PART I

INTENSIONALITY, ATTITUDES, MODALS, CONDITIONALS

Chapter One
Beginnings

> Language is the main instrument of man's refusal to accept the world as it is.
>
> George Steiner, *After Babel*, p. 228

We introduce the idea of extension vs. intension and its main use: taking us from the actual here and now to past, future, possible, counterfactual situations. We develop a compositional framework for intensional semantics.

1.1 Displacement

Hockett (1960) in a famous article (and a follow-up, Hockett & Altmann (1968)) presented a list of DESIGN FEATURES OF HUMAN LANGUAGE. This list continues to play a role in current discussions of animal communication. One of the design features is DISPLACEMENT. Human language is not restricted to discourse about the *actual here and now*.[1]

Charles F. Hockett. 1960. The origin of speech. *Scientific American* 203. 89–96

How does natural language untie us from the actual here and now? One degree of freedom is given by the ability to name entities and refer to them even if they are not where we are when we speak:

(1) Thomas is in Hamburg.

This kind of displacement is not something we will explore here. We'll take it for granted.

Consider a sentence with no names of absent entities in it:

(2) It is snowing (in Cambridge).

1 Steiner (1998: 226) writes: "Hypotheticals, 'imaginaries', conditionals, the syntax of counterfactuality and contingency may well be the generative centres of human speech".

On its own, (2) makes a claim about what is happening right now here in Cambridge. But there are devices at our disposal that can be added to (2), resulting in claims about snow in displaced situations. Displacement can occur in the temporal dimension and/or in what might be called the modal dimension. Here's an example of temporal displacement:

The terms modal and m
ity descend from the
modus, "way", and
cient terms pertain
the way a proposition
necessarily, contingent

(3) At noon yesterday, it was snowing in Cambridge.

This sentence makes a claim not about snow now but about snow at noon yesterday, a different time from now.

Here's an example of modal displacement:

(4) If the storm system hadn't been deflected by the jet stream, it would have been snowing in Cambridge.

This sentence makes a claim not about snow in the actual world but about snow in the world as it would have been if the storm system hadn't been deflected by the jet stream, a world distinct from the actual one (where the system did not hit us), a merely possible world.

See Kratzer (1981, 1991) for more examples of modal constructions.

Natural language abounds in modal constructions. (4) is a so-called counterfactual conditional. Here are some other examples:

(5) Modal Auxiliaries
It may be snowing in Cambridge.

(6) Modal Adverbs
Possibly, it will snow in Cambridge tomorrow.

(7) Propositional Attitudes
Jens believes that it is snowing in Cambridge.

(8) Habituals
Jane smokes.

(9) Generics
Bears like honey.

https://en.wikipedia.org/wiki/Sherlock_Holmes

In this chapter, we will put in place the basic framework of intensional semantics, the kind of semantics that models displacement of the point of evaluation in temporal and modal dimensions. To do this, we will start with one rather special example of modal displacement:

(10) In the world of Sherlock Holmes, a detective lives at 221B Baker Street.

(10) doesn't claim that a detective lives at 221B Baker Street in the actual world (presumably a false claim), but that in the world as it is described in the Sherlock Holmes stories of Sir Arthur Conan Doyle, a detective lives at 221B Baker Street (a true claim, of course). We choose this example rather than one of the more run-

of-the-mill displacement constructions because we want to focus on conceptual and technical matters before we do serious empirical work.

The questions we want to answer are: How does natural language achieve this feat of modal displacement? How do we manage to make claims about other possible worlds? And why would we want to?

To make displacement possible and compositionally tractable, we need meanings of natural language expressions, and of sentences in particular, to be displaceable in the first place. They need to be "portable", so to speak, able to make claims about more than just the actual here and now. And need we need other natural language expressions that take that portable meaning and apply it to some situation other than the actual here and now. That is what intensionality is all about.

The basic idea of the account we'll develop is this:

- expressions are assigned their semantic values relative to a possible world;
- in particular, sentences have truth-values in possible worlds;
- in the absence of modal displacement, we evaluate sentences with respect to the "actual" world, the world in which we are speaking;
- modal displacement changes the world of evaluation;
- displacement is effected by special operators, whose semantics is our primary concern here.

A terminological note: we will call the sister of the intensional operator its PREJACENT, a useful term introduced by our medieval colleagues.

