The Semantics of Modality and Conditionals

Richmond H. Thomason
Philosophy Department
University of Michigan

rthomaso@umich.edu
http://www.eecs.umich.edu/~rthomaso/

February 8, 2010

1. Introduction

An approach to modality, conditionals, and their interpretation due to David Lewis and Angelika Kratzer has gained widespread acceptance in linguistic circles. This may be due in part to the fact that few competitors have been presented in the linguistic literature, and in part to the good ideas that the approach incorporates. In particular, I will mention three of these good ideas.

1. The approach uses possible worlds semantics.

2. Modals combine with conditionals to produce conditional modalities: constructions whose meanings are problematic in some ways, but which, one would hope, would somehow depend systematically and uniformly on the meaning of the conditional and the particular modality with which it combines. The Lewis-Kratzer account provides a theory of this.

3. in accounting for the meaning of both modals and conditionals it is useful to use relations of relative nearness among worlds. The Lewis-Kratzer account uses such relations, and, while making minimal assumptions about the properties of nearness relations, provides uniform satisfaction conditions for modalities, conditionals, and combined constructions.

But the approach has some disturbing features that, to me at least, suggest that it needs competitors and a thorough reconsideration. I will mention four of these features.

*I owe thanks to Craige Roberts to several conversations about this topic while the ideas were in an early stage of development.
1. Some things that the theory has been claimed to explain it doesn’t.

2. The satisfaction conditions that go along with this approach enforce a logic of the conditional that is controversial and has many competitors.

3. The theory does not agree well with logical work on modalities and conditional modalities.

4. The theory provides satisfaction conditions for some conditional+modal constructions that conflict with the semantic evidence about these constructions.

2. The Lewis-Kratzer semantics for modals and conditionals

The problem of “counterfactual conditionals” emerged as a problem in analytic philosophy during the first half of the 20th century.\(^1\) The development and acceptance of possible worlds semantics for modal logic allowed this problem to be reinterpreted as a logical challenge, taking the following form: provide plausible satisfaction conditions for “counterfactual” or “subjunctive” conditionals.

Two independent solutions to this logical problem appeared in the late 1960s: one due to Robert Stalnaker and the other to David Lewis.\(^2\) Both theories used orderings over worlds and appealed to the idea that a conditional \(\phi > \psi\) is true if and only if \(\psi\) is true in the closest worlds satisfying \(\psi\); but they did this in very different ways.

Stalnaker’s account is not important for the Lewis-Kratzer theory, and we can ignore it. Logical languages for modality or conditionals without quantification are interpreted with respect to a frame or modal structure \(S\) and a model \(M\) on \(S\). A Lewis frame is a pair \(S = (W, <)\), where \(W\) is a nonempty set and, for each \(w \in W\), \(<_w\) is a preorder over \(W\) with minimal element \(w\).


**Definition 2.1.** Lewis satisfaction for >.

Let \(M\) be a model on a Lewis frame \((W, <)\), let \(w \in W\) and let \(S = \{u \mid w <_w u\}\). Then:

\[M, w \models \phi > \psi \text{ iff either } S_w = \emptyset \text{ or there is a } u \in S_w \text{ such that } M \models \phi \rightarrow \psi \text{ for all } v < u.\]

Typically, satisfaction conditions in modal logic for statements that involve some sort of necessity are simple universal conditions. That is, the translation into a first-order logic

---

\(^1\)I don’t know of an extended historical survey of the literature on conditionals from this period; but see [Bennett, 2003, Chapter 20]. Two important representative works are [Goodman, 1947] and [Goodman, 1947].

\(^2\)See [Stalnaker, 1968], [Stalnaker and Thomason, 1970], and [Lewis, 1973].

\(^3\)When I’m discussing the logic of conditionals, I’ll use the notation ‘\(>\)’ as a logical primitive representing a sort of generic conditional. This conditional would figure in the logical translations of many of the conditional constructions found in natural languages. In effect, I’m assuming that we can think sensibly about the generic interpretation of natural language conditionals, or of some common sorts of natural language conditionals, without worrying about considerations having to do with mood and tense. This assumption is commonly made in the linguistic and the logical literature. Of course, there are some risks in ignoring mood and tense, but it simplifies things enormously.

