be easily related to empirically assessable entailment predictions of the semantics. Of course, this is not inevitable. There can be axiomatizations containing complicated principles that are far removed from empirically assessable entailment predictions (see, e.g., Holliday and Icard 2013 for discussion of complicated principles in the axiomatizations of probabilistic semantics mentioned in §3.4). However, in many cases, arguably including all of the axiomatizations we have presented in our case studies above, there is not a great distance between axioms of the logic and empirically assessable entailment predictions.

It is worth stepping back and realizing how desirable it would be in any scientific field to have a complete set of axioms for some theory that are themselves close to empirically assessable predictions. From this perspective, the availability of such axiomatizations in the science of meaning is rather remarkable.

4 Conclusion

We have now seen several ways in which axiomatizations can be helpful in elucidating a semantically defined consequence relation. Axiomatizations can provide assurance that a consequence relation does not overgenerate entailments by making manifest the problematic entailment predictions of a semantics; and they can be used to determine the relative strength or equivalence of two consequence relations defined by different semantics. In addition, we have seen that even when a semantics has a form quite different from those familiar from classical model theory, this might not pose any obstacle to axiomatization.

In this paper, we have assumed that the models come first and the axiomatizations are then to be discovered. Yet in some cases it may seem that the model-theoretic proposals are largely guided by the task of delivering the right entailment predictions. Given this apparent primacy of entailment and inference patterns, one might wonder whether the semanticist ought simply to focus attention on proof systems themselves and eschew model theory altogether. In that case, axiomatic systems, together with a specific deductive apparatus intended to capture natural inferential patterns, would be the main object of study. This kind of project has of course been pursued within linguistic semantics (see, e.g., Francez and Dyckhoff 2010, or Szabolcsi 2007 for an appraisal), and there is a distinguished tradition of proof theoretic semantics within philosophy of logic and language. For instance, Prawitz [2006] characterizes the project as follows:

> One very simple version of an approach of this kind is to take meaning to be determined by all the rules for a language. Restricting oneself to deductive uses of language and thinking of proofs as determined by a set of inference rules, meaning simply becomes determined by all the inference rules of the language. (p. 509)

Others are less sanguine about the idea of restricting attention to inferential patterns only. For example, Dummett [2000] writes:

> No one could think that the grasp of the meaning of an arbitrary sentence consisted solely in a knowledge of the way in which it might figure in an inference, as premiss or conclusion. . . . If we take it as the primary function of a sentence to convey information, then it is natural to view a grasp of the meaning of a sentence as consisting in an awareness of its *content*. . . . (p. 252)

In a classical picture, one captures content with truth conditions (as opposed to verification conditions as in Dummett 2000). Dowty et al. [1981] argue in favor of a truth-conditional rather proof-theoretic approach on the grounds that:

> The definition of truth with respect to a model has the advantage that it allows us to capture the definitions of logical truth, logical entailment, and related notions *and* at the same time to capture our intuitions of the essential "aboutness" of natural language. (p. 52)

Whether this "aboutness" is intended to capture actual reference, or merely some analysis of how people think about the world (in the sense of Bach's [1986] "natural language metaphysics"), by far the dominant trend in the field has been to take the models to be of independent interest. Thus, model theoretic objects including worlds, situations, possibilities, events, similarity orders, time points and intervals, mereological sums and parts, and all the rest are typically taken to be first class citizens in the study of linguistic meaning (see, e.g., Krifka 1998 for lucid discussion of this view). One might conclude from this that there is a tension between model-theoretic semantics and the study of the kinds of axiomatic systems we have been discussing (cf. Dowty et al. 1981, who speak of "preferring the semantic method to the deductive method"). However, we hope to have shown that the situation is exactly the opposite. It is precisely when the model theory becomes more intricate and more divorced from a mere "encoding" of inferential relations that axiomatization promises to be most useful and enlightening.

Acknowledgements. For helpful comments, we wish to thank Johan van Benthem, Justin Bledin, Fabrizio Cariani, Melissa Fusco, Alex Kocurek, Daniel Lassiter, Paolo Santorio, and Seth Yalcin. We also wish to thank Derek Ball and Brian Rabern for the invitation to contribute to the present volume.

References

Shiri Alon and Aviad Heifetz. The logic of Knightian games. *Economic Theory Bulletin*, 2(2):161–182, 2014.

