A matter of trust: Dynamic attitudes in epistemic logic

Rodenhäuser, L.B.

Citation for published version (APA):
Chapter 6.
Epistemic Modals

The work of this chapter begins with a change of perspective, as it introduces a second interpretation of the notion of a dynamic attitude.

So far in this dissertation, we have understood dynamic attitudes as formal correlates of reliability assessments—capturing in what way, and to what extent, a speaker trusts or distrusts a source. Reliability assessments are, conceptually speaking, part of the epistemic state of an agent. The overall idea we have pursued was that the agent uses her reliability assessment towards a particular source to determine how to change her epistemic state when receiving information from that source. The aspects of the epistemic state of an agent that we have been concerned with were thus, in the single-agent case, (1) her beliefs about the world, modeled by a (single-agent) plausibility order, and (2) her attitudes towards sources, modeled by dynamic attitudes. The informational inputs that trigger belief changes, on the other hand, were understood as representable by means of their propositional content, as sets of possible worlds.

In this chapter, we shall view dynamic attitudes not as part of the epistemic state of an agent, but as part of the informational input received by the agent. From this changed perspective, the input received by the agent is not simply a proposition, but a proposition embedded under an operator. Our proposal is that an upgrade \( \tau P \) given by applying a dynamic attitude \( \tau \) to a proposition \( P \) can be used to give a semantics for epistemic modal sentences of the form \( \text{modal} (p) \), as in “it might be the case that \( p \)”. Let me motivate the perspective change further with an example. Recall first the set of scenarios discussed in the introduction of this dissertation, repeated below. We have considered various ways in which I may receive the information that there are tigers in the Amazon jungle:

\[ \text{In the multi-agent setting we have studied in §§2.7-2.8 and §5.6 the single-agent plausibility order lives inside a "bigger" structure, a multi-agent plausibility order, also representing the agent’s information about other agents.} \]
(1)  

a. I read a somewhat sensationalist coverage in the yellow press claiming this.

b. I read a serious article in a serious newspaper claiming this.

c. I read the Brazilian government officially announcing that tigers have been discovered in the Amazon area.

d. I see a documentary on TV claiming to show tigers in the Amazon jungle.

e. I read an article in Nature by a famous zoologist reporting of tigers there.

f. I travel to the Amazon jungle, and see the tigers.

Examples of this kind have been used to motivate the concept of a dynamic attitude (cf. the introduction). The information that there are tigers in the Amazon jungle may be received from a variety of sources; and since we trust these sources to varying degrees, our “epistemic response”, i.e., the way we change our epistemic state on receiving the information that there are tigers in the Amazon jungle, will differ depending on the particular source.

Let us now consider a different type of scenario. Suppose that I talk to a trusted friend about wildlife in the Amazon jungle. Consider six variants of what my friend might tell me:

(2)  

a. There might be tigers in the Amazon jungle.

b. There could be tigers in the Amazon jungle.

c. There may be tigers in the Amazon jungle.

d. There should be tigers in the Amazon jungle.

e. There must be tigers in the Amazon jungle.

f. There are tigers in the Amazon jungle.

In all six cases, I receive information about tigers in the Amazon jungle. Also, in all six cases, I receive information about tigers in the Amazon jungle from the same source, my trusted friend. But in all six cases, the way I change my epistemic state on receiving that information from that source seems to be different. Since the source of information is the same in each case, the fact that the information change is different would seem to be due to the fact that the information received is different: the epistemic modal auxiliary (might, could, should, etc) modulates the “information uptake” on the side of the recipient.
As a result, depending on which modal is used, I will adopt different stances towards the proposition that there are tigers in the Amazon jungle.

For a preliminary diagnosis of the flavour of these different epistemic stances, a first observation to make is that the assertions lower in the list tend to make the assertions higher in the list redundant. We can see this by composing pairs of the statements: in (a) below, the second assertion is informative after the first; but not so in (b):

(3) a. There may be tigers in the Amazon jungle. In fact, there are tigers in the Amazon jungle!

b. There are tigers in the Amazon jungle. ?In fact, there may be tigers in the Amazon jungle!

Also, consider this variation on the data in (2), pointing in the same direction:

(4) a. Scientists found out that there may be tigers in the Amazon jungle. Then, they found out that there are tigers in the Amazon jungle.

b. Scientists found out that there are tigers in the Amazon jungle. ?Then, they found out that there may be tigers in the Amazon jungle.

The sequence in (a) reports progress; the sequence in (b) is hard to make sense of.

Another observation is that I may cite my friend’s assertion that there are tigers in the Amazon jungle to justify my belief that there are tigers there. But it would be highly odd to cite his assertion that there may be tigers in the Amazon jungle in the same way:

(5) a. I believe that there are tigers in the Amazon jungle, because my friend told me that there are tigers in the Amazon jungle.

b. ?I believe that there are tigers in the Amazon jungle, because my friend told me that there may be tigers in the Amazon jungle.

The assertion that there are tigers in the Amazon jungle is thus stronger (in some relevant, not yet precise sense) than the assertion that there may be tigers in the Amazon jungle; and this is reflected in the change in my epistemic state induced by integrating each sentence into my epistemic state.

As usual, the judgement diacritic “?” indicates contextual infelicity of the sentence labeled with it.
So epistemic modality in natural language and assessments of reliability as we have studied them so far share an important property: a notion of strength seems to apply to them in a similar way.

Another upshot of the discussion so far is that, in natural language use, information change is not simply induced by transmitting propositions from source to recipient. Rather, we may distinguish between different ways in which the same proposition may be embedded under a modal operator. Intuitively, the modal operator seems to modulate the “force” with which the epistemic state of the recipient is changed.

Phrasing this in terms of our formal setting, the picture is then that epistemic states of single agents (recipients of information) may still be captured by (single-agent) plausibility orders; however, the role assumed by dynamic attitudes changes. Given a plausibility order $S$, an upgrade $\tau P$ will take $S$ to some new order $S^{\tau P}$. The suggestion is now, as indicated above, that $\tau$ may figure as the semantic correlate of some epistemic modal. So dynamic attitudes can be seen as supplying a semantics for epistemic modals.