\(^4\)A preorder is a transitive, trichotomous relation over \(W\). A relation \(<\) over \(W\) trichotomous if for a all \(u, v \in W\), \(u < v\) or \(v < u\) or \(v = u\).
with explicit relations over possible worlds of the satisfaction condition for $\square \chi$, where $\chi$ is modal-free, has the form $\forall x \chi$, where $\chi$ is quantifier-free. Lewis has replaced this with a more complicated $\exists \forall$ first-order satisfaction condition.

The reason for the added complexity is Lewis’ worry about the limit condition. Intuitively, it is more natural to say that the truth of $\phi > \psi$ at $w$ has to do with the closest worlds to $w$ that satisfy $\psi$. But there may be no closest worlds, even if there are many close worlds that satisfy $\phi$. Lewis cites the example ‘If I were taller than I am . . .’, and uses this to motivate the thought that it would be wrong to postulate the limit assumption. For Lewis, then, the limit assumption problem leads to an iterated-quantifier satisfaction condition.

Kratzer alters Lewis’ ideas about satisfaction conditions by relaxing constraints on the underlying ordering relation, complicating the iterated-quantifier satisfaction condition, and generalizing the approach to modals and combinations of modals with conditionals. The ideas can be found in [Kratzer, 1977], [Kratzer, 1979], [Kratzer, 1981], and [Kratzer, 1991]; they are most clearly expressed in [Kratzer, 1991].

Kratzer gives up Lewis’ trichotomy constraint on the ordering relation. She wishes to provide a more or less uniform semantics for modals and conditionals, and doesn’t think that trichotomy is appropriate for some modalities. In the case of deontic modality, for instance, she thinks that the ordering can be determined by an background set of imperatives, with $u < v$ iff every applicable imperative that is satisfied in $v$ is also satisfied in $u$. Cases in which there are contradictions between some applicable imperatives provide counterexamples to trichotomy.

Without trichotomy, Lewis’ 2-quantifier condition gives anomalous results, failing to validate the inference from $\phi > \psi$ and $\phi \rightarrow \chi$ to $\phi \rightarrow (\psi \land \chi)$ for conditionals, and from $O\phi$ and $O\psi$ to $O(\psi \land \psi)$ for modal operators. Her solution to such problems is to retreat to a 3-quantifier satisfaction condition.

For conditionals, the satisfaction condition can be stated as follows.

**Definition 2.2.** Kratzer satisfaction for $>$. Let $M$ be a model on a Kratzer frame $\langle W, < \rangle$ and let $w \in W$. Then:

\[
M, w \models \phi > \psi \text{ iff for all } t \in W \text{ such that } M, t \models \phi \text{ there is a } u \in W \text{ such that } u <_w t \text{ and } M, u \models \phi \text{ such that for all } v \in W \text{ such that } v <_w t, M, v \models \phi \rightarrow \psi.
\]

Kratzer’s satisfaction condition for deontic $O$ is analogous.

**Definition 2.3.** Kratzer satisfaction for $O$. Let $M$ be a model on a Kratzer frame $\langle W, < \rangle$ and let $w \in W$. Then:

\[
M, w \models O\phi \text{ iff for all } t \in W \text{ such that } M, t \models \phi \text{ there is a } u \in W \text{ such that } u <_w t \text{ and } M, u \models \phi \text{ such that for all } v \in W \text{ such that } v <_w t, M, v \models \phi.
\]

3. Lewis-Kratzer conditional modality

This section is a sketch.

Take a simple conditional ought sentence, like

(3.1) *If it’s raining, you ought to take an umbrella.*
Let \( \phi \) be a logical representation of ‘It’s raining’ and \( \psi \) be a logical representation of ‘You take an umbrella’. On Kratzer’s theory, the logical target of (3.1) would have the following satisfaction conditions at world \( w \), where \( <_O^u \) is the ordering relation associated with \( O \) in world \( u \).

\[
(3.2) \text{ for all } t \in W \text{ such that } M, t \models \phi \text{ there is a } u \in W \text{ such that } u <_O^w t \text{ and } M, u \models \phi \text{ such that for all } v \in W \text{ such that } v <_O^w t, M, v \models \phi.
\]

This suggests that the contribution of the conditional+protasis to the combination is restriction of the quantifiers in the satisfaction condition to worlds satisfying the protasis. The contribution of \( O \) is a particular ordering relation. This idea can be generalized to provide a uniform account of how the conditional interacts with modals.