1.2 An Intensional Semantics in 10 Easy Steps

1.2.1 Laying the Foundations

Step 1: Possible Worlds. Our first step is to introduce possible worlds. This is not the place to discuss the metaphysics of possible worlds in any depth. Instead, we will just start working with them and see what they can do for us. Basically, a possible world is a way that things might have been. In the actual world, there are two coffee mugs on my desk, but there could have been more or less. So, there is a possible world — albeit a rather bizarre one — where there are 17 coffee mugs on my desk. We join Heim & Kratzer in adducing this quote from D. Lewis (1986: 1f.):

David Lewis

> The world we live in is a very inclusive thing. Every stick and every stone you have ever seen is part of it. And so are you and I. And so are the planet Earth, the solar system, the entire Milky Way, the remote galaxies we see through telescopes, and (if there are such things) all the bits of empty space between the stars and galaxies. There is nothing so far away from us as not to be part of our world. Anything at any distance at all is to be included. Likewise the world is inclusive in time. No long-gone ancient Romans, no long-gone pterodactyls, no long-gone primordial clouds of plasma are too far in the past, nor are the dead dark stars too far in the future, to be part of the same world. . . .
>
> The way things are, at its most inclusive, means the way the entire world is. But things might have been different, in ever so many ways. This book of mine might have been finished on schedule. Or, had I not been such a commonsensical chap, I might be defending not only a plurality of possible worlds, but also a plurality of impossible worlds, whereof you speak truly by contradicting yourself. Or I might not have existed at all — neither myself, nor any counterparts of me. Or there might never have been any people. Or the physical constants might have had somewhat different values, incompatible with the emergence of life. Or there might have been altogether different laws of nature; and instead of electrons and quarks, there might have been alien particles, without charge or mass or spin but with alien physical properties that nothing in this world shares. There are ever so many ways that a world might be: and one of these many ways is the way that this world is.

Previously, our "metaphysical inventory" included a domain of entities and a set of two truth-values and increasingly complex functions between entities, truth-values, and functions thereof. Now, we will add possible worlds to the inventory. Let's assume we are given a set W, the set of all possible worlds, which is a vast space since there are so many ways that things might have been different from the way they are. Each world has as among its parts entities like you and me and these coffee mugs. Some of them may not exist in other possible worlds. So, strictly speaking each possible worlds has its own, possibly distinctive, domain of entities. What we will use in our system, however, will be the grand union of all these world-specific domains of entities. We will use D to stand for the set of all possible individuals.

Among the many possible worlds that there are — according to Lewis, there is a veritable plenitude of them — is the world as it is described in the Sherlock Holmes stories by Sir Arthur Conan Doyle. In that world, there is a famous detective Sherlock Holmes, who lives at 221B Baker Street in London and has a trusted sidekick named Dr. Watson. Our sentence *In the world of Sherlock Holmes, a detective lives at 221B Baker Street* displaces the claim that a famous detective lives at 221B Baker Street from the actual world to the world as described in the Sherlock Holmes stories. In other words, the following holds:[2]

(11) The sentence *In the world of Sherlock Holmes, a detective lives at 221B Baker Street* is true in a world w iff the sentence *a detective lives at 221B Baker Street* is true in the world as it is described in the Sherlock Holmes stories.

What this suggests is that we need to make space in our system for having devices that control in what world a claim is evaluated. This is what we will do now.

Step 2: The Evaluation World Parameter. Recall from H& K that we were working with a semantic interpretation function that was relativized to an assignment function g, which was needed to take care of pronouns, traces, variables, etc. From now on, we will relativize the semantic values in our system to possible worlds as well. What this means is that from now on, our interpretation function will have two superscripts: a world w and an assignment g: $[\![\cdot]\!]^{w,g}$.

So, the prejacent embedded in (10) will have its truth-conditions described as follows:[3]

(12) $[\![\text{a famous detective lives at 221B Baker Street}]\!]^{w,g} = 1$
iff a famous detective lives at 221B Baker Street in world w.

It is customary to refer to the world for which we are calculating the extension of a given expression as the evaluation world. In the absence of any shifting

2 We will see in Section 1.3.2 that this is not quite right. It'll do for now.

3 Recall from H& K, pp.22f, that what's inside the interpretation brackets is a mention of an object language expression. They make this clear by bold-facing all object language expressions inside interpretation brackets. In these notes, we will follow common practice in the field and not use a special typographic distinction, but let it be understood that what is interpreted are object language expressions.

devices, we would normally evaluate a sentence in the actual world. But then there are shifting devices such as our *in the world of Sherlock Holmes*. We will soon see how they work. But first some more pedestrian steps: adding lexical entries and composition principles that are formulated relative to a possible world. This will allow us to derive the truth-conditions as stated in (12) in a compositional manner.

STEP 3: LEXICAL ENTRIES. Among our lexical items, we can distinguish between items which have a WORLD-DEPENDENT semantic value and those that are world-independent. Predicates are typically world-dependent. Here are some sample entries.