This picture will be immediately recognizable for anyone familiar with the dynamic semantics literature. There, the idea is that “knowing the meaning of a sentence is knowing the change it brings about in the cognitive state of anyone who wants to incorporate the information conveyed by it.” (Veltman 2005) Here, we study a specific type of sentence, which we assume to have the logical form

$$\text{MODAL}(p),$$

with $\text{MODAL}$ some modal auxiliary, and $p$ a sentence embedded under the former. The proposal here is that the semantics of such a sentence can be given in terms of an upgrade

$$\tau P,$$

where $\tau$ is a dynamic attitude capturing the meaning of $\text{MODAL}$, and $P$ is a proposition capturing the meaning of $p$.

The following sections develop our account in more detail. We focus here on the modal auxiliaries which have wide-spread uses with an epistemic reading. This excludes epistemic uses of would that seem to occur only in the context of the so-called “epistemic would equatives” (Ward, Birner, and Kaplan 2003):

(6) a. Who is the man with the microphone?

b. That would be the Dean.
It also excludes the modal auxiliary *can* which is generally assumed to only admit non-epistemic readings (Hacquard 2011), expressing, for example, an ability:

(7) a. John can do the job.
    b. Mary can lift 200 pounds.

We are then left with *must, may, might, should* and *could*. We shall, as a rule, ignore other readings of these modals (for example, deontic readings of *should*), as they fall outside of the scope of the framework developed in this dissertation.

Besides the modals themselves, the question which semantics should be assigned to non-modalized indicatives is pertinent, based on the hypothesis that epistemic modal statement live on an ascending scale of certainty which is topped by categorical, flat-out statements. In other words (as we shall argue below): modalized claims induce a weaker form of acceptance in the hearer than a categorical claim that something is the case. While in its general outlines, this assumption is not a matter of debate in the literature, particular details are controversial, and we will discuss those below.

Our aim is thus to provide a semantics for sentences of the form

\[
\text{ACCEPT}(p)
\]

in terms of upgrades \(\tau P\), where \(\tau\) is a dynamic attitude understood as interpreting

\[
\text{ACCEPT} \in \{\text{is, must, should, may, could, might}\},
\]

and \(P\) is a proposition interpreting what is traditionally called the “prejacent”, the sentence \(p\) embedded under the modal.

We shall generally assume that a “typical” sentence \(p\) is satisfied in all and only the \(P\)-worlds (same character, upper case), where the worlds in \(P\) are drawn from some arbitrary but fixed set of possible worlds \(W\).

The following analysis tries to accumulate empirical and formal arguments, so that in the end, operations can be identified that plausibly capture the meaning of the members of \(\text{ACCEPT}\). In the end, we do not actually arrive at one proposal, but at a selection of alternative ones. This underlines the exploratory nature of this chapter, whose main point is to make the case that the setting developed in this dissertation may provide an interesting toolbox for formal semantics. We will approach the issue from three directions. §6.1 contrasts our approach with two main approaches in the literature, the proposal that epistemic modals express *degrees of commitment* on behalf of the
speaker, and the proposal that epistemic modals are contextually restricted quantifiers. §6.2 adopts a “bottom-up” perspective, identifying distinctions between the various “acts of acceptance” that epistemic modals correspond to on our account; and §6.3 proceeds in a more “top-down” fashion, by considering intuitions about what all acts of acceptance considered here have in common.

6.1. Epistemic Modals as Dynamic Operators

A key intuition voiced above as well as in the literature is that different modal expressions carry a distinct modal force. On the view developed here, epistemic modals correspond to types of transformations of plausibility orders, with the latter representing the epistemic state of an agent processing a modalized sentence. The central claim is the following: epistemic modals transform the hearer’s plausibility assessment of the prejacent. The force of a modal is then characterized by the extent to which accepting a modal sentence modal\( \tau(p) \) transforms the recipient’s information state. It should be already clear at this point that the framework developed in the preceding chapters lends itself well to arriving at a formal conception of modal force. In fact, our setting offers two perspectives on the notion. First, the force of a modal \( \tau \) can be understood in terms of the fixed point \( \bar{\tau} \) of \( \tau \) (cf. §1.7): a modal \( \sigma \) is “more forceful” than another modal \( \tau \) if the fixed point of \( \sigma \) entails the fixed point of \( \tau \), which is to say that whenever \( \sigma P \) is satisfied, then so is \( \tau P \). But we can also understand modal force directly in terms of the transformations themselves: \( \sigma \) is “more forceful” than \( \tau \) iff \( \sigma \) subsumes \( \tau \) (cf. §1.8). This means that after applying an upgrade \( \sigma P \), applying an upgrade \( \tau P \) is redundant (for example: after accepting that John must be in London, hearing that John might be in London provides no new information). By Theorem 11 both perspectives amount to the same thing.

The idea of cashing in the notion of modal force in terms of one transformation making another redundant seems very plausible (in the introduction to this chapter, we have already used this idea implicitly). The idea is also inherent in previous work on modality in dynamic frameworks, even though previous research has generally tended to (1) solely focus on the epistemic modals must and might (Veltman 1996, Willer 2013, forthcoming), and (2) study modals as dynamic operators, without making the connection to propositional attitudes explicit.

Before developing the view in more formal detail, this section argues that the position defended here has empirical bite, and allows us to make sense of
certain empirical phenomena that otherwise remain obscure. I will do so by contrasting the proposal with two important perspectives on epistemic modality in the literature: the view (popular in the descriptive linguistics tradition) according to which epistemic modals express degrees of commitment, and the view (predominant in formal semantics) according to which epistemic modals are contextually restricted quantifiers.

6.1.1. Degrees of Commitment. According to a wide-spread view, “epistemic modality in natural language marks the degree and/or source of the speaker’s commitment to the embedded proposition.”

This proposal suggests, for example, that a speaker’s communicating that there may be tigers in the Amazon jungle expresses a weaker commitment on behalf of the speaker than a speaker’s saying, flat-out, that there are tigers in the Amazon jungle. Epistemic modals are thus assigned a clear job description: they are to be used to communicate degrees of (epistemic) commitment. This proposal is very attractive in that it provides us with an immediate explanation of modals having particular forces lying on a scale: the force of a modal simply resides in the degree of commitment of the speaker that the modal expresses. Consequently, the strength relations between modals derive from the fact that expressing a particular degree of strength is, essentially, what modals are supposed to do.