4. Deontic logic

This section is a quick first draft.

Quite independently of the linguistic work on modals (and not mentioned, for instance, by Kratzer) there is a tradition of work in deontic logic, carried out for most part by philosophical logicians and especially (because of von Wright’s influence) by Scandinavian logicians. There is a more recent and somewhat independent literature on the topic by computational logicians, interested in applications to artificial intelligence.

Early in the deontic logic tradition it was recognized that the formalization of conditional obligation was problematic in some ways, and that this was connected with the ethical concept of reparational obligations. In [Prior, 1958], for instance, Arthur Prior formulated the Good Samaritan Paradox: you ought to help someone who is robbed, but this seems to imply that someone ought to be robbed. This is sometimes resolved by pointing out that the reparational obligation is properly formalized as a conditional obligation.

At first, I believe deontic logicians believed that the problems of conditional obligation could be solved by using the conditional, monadic \( O \), and making scope distinctions or distinctions between deontic and logical consequence. For instance, [Hintikka, 1969] falls into this category.

But it gradually came to be accepted that there was a salient and relevant interpretation of conditional obligation that could not be accounted for as a combination of \( O \) and the conditional. This led to a rather extensive literature on the logic of “dyadic deontic logic,” in which a 2-place modal operator \( O \) and the logical properties of formulas having the form \( O(\phi, \psi)^6 \) were investigated. Some papers from this literature include [Hansson, 1969], [Lewis, 1974], [van Fraassen, 1972], [Spohn, 1975], [von Kutschera, 1975], and [Prakken and Sergot, 1997]. There is a good, recent survey of work in deontic logic, including dyadic deontic logic, [McNamara, 2009]. A very recent paper, [Gabbay and Schlechta, 2010], uses sophisticated techniques from nonmonotonic logic and a variety of realistic examples to motivate a theory of conditional obligation.

As far as I know, none of the theories in the deontic logic literature—including the ones discussed in [Lewis, 1974]—use Kratzer’s account.

---

\(^6\)The most usual notation for “It ought to be that \( \psi \), given \( \phi \)” is \( O(\psi/\phi) \), in analogy to a common notation for conditional probability. I prefer to use \( O(\phi, \psi) \).
This is unacceptable. I am reluctant to believe that thoughtful, well-motivated logical theories of the truth-conditions for dyadic $O$ should conflict with linguistic evidence about the meaning of modals and their interaction with the conditionals. This provides a very good reason to examine the linguistic evidence for Kratzer’s theory.

5. The semantic evidence for Kratzer’s theory

*This section is placeholder.*

I want to say that Kratzer actually provides little or no semantic evidence for her framework, and especially for her satisfaction condition, and it is not motivated by a careful examination of inferential relations, of the sort that you find in the best logical work. The theory, in fact, is presented without much motivation. Once it is presented, Kratzer indicates some things that it explains, the tacit assumption being that these explanations provide support for the theory. Alternative theories are not mentioned, and Kratzer does not cite the work in deontic logic.

6. Explanatory problems

*This section is a quick first draft.*

Some of the explanations that Kratzer provides are flawed. I’ll just give one example here. She motivates the use of orderings in her semantics by arguing that this accounts for certain modifications of modalities. For instance, in [Kratzer, 1991], she says that you can account for the meanings of sentences involving qualified modality using different quantifier alterations. In particular, ‘There is a good possibility in $w$ that $\phi$’ is supposed to correspond to

$$\exists u \in f(w) \forall v \in f(v) [v \leq_{g(w)} u \rightarrow v \in [\phi]].$$

According to this idea, there is a good possibility that I’ll win a lottery if I hold a ticket. (The worlds are equally likely here. So in a world where my ticket wins, it wins in all more likely worlds.)

I want to conclude that the motivation for the details of the theory is weak or nonexistent. The main support comes from the fact that the theory provides a unified, more or less compositional account of modals, conditionals, and their interaction. The support falls away, then, if there are alternative theories that do this.

7. A deontic problem

*This section is a quick first draft.*

Since I am concentrating on linguistic ‘should’ and deontic $O$, I want to explain why Kratzer’s theory is insufficiently flexible in this particular case.