Shiri Alon and Ehud Lehrer. Subjective multi-prior probability: A representation of a partial likelihood relation. *Journal of Economic Theory*, 151:476–492, 2014.

Emmon Bach. Natural Language Metaphysics. In Ruth Barcan Marcus, Georg J.W. Dorn, and Paul Weingartner, editors, *Logic, Methodology and Philosophy of Science VII*, pages 573–595. North-Holland, Amsterdam, 1986.

Johan van Benthem. *Essays in Logical Semantics*. D. Reidel Publishing Company, Dordrecht, 1986.

Johan van Benthem. Correspondence Theory. In Dov M. Gabbay and F. Guenthner, editors, *Handbook of Philosophical Logic*, volume 3, pages 325–408. Springer, Dordrecht, 2nd edition, 2001.

Johan van Benthem. *Logical Dynamics of Information and Interaction*. Cambridge University Press, Cambridge, 2011.

Patrick Blackburn and Johan Bos. *Representation and Inference in Natural Language: A First Course in Computational Semantics*. CSLI Press, Stanford, 2005.

Patrick Blackburn, Maarten de Rijke, and Yde Venema. *Modal Logic*. Cambridge University Press, New York, 2001.

Justin Bledin. Logic Informed. *Mind*, 123(490):277–316, 2014.

Rachael Briggs. Interventionist counterfactuals. *Philosophical Studies*, 160(1):139–166, 2012.

John P. Burgess. Quick Completeness Proofs for Some Logics of Conditionals. *Notre Dame Journal of Formal Logic*, 22(1):76–84, 1981.

Fabrizio Cariani. 'Ought' and Resolution Semantics. *Noûs*, 47(3):534–558, 2013.

Fabrizio Cariani. Consequence and Contrast in Deontic Semantics. *Journal of Philosophy*, 2017. Forthcoming.

Gennaro Chierchia and Sally McConnell-Ginet. *Meaning and Grammar*. MIT Press, Cambridge, Mass., second edition, 2001.

M. J. Cresswell. The Autonomy of Semantics. In Stanley Peters and Esa Saarinen, editors, *Processes, Beliefs, and Questions*, pages 69–86. D. Reidel Publishing Company, Dordrecht, 1982.

Hans van Ditmarsch, Wiebe van der Hoek, and Barteld Kooi. *Dynamic Epistemic Logic*. Springer, Dordrecht, 2008.

David Dowty. The Role of Negative Polarity and Concord Marking in Natural Language Reasoning. In Mandy Harvey and Lynn Santelmann, editors, *Proceedings of Semantics and Linguistic Theory (SALT) 4*, pages 114–144. Cornell University, Department of Modern Languages and Linguistics, Ithaca, N.Y., 1994.

David R. Dowty, Robert E. Wall, and Stanley Peters. *Introduction to Montague Semantics*. Kluwer, Dordrecht, 1981.

Michael Dummett. *Elements of Intuitionism*. Clarendon Press, Oxford, 2nd edition, 2000.

Terrence L. Fine. *Theories of Probability*. Academic Press, New York, 1973.

Nissim Francez and Roy Dyckhoff. Proof-theoretic semantics for a natural language fragment. *Linguistics and Philosophy*, 33(6):447–477, 2010.

David Galles and Judea Pearl. An Axiomatic Characterization of Causal Counterfactuals. *Foundations of Science*, 3(1):151–182, 1998.

Daniel Gallin. *Intensional and Higher-Order Modal Logic*. North-Holland, Amsterdam, 1975.

Peter Gärdenfors. Qualitative Probability as an Intensional Logic. *Journal of Philosophical Logic*, 4(2): 171–185, 1975.

Michael Glanzberg. Lexical meaning, concepts, and the metasemantics of predicates. In Derek Ball and Brian Rabern, editors, *The Science of Meaning: Essays on the Metatheory of Natural Language Semantics*. Oxford University Press, Oxford, 2017.

Valentin Goranko. Refutation Systems in Modal Logic. *Studia Logica*, 53(2):299–324, 1994.

Joseph Y. Halpern. Defining Relative Likelihood in Partially-Ordered Preferential Structures. *Journal of AI Research*, 7:1–24, 1997.

Joseph Y. Halpern. *Reasoning about Uncertainty*. MIT Press, Cambridge, Mass., 2003.

Joseph Y. Halpern. From Causal Models to Counterfactual Structures. *Review of Symbolic Logic*, 6(2): 305–322, 2013.