The view has problems, however (cf. Papafragou (2006) for more discussion). The following issue is the central one for our purposes. Consider the following example. Bob is watching the second season of *Homeland*. Alice has already seen all episodes, but is keeping Bob company. At some point, the following dialogue unfolds:

1. **Bob:** Could Galvez be a mole?
2. **Alice:** He could be.

Let us assume that the commitment view on epistemic modals is correct. Then the first sentence uttered by Alice reveals, it seems safe to say, some rather weak form of commitment, on Alice’s behalf, to the fact that Galvez is a mole. One gets the impression that she regards it as a remote possibility that Galvez is a mole, at the very least: a possibility not to be excluded.

Actually, however, this is only part of what Alice says, as I have omitted the last bit of the dialogue. In fact, the dialogue runs as follows:

---

3This quote is taken from Papafragou (2006), who, however, does not endorse this view, but merely reports it, and is actually critical of it. As cited by Papafragou, the view is endorsed, for example, by Halliday (1970), Palmer (1990) and Bybee and Fleischman (1995).
a. *Bob:* Could Galvez be a mole?

b. *Alice:* He could be. But as a matter of fact, he isn’t.

Continuing to assume that the commitment view is correct, we are now in a bit of trouble. Let’s see why. Clearly, the second sentence uttered by Alice reveals that she is committed to Galvez *not* being a mole. Combining this with the previous observation that Alice regards it as a remote plausibility that Galvez is a mole, Alice is contradicting herself. However, I find no reason to take issue with what Alice says: it does not seem to me that she is contradicting herself. And this indicates that the commitment view is wrong.

Notice also that the coherence of Alice’s order is sensitive to the order. The following reply to Bob’s question is marked (i.e., stands out as unusual and uncommon, “feels wrong”) and sounds rather incoherent:

\((10)\) *Alice:* He isn’t. ??But as a matter of fact he could be.

The following variants do not sound better:

\((11)\) a. *Alice:* He isn’t. ?But he could be.

b. *Alice:* As a matter of fact he isn’t. ?But he could be.

The speaker commitment approach provides no clue why this should be so; we will return to this issue below.

6.1.2. **MODALS AS RESTRICTED QUANTIFIERS.** Another perspective on epistemic modals is provided by the predominant view in theoretical linguistics originating in Kratzer’s work \((\text{Kratzer 1981, 2012})\). On this view, epistemic modals carry truth-conditional content on the one hand, but are context-dependent on the other hand.\(^4\) The context, essentially, provides an ordering on a set of possible worlds, which in turn is derived from a number of propositions (a “premise set”). For an example, let such an order \(\leq\) be given (usually, \(\leq\) is assumed to be a reflexive and transitive but not necessarily total relation: some worlds may be incomparable). The main idea is then that a sentence like *there may be tigers in the Amazon jungle* is true iff, among the *best* worlds in \(\leq\) (the worlds \(w\) such that we can find no \(v\) with \(v < w\)), some worlds are worlds in which there are tigers in the Amazon jungle. So *may* functions

\(^4\)In fact, Kratzer’s account provides a general framework for dealing not only with epistemic, but with any kind of modality in natural language. But for present purposes, it is solely the epistemic readings of modals that matter, given our limited scope in this chapter.
as a “restricted quantifier”, i.e., an existential quantifier quantifying over the best worlds in the order.

The account assumes that in linguistic practice, context supplies a set of premises from which the order \( \leq \) can be “projected”. For example, in a certain context, it may be the current state of scientific knowledge about the Amazon jungle that figures as a conversational background, and the above sentence would then be paraphrasable as:

(12) In view of the current state of scientific knowledge, there may be tigers in the Amazon jungle.

Since the latter paraphrase suggests that the speaker herself endorses the view that there may be tigers in the Amazon jungle, it has usually been assumed that, at least in the majority of cases, the evidence of the speaker should be considered part of the conversational background (the “speaker inclusion constraint”), and, in fact, the so-called “speaker-centric” reading, in which it is solely the speaker’s evidence that matters, is sometimes taken to be the default reading of an epistemic modal sentence.

The Kratzerian account offers a simple account of modal force. Since modal sentences carry truth-conditions, it is simply truth-conditional entailment that is responsible for modal force: that \textit{must} has a stronger force than \textit{may} is due to the fact that whenever it is true that something must be the case it is also true that that same thing may be the case.

Returning to the above discussion, the first observation to be made is that the Kratzerian account can make sense of the data that turned out to be puzzling for the speaker commitment approach above. Namely, it could be argued from this perspective that Alice’s statement (that Galvez could be a mole, but as a matter of fact isn’t a mole) is contextually decoded in the following way:

(13) In view of your evidence, he could be a mole. But in view of my evidence, he isn’t.

This seems to express in a fairly intuitive way what the discourse in (9b) intuitively communicates. While Bob’s evidence does not yet exclude that Galvez is a mole, Alice is in a more privileged epistemic position. Remember that she has already seen all episodes. And since, in later episodes, it turns out that Galvez is not a mole, he is not a mole in view of Alice’s evidence. This rather straightforward account favours the more flexible Kratzerian approach over the speaker commitment view discussed earlier.
Notice, however, that the quantificational view does not explain why reversing the order would matter for the markedness of the discourse:

\[ (14) \text{ In view of my evidence, he isn’t a mole. But in view of your evidence, he could be a mole.} \]

\[ (14) \text{ is not more marked than (13), which is unexpected, given that, as observed above, (10) and (9b) differ strongly w.r.t. markedness. Since the whole idea was that (14) elicits the contextual meaning of (10), and (13) the contextual meaning of (9b), this result is unsatisfactory.} \]

Notice, on the other hand, that there is nothing problematic about the initial example from the perspective of the dynamic account sketched so far. A speaker can easily accommodate the information that Galvez is not a mole after first accepting that she could be. This happens all the time: whenever we gain information, the space of options shrinks and previously possible, or plausible options come to be impossible, or implausible. The dynamic account also explains why reversing the order should make a difference: having discarded the possibility that Galvez is a mole (after accepting Alice’s first statement), the hearer has no room anymore to accommodate the possibility that he could be a mole. So from the perspective of any hearer accepting both statements in turn, the discourse must be incoherent.\(^5\)

Furthermore, on the dynamic account, a simple pragmatic explanation for the intuition that epistemic modals reflect speaker commitments is possible. In a context where an information exchange may be presumed by the participants to proceed in a cooperative fashion in the sense of Grice, it is natural to assume that the information change induced in the hearer by accepting an epistemic modal sentence is matched by a corresponding propositional attitude of the speaker.\(^6\) In other words, if I come to believe that John must be in London based on information obtained from you, it is usually safe to assume that you, as the source of information, also believe that John must be in London.