Suppose there are just 2 worlds, and you are offered even odds on a bet; you are paid $100 if $p$ is true, and you pay $100 if $p$ is false. You ask yourself “should I bet on $p$?” Here,
this practical use of ‘should’ is true if and only if \( p \) is more likely than \( \overline{p} \). There is no way to order the two worlds to get this result.

This example depends on two things: (1) in some cases, ‘should \( \phi \)’ takes on an interpretation that is equivalent to ‘\( \phi \) is more likely than not’. (2) Whether \( \phi \) is more likely than not depends on the probability mass of the worlds in which \( \phi \) is satisfied. This cannot be accounted for by a condition expressible by a first-order translatable satisfaction condition. No iteration of quantifiers, including Kratzer’s will accomplish this.

### 8. Maximum flexibility

In the logical literature the semantics of the conditional is still controversial, and as we saw, we need to accommodate probabilistic interpretations of modals. The probabilistic cases extend to modal-conditional combinations.

A conditional probability is a number, not a truth-value, but even so the treatment of conditional probability is instructive. Where \( \phi \) and \( \psi \) are formulas, \( P(\psi, \phi) \) is an expression representing the probability of \( \phi \) conditional on \( \psi \). Neither absolute nor conditional probability can be interpreted using an accessibility relation or an ordering over worlds. Instead, where \( W \) is the set of worlds, you need a measure \( \mu \) defined on a subset of the power set of \( W \)—a set which would be closed under boolean operations and any other operation definable in the language—and returning a real number in \([0, 1]\) for each set in its domain. Where \([\phi]\) is the set of worlds where \( \phi \) is true, \( P(\phi) \) is the \( \mu([\phi]) \), and \( P(\psi, \phi) = \frac{P(\phi \land \psi)}{P(\phi)} \), supposing \( P(\phi) \neq 0 \).

I don’t believe that the deontic logic community has ever settled on a single satisfaction condition for \( O(\psi, \phi) \). For that reason, I think it’s wisest to begin with the most general semantics for \( O(\phi, \psi) \) that the possible worlds framework will allow. If a more specialized semantics is wanted for some particular modal conditional, or some particular use of a modal+conditional, it can be treated as a special case.

This means—since the arguments of conditional \( O \) are formulas and correspond to propositions (sets of worlds)—that dyadic \( O \), expressed in English as a combination of ‘if’ and ‘ought’, must be fundamentally relational, and is not to be defined in terms of monadic \( O \). We can obtain the unconditional from the conditional \( O \) by conditionalizing on the necessarily true proposition.

If we can tell a plausible story about the interaction between unconditional and conditional \( O \) at this very general level, then we should be able to adopt this story to any modal or modal-like locution, and to most or maybe even all semantic theories of these locutions that use possible worlds. That is my hope, and this is the plan I will pursue.

### 9. Implicit deontic conditionality

The changes I’ll recommend in how to approach conditionality will look pretty sweeping when we’re done, but if we begin with a simple, central problem the main idea can be made to look pretty plausible.

Here’s the problem. Statements of obligation using ‘ought’ are often unconditional: ‘You ought to move that piece’. But we readily understand the conditional form of these
statements—‘If you touched that piece, you ought to move it’—even though (let us agree) we can’t understand this in terms of simple, semantically interpreted, syntactic complexity. It is not a matter of \( O(\phi, \psi) \) or \( \phi > O\psi \). On the other hand, there is a strong feeling that the absolute or unconditional form is somehow more basic, and the conditional form should somehow be derived from it.

If you see the problem this way, it’s a matter of how to derive a 2-place relation from a 1-place relation or property—something that, except in the most trivial cases, should be logically impossible.

But now the problem looks very similar to the one Hans Kamp addressed in his “Two theories of adjectives,” where he wondered how to derive the interpretation of the comparative form of an adjective from the interpretation of the absolute form. I think that Kamp’s idea—that the absolute forms get their interpretations by a sort of decontextualization of the absolute forms’ meanings—can be adapted to our present needs.

Kamp’s account turns on the fact that an adjective like ‘warm’ depends on a contextually assumed standard, so that ‘warm day’ can take on one temperature range in summer and another in winter.