Matthew Harrison-Trainor, Wesley H. Holliday, and Thomas F. Icard. A note on cancellation axioms for comparative probability. *Theory and Decision*, 80(1):159–166, 2016.

Matthew Harrison-Trainor, Wesley H. Holliday, and Thomas F. Icard. Preferential Structures for Comparative Probabilistic Reasoning. In Satinder Singh and Shaul Markovitch, editors, *Proceedings of the Thirty-First AAAI Conference on Artificial Intelligence (AAAI-17)*, pages 1135–1141. AAAI Publications, 2017.

Wesley H. Holliday and Thomas F. Icard. Measure semantics and qualitative semantics for epistemic modals. In Todd Snider, editor, *Proceedings of Semantics and Linguistic Theory (SALT) 23*, pages 514–534. LSA and CLC Publications, 2013.

Wesley H. Holliday and Thomas F. Icard. Indicative Conditionals and Dynamic Epistemic Logic. Manuscript, 2017.

Thomas F. Icard and Lawrence S. Moss. A Complete Calculus of Monotone and Antitone Higher-Order Functions. In *Proceedings of Topology, Algebra, and Categories in Logic (TACL) 2013*, volume 23 of *EPiC Series*, pages 96–99. Vanderbilt University, 2013.

Thomas F. Icard and Lawrence S. Moss. Recent Progress on Monotonicity. *Linguistic Issues in Language Technology*, 9(7):167–194, 2014.

John Maynard Keynes. *A Treatise on Probability.* The Macmillan Company, London, 1921.

Niko Kolodny and John MacFarlane. Ifs and Oughts. *Journal of Philosophy*, 107(3):115–143, 2010.

Philipp Koralus and Salvador Mascarenhas. The Erotetic Theory of Reasoning: Bridges Between Formal Semantics and the Psychology of Deductive Inference. *Philosophical Perspectives*, 27(1):312–365, 2013.

M. Kracht. *Tools and Techniques in Modal Logic.* Elsevier, New York, 1999.

David H. Krantz, R. Duncan Luce, Patrick Suppes, and Amos Tversky. *Foundations of Measurement*, volume 1. Academic Press, New York, 1971.

Angelika Kratzer. The Notional Category of Modality. In Hans-Jürgen Eikmeyer and Hannes Rieser, editors, *Words, Worlds, and Contexts: New Approaches in Word Semantics*, pages 38–74. de Gruyter, Berlin, 1981.

Angelika Kratzer. Conditionals. In Anne M. Farley, Peter T. Farley, and Karl-Erik McCullough, editors, *Papers from the Parasession on Pragmatics and Grammatical Theory*, pages 115–135. Chicago Linguistics Society, 1986.

Angelika Kratzer. Modality. In Arnim von Stechow and Dieter Wunderlich, editors, *Semantics: An International Handbook of Contemporary Research*, pages 639–650. de Gruyter, Berlin, 1991.

Angelika Kratzer. *Modals and Conditionals.* Oxford University Press, New York, 2012.

Manfred Krifka. The Origins of Telicity. In Susan D. Rothstein, editor, *Events and Grammar*, Studies in Linguistics and Philosophy, pages 197–236. Springer, Dordrecht, 1998.

Daniel Lassiter. Gradable Epistemic Modals, Probability, and Scale Structure. In Nan Li and David Lutz, editors, *Proceedings of Semantics and Linguistic Theory (SALT) 20*, pages 1–18. CLC Publications, 2010.

Daniel Lassiter. Measurement and Modality: The Scalar Basis of Modal Semantics. New York University PhD dissertation, 2011.

Daniel Lassiter. Epistemic Comparison, Models of Uncertainty, and the Disjunction Puzzle. *Journal of Semantics*, 32(4):649–684, 2015.

Daniel Lassiter. *Graded Modality.* Oxford University Press, Oxford, 2017.

David Lewis. General Semantics. *Synthese*, 22(1/2):18–67, 1970.

David Lewis. Completeness and Decidability of Three Logics of Counterfactual Conditionals. *Theoria*, 37 (1):74–85, 1971.

David Lewis. *Counterfactuals*. Basil Blackwell, Oxford, 1973a.

David Lewis. Counterfactuals and Comparative Possibility. *Journal of Philosophical Logic*, 2(4):418–446, 1973b.

Jan Łukasiewicz. *Aristotle's Syllogistic from the Standpoint of Modern Formal Logic*. Clarendon Press, Oxford, 1957.