The context considered in the *Homeland* example, however, is not a standard case of a cooperative information exchange: being as informative as possible is clearly not guaranteed to be the most helpful thing to do when watching a TV series. Rather, it is customarily regarded as spoiling the fun. In this sense, the “adversarial” strategy of revealing as little information as

\(^5\)This type of consideration was a main point of Veltman (1996)’s work on might, and is not a novelty of the current proposal.

\(^6\)In fact, this assumption is closely related to the notion of honesty discussed in an earlier chapter of this dissertation, cf. \(\S 2.8\).
possible may be the really cooperative one in such instances. In such circumstances, speakers need not be committed themselves to the epistemic modal claims they are making.\footnote{Cf. Verbrugge and Mol (2008) for a more systematic perspective on the relationship between cooperative and adversarial communication.}

Overall, we may conclude that in a standard, Gricean context, speakers will indeed tend to express degrees of commitment by epistemic modal talk, and they will aim at converting the hearer to the same degree of commitment—as predicted by the commitment view. But the dynamic account also handles deviant cases with ease.

The view advocated here has another advantage over the Kratzerian account, which I discuss next. Noticeably, on Kratzer’s view, epistemic modals report about what is compatible with, plausible from the perspective of, or entailed by the information of a group. It seems to me that this leads to conceptual problems. Consider the following example.

\begin{enumerate}
\item[(15)]
\begin{enumerate}
\item Peggy: Don must be in his office.
\item Peggy: I believe that Don is in his office.
\end{enumerate}
\end{enumerate}

Let us assume that the context supplies a speaker-centric reading for Peggy’s utterance that Don must be in his office. The rough truth-conditions for the sentence (15a) on a Kratzerian analysis would then run as follows: the sentence is true at a world $w$ iff it follows from Peggy’s evidence in $w$ that Don is in his office. Now what about (15b), where Peggy is reporting her belief that Don is in his office? It seems that the truth conditions of the sentence I believe that Don is in his office uttered by Peggy in the same context amount to the very same thing. It seems, then, that the two sentences are equivalent (in this context). However, the following is only felicitous as a reply to (15b), not to (15a):

\begin{enumerate}
\item[(16)] Yes, I know. But I don’t believe it. I have seen him in the conference room.
\end{enumerate}

This observation seems hard to make sense of from the point of view of the Kratzerian framework. The dynamic account, on the other hand, does not face a problem. Here is a sketch of an analysis: Peggy’s saying that Don must be in his office amounts, when accepted, to increasing the plausibility of worlds where Don is in his office over the plausibility of worlds where he isn’t (just in what particular way is less important for the concrete example: details follow below). On the other hand, Peggy’s saying that she believes that Don is
in his office communicates, when accepted, that worlds in which Peggy does *not* believe that Don is in his office may be excluded from consideration. And, of course, one may agree that worlds where Peggy believes that Don is not in his office may be excluded from consideration without finding it plausible that Don is in his office.

The advantages of the dynamic account that have been highlighted so far are thus, in my view, the following: (1) it allows us to preserve the plausible intuition that epistemic modals are usually related to the speaker’s beliefs in one way or another; (2) it avoids the pitfall posed by the observation that using a modal does not necessarily require a particular degree of commitment (the *Homeland* example); (3) it allows us to make sense of the order-dependence of modal discourse (as has already been observed by Veltman); (4) it allows us to draw a clear distinction between modals and reports about the propositional attitude of a speaker (the *Peggy* example).

A final point to make is that our account, with its emphasis on fixed points of dynamic transformations, can be seen as a quite conservative “remodeling” of Kratzer’s approach. In the Kratzerian tradition, epistemic modals essentially correspond to what has been called propositional attitudes in this thesis. The claim underlying my proposal—very much in the overall spirit of this dissertation—is that rather than corresponding to propositional attitudes, epistemic modals realize propositional attitudes in the hearer, as the dynamic output of accepting an epistemic modal claim.

### 6.2. The Modal Scale

A main source of evidence about epistemic modals are speaker judgements about entailment relations among modalized sentences. On balance, these judgments suggest that the group of modals we are concerned with here *live on a scale* that is topped by the force of a flat-out assertion that something *is* the case. Let us look into the matter in more detail.

#### 6.2.1. *Is* vs. *Must*

Consider the following pair of sentences:

(17) a. John must be in London.
    b. John is in London.

The conviction that (17a) has a more “tentative” flavour than (17b) has been voiced many times in the literature. Recently, von Fintel and Gillies
have challenged this position. Von Fintel and Gillies claim that saying that something must be the case is just as strong as saying that something is the case, the difference being that *must* carries an evidential signal to the extent that the claim is based on an indirect inference. And, indeed, *must*-claims cannot be based on direct inference:

(18) *Looking at the rain pouring down: ??It must be raining.*

However, the arguments put forward by von Fintel and Gillies leave open the possibility that English tends to present claims that are marked as derived from indirect evidence as “weak”. The question is what the character of this weakness should amount to. I find it quite difficult to judge whether there can be situations where it is *true* that a particular state of affairs *must* obtain, while it is false that this particular state of affairs *does* obtain (or vice versa, for that matter). That judgements of this kind are difficult may point to the fact that we are not trained to think about epistemic modals in terms of their truth conditions.