The analogous contextualization for statements of obligation is a form of what Craige Roberts calls modal subordination.⁷ Adapting one of her examples, the phenomenon is illustrated by discourses like this.

\[
\begin{align*}
\text{If Edna’s bird feeder gets emptied, the birds will go hungry.} \\
\text{She ought to feed it.}
\end{align*}
\]

Adapting an example that I found in an insurance company’s web pages (with ‘should’ substituted for imperatives), we get this.

\[
\begin{align*}
\text{Customer care in care in case of accident.} \\
\text{You should try to stay calm.} \\
\text{You should not admit responsibility.} \\
\text{You should not discuss the scope of your insurance coverage.} \\
\text{You should not leave the scene of an accident.}
\end{align*}
\]

I take this evidence to support the idea that monadic ‘ought’ is context-sensitive, not only in other ways that have been motivated in the literature, such as with respect to the applicable norms, but also to an implicit condition. In this dimension of their context-sensitivity, ‘ought’ statements are hypothetical or enthymematic or implicitly conditional.

The idea would be that a context \( c \) supplies a deontic condition \( DC(c) \), which is simply a proposition. The semantics for monadic \( O \) interprets it relationally: \( [O]_M \) has type (\( \langle \langle s, st \rangle, \langle s, t \rangle, t \rangle \)). A model \( M \) satisfies \( O\psi \) in context \( c \), \( M, c \models O\psi \), if and only if \( [O]_M([\psi]_M, DC(c)) = T \).

In default contexts \( c \), we can assume that \( DC(c) \) is the absolutely necessary proposition \( P_T \) that is true in all worlds. If we adopt the natural equivalence of \( O\psi \) and \( O(\psi, P_T) \), this

⁷References to be supplied.
means that in the default case, context-sensitive $O$ is equivalent to simple monadic $O$. We can treat the other cases as somehow marked, and indeed they typically require some sort of explicit linguistic marking in English.

10. The general theory

This section is a rough sketch.

The idea is to give all sentences of the source language a conditional type. The conditional is treated as a decontextualizer; but to account for iterated conditionals the decontextualization needs to be followed up with a recontextualization.

I’ll try to illustrate how the approach would deal with (1) simple nonmodal, nonconditional sentences, (2) simple sentences with modal ‘should’, (3) simple conditional sentences, and (4) combined modal and conditional sentences. We can illustrate these patterns as follows. To simplify things, I’ll ignore tense and just use the infinitive form in the main clause of the conditional.

(10.1) It’s raining.
(10.2) You take an umbrella.
(10.3) You should take an umbrella.
(10.4) If it’s raining you take an umbrella.
(10.5) If it’s raining you should take an umbrella.

Let $p$ be the type $\langle s, t \rangle$ of (simple) propositions, and $cp$ be the type $\langle p, p \rangle$ of (implicitly) conditional propositions. In this theory, (10.1) and (10.2) receive type $cp$. We produce their logical formalizations by compositionally producing formulas $\phi$ and $\psi$ of type $p$ corresponding to (10.1) and (10.2), respectively. Implicit conditionals are interpreted using the context. It makes the formalization clearer to assume that these contain a special variable $z$ of type $p$ that is bound to a specific proposition by the context. Then the formulas corresponding to (10.1) and (10.2) are given by the following formulas.

(10.6) $z > \phi$
(10.7) $z > \psi$

The modal ‘should’ is translated by the dyadic operator $O$, which retains the type $\langle p, p \rangle$. The formalization of (10.3) is then given by (10.8).

(10.8) $O(z, \phi)$

The conditional is translated by $>$, so that (10.4) has the following formalization.

(10.9) $z > [\phi > \psi]$

Finally, the conditional+modal construction restores the dyadic modal by abstracting on the contextual variable $z$, so that the formalization of (10.5) looks like this.

(10.10) $z > \lambda z [O(z, \psi)] \phi$

And (10,10), of course, is equivalent to (10.11).

(10.11) $z > O(\phi, \psi)$
This approach manages, I think, to give a uniform account of how the conditional interacts with modals; it exchanges their implicit condition for the explicit condition supplied by the protasis. This, too, is what the conditional does for nonmodal sentences like (10.2)—but in this case, the result is a simple conditional. The price that is paid for this uniformity is the contextual apparatus and the retyping of sentences from $p$ to $sfcp$. But we know how to do this sort of retyping, and the apparatus is independently motivated by the phenomenon of modal subordination. The framework does not commit us to any particular model theoretic semantics for either conditionals or modals. But I have argued that this is an advantage.

Bibliography