David A. McAllester and Robert Givan. Natural language syntax and first-order inference. *Artificial Intelligence*, 56(1):1–20, 1992.

Richard Montague. Pragmatics and Intensional Logic. *Synthese*, 22(1/2):68–94, 1970.

Lawrence S. Moss. Natural Logic and Semantics. In Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz, editors, *Logic, Language, and Meaning*, pages 84–93. Springer, Berlin, 2010.

Reinhard Muskens. Higher Order Modal Logic. In Patrick Blackburn, J.F.A.K. van Benthem, and Frank Wolter, editors, *The Handbook of Modal Logic*, pages 621–654. Elsevier, Amsterdam, 2007.

Ian Pratt-Hartmann and Lawrence S. Moss. Logics for the Relational Syllogistic. *Review of Symbolic Logic*, 2(4):647–683, 2009.

Dag Prawitz. Meaning Approached via Proofs. *Synthese*, 148(3):507–524, 2006.

D. Ríos Insua. On the foundations of decision making under partial information. *Theory and Decision*, 33 (1):83–100, 1992.

Lance J. Rips. *The Psychology of Proof: Deduction in Human Thinking*. MIT Press, Cambridge, Mass., 1994.

Paolo Santorio. Filtering Semantics for Counterfactuals. In Todd Snider, Sarah D'Antonio, and Mia Weigand, editors, *Proceedings of Semantics and Linguistic Theory (SALT) 24*, pages 494–513. LSA and CLC Publications, 2014.

Paolo Santorio. Alternatives and Truthmakers in Conditional Semantics. *Journal of Philosophy*, 2017a. Forthcoming.

Paolo Santorio. Interventions in Premise Semantics. *Philosophers' Imprint*, 2017b. Forthcoming.

Krister Segerberg. Qualitative Probability in a Modal Setting. In J. E. Fenstad, editor, *Second Scandinavian Logic Symposium*, pages 341–352. North-Holland, Amsterdam, 1971.

Robert C. Stalnaker. A Theory of Conditionals. In Nicholas Rescher, editor, *Studies in Logical Theory*, volume 2, pages 98–112. Basil Blackwell, Oxford, 1968.

Robert C. Stalnaker and Richmond H. Thomason. A semantic analysis of conditional logic. *Theoria*, 36(1): 23–42, 1970.

Patrick Suppes. The Axiomatic Method in the Empirical Sciences. In Leon Henkin, John Addison, C. C. Chang, William Craig, Dana Scott, and Robert Vaught, editors, *Proceedings of the Tarski Symposium*, pages 465–479. American Mathematical Society, Providence, RI, 1974.

Zoltán Gendler Szabó, editor. *Semantics versus Pragmatics.* Oxford University Press, Oxford, 2005.

Anna Szabolcsi. Questions about Proof Theory vis-à-vis Natural Language Semantics. ESSLLI, 2007. URL http://semanticsarchive.net/Archive/Dk1MTNkO/szabolcsi_questions_about_proof_theory_2007.pdf.

Alfred Tarski. Über den Begriff der logischen Folgerung. *Actes du Congrès International de Philosophie Scientifique*, 7(394):1–11, 1936.

Frederik Van De Putte. Coarse deontic logic (short version). In Olivier Roy, Allard Tamminga, and Malte Willer, editors, *Deontic Logic and Normative Systems: 13th International Conference, DEON 2016, Bayreuth, Germany, July 18-21, 2016.* College Publications, 2016.

Frank Veltman. *Logics for Conditionals.* PhD thesis, University of Amsterdam, 1985.

Samuel C. Wheeler. Attributes and their Modifiers. *Noûs*, 6(4):310–334, 1972.

Malte Willer. Simplifying Counterfactuals. In Thomas Brochhagen, Floris Roelofsen, and Nadine Theiler, editors, *Proceedings of the 20th Amsterdam Colloquium*, pages 428–437, 2015.

Timothy Williamson. *Modal Logic as Metaphysics.* Oxford University Press, Oxford, 2013.

Seth Yalcin. Epistemic Modals. *Mind*, 116(464):983–1026, 2007.

Seth Yalcin. Probability Operators. *Philosophy Compass*, 5(11):916–937, 2010.

Seth Yalcin. A Counterexample to Modus Tollens. *Journal of Philosophical Logic*, 41(6):1001–1024, 2012.