The position suggested by the account presented here is that the difference between *must* and *is* lies in the fact that claims of the former kind are marked as defeasible, while claims of the latter kind are not. Consider the following example:

(19)  

a. John must be in his office. Let’s go see if he’s there.

b. John is in his office. Let’s go see if he’s there.

If John is in his office, there is no need to check whether he is here. Suggesting such a check seems pointless; so 19b sounds strange. But 19a does not sound strange at all. In particular, it does not sound like the speaker is retreating to a weaker position, as he would be by saying:

(20) John must be in his office. If you don’t believe me, let’s go see if he’s there.

In other words: if one is convinced that something must be the case, one may still find oneself in a position where one wants to check whether that

---

8Cf. von Fintel and Gillies (2010), who extensively reference the literature in support of the “mantra”, as they call it.
9Just as a judge in a court would tend to regard an elaborate chain of inferences in support of the fact that John murdered Bill as weak compared to an eyewitness account of the actual incident of John strangling Bill.
something indeed is the case. And this seems to me to point to the fact that the “tentativeness” of must-claims derives from the fact that they are, as it were, ear-marked for defeasibility. In using such a claim, a speaker points to the fact that the claim made might be overridden by later, more conclusive evidence.

The process of marking claims as defeasible can be overridden by conclusively establishing that the prejacent of the must-claim holds:

(21) The ball is in box A, B or C. It is not in box A. It is not in box B. So it must be in box C.

After processing the first three sentences, the final sentence merely states the obvious, marking that a conclusion has been drawn: so, in this example, must is reduced to the evidential signal diagnosed by von Fintel and Gillies.

6.2.2. Must vs. Should. Next, we consider the force relation between must and should. This matter is easier to decide, as the notion that must is stronger than should does not seem to be contested in the literature. So we can proceed much more quickly here.

The observation that should can be strengthened to must, but not vice versa, provides empirical evidence that the latter is indeed stronger than the former:

(22) Where are my diamonds?
   a. They should be in the safe. In fact, they must be in the safe.
   b. They must be in the safe. ?In fact, they should be in the safe.

That a stronger claim is made by saying that something must be the case than by saying that something should be the case is also plausible in view of the parallel to deontic modality. The following pair is taken from Silk (2012):

(23) a. I should help the poor. In fact, I must.
   b. I must help the poor. ?In fact, I should.

In a recent paper, Krzyżanowska, Wenmackers, and Douven (2013) argue that must and should carry distinct evidential signals. The authors suggest that must signals the presence of an abductive inference, while should signals an inductive inference. Examples like the following two support this hypothesis:
6.2. The Modal Scale

(24) a. John has done many assignments well. He should be able to handle this one, too.

b. John has done many assignments well. He must be able to handle this one, too.

John’s past performance provides the basis for an inductive inference. The fact that (24a) is considerably less marked compared to (24b) supports the position of Krzyżanowska et al. (2013) (note that (24b) has a deontic reading on which it is quite acceptable). And conversely:

(25) a. Only John had access to the kitchen yesterday. He must have stolen the cookies.

b. Only John had access to the kitchen yesterday. He should have stolen the cookies.

The observation that access to the kitchen was restricted to John sets up an abductive inference to the extent that John is the one who took the cookies, as no other equally good explanation seems to be available.

Returning to the cookie example, we note that the conviction that John did it is easily defeated by taking further aspects into account:

(26) But perhaps the cookies were already stolen the day before.

This seems to me also the case where must is weak in the relevant sense: an information state that supports the claim that John must have stolen the cookies may easily evolve into a state in which this claim is no longer upheld.

Notice, further, that the question just which type of evidence gives naturally rise to must vs. should claims is independent of the claim that must carries stronger conviction than should. If Krzyżanowska et al. (2013)’s claim is correct, the conclusion to draw would be then that natural language treats abductive inferences as stronger than inductive ones.\footnote{It seems to me worth pondering whether the taxonomy used in Krzyżanowska et al. (2013) is comprehensive. Consider (Frank Veltman, p.c.):

(27) Normally, John does well on assignments. He should be able to pass this exam, too.

The inference from John’s usual success to the particular case of this exam does not seem to be inductive. Rather, it seems to be a default inference. Veltman (1996).

Interestingly, Veltman (1996) saw a tight link between epistemic must and default reasoning. But there are some doubts whether the following discourse is really felicitous:}
6.2.3. Should vs. May. That should is stronger than may almost goes without saying. The former is a “necessity modal”, while the latter is a “possibility modal.” An empirical test to decide to which class a given modal belongs is to check whether enumerating a number of options is acceptable with a particular modal:

\[(29)\]

\[\begin{array}{l}
\text{a. John might be in London and he might be in Paris.} \\
\text{b. John may be in London and he may be in Paris.} \\
\text{c. John could be in London and he could be in Paris.} \\
\text{d. ?John should be in London and he should be in Paris.} \\
\text{e. ?John must be in London and he must be in Paris.}
\end{array}\]

As it turns out, claiming that two inconsistent facts should or must be the case is incoherent; but not so for might, may and could. This makes them qualify as “existential”.

In view of this observation, the following data is not surprising:

\[(30)\]

\[\text{Recursion Theory is so difficult.} \]

\[\begin{array}{l}
\text{a. Don’t worry. You should be able to pass it.} \\
\text{b. Don’t worry. ?You may be able to pass it.}
\end{array}\]

The second discourse above is marked, and the explanation seems to be that learning that one may be able to pass an exam is not bound to diffuse worries of failure. On the other hand, learning that one should be able to pass seems to instill at least some confidence in the hearer. That is: the should-claim gives rise to a higher “degree of acceptance” of the fact that one will pass the exam than the corresponding may-claim.

6.2.4. May vs. Could vs. Might. We turn to the trio of “existential modals”: may, could and might. Here, the situation is again a bit more complicated. My own feeling is that may is stronger than could, which is in turn stronger than might. But finding empirical judgements that unequivocally support this position is difficult.

\[(28)\]

\[\text{Normally, John does well on assignments. ?He must be able to pass this exam, too.}\]

\[\text{So is perhaps should the epistemic modal that goes best with default reasoning? I will leave the matter unresolved.}\]
6.2. The Modal Scale

The following is the best I can come up with. Which of the following stories is more scary?[12]

(31)  a. A wolf might come in. It would eat you first.
    b. A wolf could come in. It would eat you first.
    c. A wolf may come in. It would eat you first.

My sense is that they are increasingly scary from (a.) to (c.). To me, (a.) seems to talk about an “abstract possibility”, while (c.) points to a “real option”. And (b.) lies somewhere inbetween. I draw the tentative conclusion that may is indeed stronger than could, which is in turn stronger than might. This account may need to be revised, which makes the following account a bit tentative.

Summarizing the preceding discussion, we arrive at the following picture:

Is < Must < Should < May < Could < Might

Under the assumption that the order < is transitive, we conclude that our five modals lie on a linearly ordered strength scale, topped by categorical claims that something is the case.[13] So we have entailment from top to bottom in the following list:[14]

(32)  a. John is in his office.
    b. John must be in his office.
    c. John should be in his office.
    d. John may be in his office.
    e. John could be in his office.
    f. John might be in his office.

[12]Frank Veltman (p.c) tells me that the third discourse is infelicitous, and needs to be replaced by “A wolf may come in. It will eat you first.” More empirical investigation seems needed to resolve this matter.

[13]With the proviso that, as admitted above, the evidence about the existential modals I have presented is rather sketchy.

[14]Not everyone is fully happy with this picture. Frank Veltman (p.c) thinks that, while may is stronger than might, the two modals could and might are really equally strong in their epistemic use. This would correspond to the following picture:

Is < Must < Should < May < Could ≈ Might

I will have to leave the matter unresolved, pointing out once more the tentative flavour of my analysis of the existential modals.
6.3. Acts of Acceptance

We move towards a more technical account. The purpose of the previous section was to argue that the modal auxiliaries lie on a strength scale that is topped by categorical claims that something is the case. In this section, we try to extract constraints on the dynamic attitudes that could in principle be seen as capturing “acts of acceptance” in the relevant sense of providing a semantics for epistemic modal claims. Compared to the previous section, the approach is rather top-down in fashion, as it looks at what the auxiliary verbs we consider have in common. I will work with the simplifying assumption that our target class of five modals (plus the copula *is*) is representative for epistemic modals more generally. So we try to extrapolate properties of acts of acceptance from linguistic judgements about the members of our small class. We identify five such properties (strictness, triviality, affirmativity, defeasibility, and informativeness) and discuss to what extent they apply to all verbs we are considering, or just to some of them. Towards the end of the section, we evaluate what choices the analysis leaves open in terms of a suitable semantics for epistemic modals.

6.3.1. Strictness. A dynamic attitude $\tau$ is strict iff for any plausibility order $S$, and proposition $P$:

$$\text{If } P \cap S = \emptyset \text{ then } S^{\tau P} = \emptyset.$$  

More simply, but equivalently:

$$S^{\tau \emptyset} = \emptyset.$$

This property is motivated by the observation that a positive claim, embedded under a modal or not, is unacceptable after its negation has been accepted:

(33)  

a. John is not in London. ?He is in London.

b. John is not in London. ?He must be in London.

c. . . .

d. John is not in London. ?He might be in London.

6.3.2. Triviality. A dynamic attitude $\tau$ satisfies triviality iff for any plausibility order $S$ and proposition $P$:

$$\text{If } P \cap S = S \text{ then } S^{\tau P} = S.$$
6.3. Acts of Acceptance

More simply, but equivalently:

$$S^{\tau W} = S.$$ 

The motivation for this property is that a positive claim, embedded under a modal or not, is trivial (uninformative) after that same positive claim has been accepted unembedded:

(34) a. John is in London. He is in London.
    b. John is in London. He must be in London.
    c. . . .
    d. John is in London. He might be in London.

6.3.3. Affirmativity. A dynamic attitude \( \tau \) is affirmative iff for all plausibility orders \( S \), propositions \( P \) and worlds \( w, v \in S^{\tau P} \):

— if \( w \in P \) and \( v \notin P \), then \( w \leq_S v \) implies \( w \leq_{S^{\tau P}} v \), and
— if \( w \notin P \) and \( v \in P \), then \( w \leq_{S^{\tau P}} v \) implies \( w \leq_S v \), and
— if \( w \in P \) iff \( v \in P \), then \( w \leq_S v \) iff \( w \leq_{S^{\tau P}} v \).

The motivation for the first two clauses of this property is that after accepting a positive claim, embedded under a modal or not, an agent may come to regard it as more plausible that the prejacent is satisfied, but certainly not less.

(35) a. John is in London this weekend.
    b. John should be in London this weekend.
    c. John might be in London this weekend.

Accepting that John is in London this weekend, but at the same time finding it less plausible that John is in London this weekend before this very act of acceptance does seem to be extremely odd. This intuition seems to me to be deeply embedded into our understanding, not only of categorical claims, but of all epistemic modals under consideration. It also echoes the idea that the purpose of epistemic \textit{might}-sentences is to “raise the possibility” of \( P \), that has been advanced, for example, by Swanson (2006).

We have already met the third clause of the property under the name of conservation (cf. §3.7.1): accepting, for example, that John may be in London, does not give us any reason to re-evaluate the relative plausibility of two
Chapter 6. Epistemic Modals

worlds where John is in London, nor does it give us any reason to re-evaluate the relative plausibility of two worlds where John is not in London. So performing an upgrade $\tau P$ really only comes down to affirming $P$, but nothing else.

6.3.4. Informativeness. A dynamic attitude $\tau$ is informative iff for any order $S$ and $S$-substantial proposition $P$: if $\neg P <_S P$, then $S^{\tau P} \neq S$.

Above, I have argued that epistemic modals (and the copula is) are subject to an affirmativity property. However, as it is formulated, the affirmativity of $\tau$ may be trivially satisfied. Notice that, for example, neutrality $id$ (given by $S \rightarrow S$ for any proposition $P$) is affirmative! But $id$ is certainly not adequate to provide a semantics for any kind of modal! This is amended by informativeness.

Informativeness requires that, at least in a situation where all $\neg P$-worlds are strictly more plausible than all $P$-worlds, performing the upgrade $\tau P$ will change the current plausibility order $S$ in some way. In conjunction with the above affirmativeness property, this amounts to saying that epistemic modals (and is) raise the plausibility of their prejacent in a non-trivial way.

Should all epistemic modals under consideration satisfy this property? Veltman’s $might$ operator violates informativeness. We return to the issue below.

6.3.5. Defeasibility. A dynamic attitude $\tau$ is defeasible iff for any $S$ and $P$: If $P \cap S \neq \emptyset$, then $S^{\tau P} = S$.

Defeasibility of a dynamic attitude $\tau$ requires, essentially, any upgrade $\tau P$ to provide soft information only, information that may be retracted as further information comes along (cf. Chapter $\S 1$ for more detail on the notion of “soft information”). Above, I have argued that epistemic $must$ serves to mark its prejacent as representing defeasible evidence. In the previous section, our claim was that the epistemic modals under consideration are ordered by their modal force as follows:

$$\text{Is} < \text{Must} < \text{Should} < \text{May} < \text{Could} < \text{Might}$$

One thus expects that epistemic modals weaker than $must$ are also defeasible. But should a claim that something is the case be construed as defeasible? The dynamic semantics tradition says no: traditionally, the interpretation of

$^{15}$Recall the notation: given a plausibility order $S$, we write $P <_S Q$ iff for any $w, v \in S$: if $w \in P$, $v \in Q$, then $w <_S v$. Recall also from $\S 3.2.4$ that a proposition $P$ is $S$-substantial (in a plausibility order $S$) if neither $P \cap S = \emptyset$ nor $P \cap S = S$. 
6.3. Acts of Acceptance

categorical claims like “John is in London” has been given in terms of world elimination, which amounts to a violation of defeasibility in the above sense. Again, we return to the issue below.

6.3.6. Summary. I have argued that all verbs under consideration should be interpreted by means of dynamic attitudes that are strict, satisfy triviality, and are affirmative; further, all the verbs which are strictly stronger than might should satisfy informativeness; finally, modals weaker than must (and including must) should receive a defeasible interpretation. And the question whether might should satisfy informativeness has been left open, as well as the question whether categorical claims could plausibly be construed as defeasible.

As for the last two points: I think how one decides on these questions depends largely on how one draws the line between semantics and pragmatics. I will adopt what I call the “classical view” here, but also comment on what I call the “pragmatic view”, which seems promising to explore further.

6.3.7. The Classical View. Suppose one where only to commit to the position that acts of acceptance should satisfy strictness and triviality, but nothing more. Then we observe:

**Proposition 103.** For any dynamic attitude \( \tau \) satisfying strictness and triviality:

\[ ! \leq \tau \leq !^\sim . \]

**Proof.** Suppose that \( \tau \) is satisfies strictness and triviality. We show first that \( ! \leq \tau \). To show this, we have to prove that \( (S^{IP})^{\tau P} = S^{IP} \), for any plausibility order \( S \) and proposition \( P \). But observing that \( P \cap S^{IP} = S^{IP} \), by the triviality constraint, it follows that \( (S^{IP})^{\tau P} = S^{IP} \). So our claim holds, and this shows that \( ! \leq \tau \).

Next, we show that \( \tau \leq !^\sim \). To show this, we have to prove that \( (S^{\tau P})^{\Gamma P} = S^{\tau P} \), for any plausibility order \( S \) and proposition \( P \). We first observe that if \( S^{\tau P} = \emptyset \), then also \( (S^{\tau P})^{IP} = \emptyset \), and our claim holds. We may thus assume that \( S^{\tau P} \neq \emptyset \). By the fact that \( \tau \) is strict, this implies that \( S^{\tau P} \cap P \neq \emptyset \). By definition of \( !^\sim \), it follows that \( (S^{\tau P})^{\Gamma P} = S^{\tau P} \), and, again, the claim holds. This shows that \( \tau \leq !^\sim \).

Assuming we have not overlooked an act of acceptance that is stronger than the act of accepting a categorical claim, or an act of acceptance that is weaker than accepting that something might be the case, in view of the previous proposition it is but a small step to conclude that \( ! \) provides a reasonable
semantics for claims that something is the case, and $\sim$ provides a reasonable semantics for claims that something might be the case. This provides an underpinning for the “classical” position in update semantics: under our assumptions, $!$ and $\sim$ are the natural choices, and they are also the choices made in Veltman (1996).

6.3.8. The “Pragmatic” View. The pragmatic approach centers on two ideas: in using natural language to interact, we generally obtain information that is defeasible, adopting views that we may change as new evidence comes into view. Adopting such a view seems pertinent if one wants to analyze simple dialogues like the following:

(36)  a. A: John is in London.
    b. B: No, he might be in Paris, too.

Consider the perspective of a hearer witnessing this dialogue. The hearer trusts both speakers. So based on what speaker A says, the hearer first comes to accept that John is in London. But then, she comes to accept that John might be in Paris, too. So she revises her beliefs. To accomodate this sort of phenomena, interpreting acts of acceptance as defeasible, i.e., “up for revision” in general, seems useful.

The second idea is based on the intuition that might claims sometimes provide non-trivial, genuinely usable information. The following example is due to Paul Dekker:

(37)  I told you it might rain!

Here, the speaker seems to point out that his past act of telling the hearer that it might rain was meant as a warning. And since warnings are typically meant to provide genuine information—how does that square with an account of might that construes it as generally uninformative?

But the question is, of course, what we mean by “informative” here. Our formal notion of informativeness defined above formalizes the notion in terms of changing the relative plausibility hierarchy among worlds. Is that really what might does?

(38)  A wolf might come in.

Does a hearer who accepts that a wolf might come in perform some mental operation akin to advancing the plausibility of some worlds where a wolf
6.3. Acts of Acceptance

comes in? I think what is rather required of the hearer is to admit that the possibility of a wolf coming in is not beyond the conceivable.

Why then is might felt to be informative? An answer is suggested in the literature: a claim that it might be that \( P \) draws attention to the possibility that \( P \) [Groenendijk et al. 1996, Clardelli, Groenendijk, and Roelofsen 2009]. While this aspect is neither captured by our setting in general, nor by our (formal) notion of informativeness in particular, adapting the present setting in a way that could account for “attentive meaning” seems feasible, and in fact, combining the two would seem to be an interesting project for future research.

Regardless on how one decides on these matters, the question is of interest what dynamic attitudes fall out if one assumes not only affirmativeness, but also defeasibility and informativeness. Here, we observe:

**Proposition 104.** For any dynamic attitude \( \tau \) satisfying affirmativeness, defeasibility and informativeness: \( \uparrow^+ \leq \tau \leq \downarrow^{++} \).

**Proof.** Let \( \tau \) be a genuine modal. We first show that \( \uparrow^+ \leq \tau \). Let \( S \) be a plausibility order, and \( P \) a proposition. If \( S \cap P = \emptyset \), it follows that \( S^{\uparrow^+} P = \emptyset = \emptyset^{\uparrow^+} = (S^{\uparrow^+} P)^{\uparrow^+} \), and our claim holds. So we may suppose that \( S \cap P \neq \emptyset \). By definition of \( \uparrow^+ \), we have \( S^{\downarrow^+} = S \), and since \( \tau \) is defeasible, \( (S^{\uparrow^+} P)^{\tau} = S^P \). It remains to be shown that for any \( w, v \in S \): \( (w, v) \in S^{\uparrow^+} \) iff \( (w, v) \in (S^{\uparrow^+} P)^{\tau} \).

So let \( w, v \in S \). If \( w \in P \) iff \( v \in P \), the claim holds since \( \tau \) is affirmative. So suppose that \( w \in P \), \( v \notin P \). Then \( (w, v) \in S^{\uparrow^+} \) by definition of \( \uparrow^+ \), and, again since \( \tau \) is affirmative, \( (w, v) \in (S^{\uparrow^+} P)^{\tau} \), so again, the claim holds. Finally, suppose that \( w \notin P \), \( v \in P \). Then \( (w, v) \notin S^{\uparrow^+} \) by definition of \( \uparrow^+ \), and, again since \( \tau \) is affirmative, \( (w, v) \notin (S^{\uparrow^+} P)^{\tau} \), so, again, the claim holds. Hence we have shown that \( (w, v) \in S^{\uparrow^+} \) iff \( (w, v) \in (S^{\uparrow^+} P)^{\tau} \), from which we conclude that \( S^{\uparrow^+} P = (S^{\uparrow^+} P)^{\tau} \). It follows that \( \uparrow^+ \leq \tau \).

As the second part of the proof, we show that \( \tau \leq \downarrow^{++} \). Let \( S \) be a plausibility order, and \( P \) a proposition. If \( S \cap P = \emptyset \), then \( S^{\tau} P = \emptyset \), since \( \tau \) is strict. But \( \emptyset^{\tau} = \emptyset \), so our claim holds. We may thus suppose that \( S \cap P \neq \emptyset \). Assuming that it is not the case that \( \neg P \leq S \ P \), it follows that \( S \not\models SbP \) (i.e., \( P \) is remotely plausible in \( S \)), and since \( \downarrow^{++} = Sb \), our claim holds. Suppose, then that \( \neg P \leq S \ P \). By the fact that \( \tau \) is informative, it follows that \( S^{\tau} P \neq S \). Since \( \tau \) is defeasible, \( S^{\tau} P = S \). So there exists a pair \( (w, v) \in S \times S \) such that \( (w, v) \in S^{\tau} P \), \( (w, v) \notin S \). Since \( \tau \) is informative, it is impossible that \( (a) \ w, v \in P \), \( (b) \ w, v \in \neg P \), or \( (c) \ w \in \neg P, v \in P \). Hence \( w \in P, v \in \neg P \). But then, \( S \not\models Sb\neg P \), hence \( S^{\uparrow^+} = S \).

From the “pragmatic” point of view, one could take this result to support the claim that categorical claims should be semantically represented using \( \uparrow^+ \).
while epistemic *might* claims should be represented using $\downarrow^{\sim+}$. But notice that the preceding result is also useful from the point of view of the “classical” position, in that the “classical” theorist could take it as providing support for the claim that $\downarrow^+$ should figure as providing a semantics for *must*, while $\downarrow^{\sim+}$ should figure as providing a semantics for *may*!

Incidentally, this perspective promises to provide an explanation for the intuition voiced above that *may* is stronger than *might*: while *might*, at best, serves to draw attention to a possibility, existential modals stronger than *might* serve to raise the plausibility of the prejacent.

### 6.3.9. Semantics for Epistemic Modals

As announced above, we adopt what I have called the classical view here. Then, the following picture emerges:

1. $!$ (fixed point: $K$) — is
2. $\downarrow^+$ (fixed point: $Sb$) — *must*
3. $\uparrow^+$ (fixed point: $B$) *should*
4. $\uparrow^{\sim+}$ (fixed point: $B^\sim$) — *may*
5. $\downarrow^{\sim+}$ (fixed point: $Sb^\sim$) — *could*
6. $!^\sim$ (fixed point: $K^\sim$) — *might*

Of these, the dynamic attitudes in (1.)–(5.) satisfy informativeness, while the dynamic attitudes in (2.)–(6.) satisfy defeasibility. All dynamic attitudes in (1.)–(6.) satisfy strictness, triviality and affirmativity.

The choices made in (1.), (2.), (5.) and (6.) are motivated by Proposition 103 and Proposition 104 above, in conjunction with the observation that $!, \downarrow^+$ and $!^\sim$ are canonical for their fixed point (Proposition 59), and $\downarrow^{\sim+}$ is the unique dynamic attitude that is positionally optimal for its fixed point (Proposition 67). Choices (3.) and (4.) are motivated by the fact that *simple belief* $B$ corresponds to the weakest natural form of (static) acceptance in our setting (essentially derived from the privileged position of the *most plausible worlds* in a plausibility order)—and the fixed point of $\uparrow^+$ is belief (Proposition 8); on the other hand $B^\sim$ is, as the dual of $B$, the strongest form of affirmation that falls short of proper acceptance. Note that both $\uparrow^+$ and $\uparrow^{\sim+}$ are also the unique dynamic attitudes that are positionally optimal for their fixed point (Proposition 67).