(13) For any $w \in W$ and any assignment function g:
a. $[\![\text{famous}]\!]^{w,g} = \lambda x \in D.\ x$ is famous in w.[4,5]
b. $[\![\text{detective}]\!]^{w,g} = \lambda x \in D.\ x$ is a detective in w.
c. $[\![\text{lives-at}]\!]^{w,g} = \lambda x \in D.\ \lambda y \in D.\ y$ lives-at x in w.

The set of detectives will obviously differ from world to world, and so will the set of famous individuals and the set of pairs where the first element lives at the second element.

Other items have semantic values which do not differ from world to world. The most important such items are certain "logical" expressions, such as truth-functional connectives and determiners:

(14) a. $[\![\text{and}]\!]^{w,g} = \lambda u \in D_t.\ \lambda v \in D_t.\ u = v = 1.$
b. $[\![\text{the}]\!]^{w,g} = \lambda f \in D_{\langle e,t\rangle}\colon \exists! x[f(x) = 1].$ the y such that $f(y) = 1.$
c. $[\![\text{every}]\!]^{w,g} = \lambda f_{\langle e,t\rangle}.\ \lambda h_{\langle e,t\rangle}.\ \forall x_e\colon f(x) = 1 \rightarrow h(x) = 1.$
d. $[\![\text{a/some}]\!]^{w,g} = \lambda f_{\langle e,t\rangle}.\ \lambda h_{\langle e,t\rangle}.\ \exists x_e\colon f(x) = 1\ \&\ h(x) = 1.$

Note the ruthless condensation of the notation in (c) and (d): variables are subscripted with the type of the domain that their values are constrained to come from.

Note that there is no occurrence of w on the right-hand side of the entries in (14). That's the tell-tale sign of the world-independence of the semantics of these items.

We will also assume that proper names have world-independent semantic values, that is, they refer to the same individual in any possible world.

(15) a. $[\![\text{Noam Chomsky}]\!]^{w,g}$ = Noam Chomsky.

4 Of course, "$\lambda x \in D.\ \ldots$" is short for "$\lambda x\colon\ x \in D.\ \ldots$". Get used to semanticists condensing their notation whenever convenient! A further step of condensation is taken below: "$\lambda x\colon\ x \in D_e.\ \ldots$" becomes "$\lambda x_e.\ \ldots$".

5 Always make sure that you actually understand what the notation means. Here, for example, we are saying that the semantic value of the word *famous* with respect to a given possible world w and a variable assignment g is that function that is defined for an argument x only if x is a member of the domain of individuals and that, if it is defined, yields the truth-value 1 if and only if x is famous in w. (13a) does *not* mean that the function maps x to "x is famous in w", which would be very weird: mapping an individual to a meta-language statement!

b. $[\![\text{Sherlock Holmes}]\!]^{w,g}$ = Sherlock Holmes.
c. $[\![\text{221B Baker Street}]\!]^{w,g}$ = 221B Baker Street.

Step 4: Composition Principles. The old rules of Functional Application, Predicate Modification, and λ-Abstraction can be retained almost intact. We just need to modify them by adding world-superscripts to the interpretation function. For example:

(16) Functional Application (FA)
If α is a branching node and $\{\beta, \gamma\}$ the set of its daughters, then, for any world w and assignment g: if $[\![\beta]\!]^{w,g}$ is a function whose domain contains $[\![\gamma]\!]^{w,g}$, then $[\![\alpha]\!]^{w,g} = [\![\beta]\!]^{w,g}([\![\gamma]\!]^{w,g})$.

The rule simply passes the world parameter down.

Step 5: Truth. Lastly, we will want to connect our semantic system to the notion of the truth of an utterance. We first adopt the "Appropriateness Condition" from Heim & Kratzer (p.243):

(17) Appropriateness Condition
A context c is *appropriate* for an LF ϕ only if c determines a variable assignment g_c whose domain includes every index which has a free occurrence in ϕ.

We then intensionalize Heim & Kratzer's definition of truth and falsity of utterances:

(18) Truth and Falsity Conditions for Utterances
An utterance of a sentence ϕ in a context c in a possible world w is *true* iff $[\![\phi]\!]^{w,g_c} = 1$ and *false* if $[\![\phi]\!]^{w,g_c} = 0$.

Exercise 1.1: Compute under what conditions an utterance in possible world w_7 (which may or may not be the one we are all living in) of the sentence *a famous detective lives at 221B Baker Street* is true. [Since this is the first exercise in this framework, please do this in excrutiating detail, not skipping any steps.] □

1.2.2 Intensional Operators

So far we have merely "redecorated" the system inherited from Heim & Kratzer. We have introduced possible worlds into our inventory, our lexical entries and our old composition principles. But with the tools we have now, all we can do so far is to keep track of the world in which we evaluate the semantic value of an expression, complex or lexical. We will get real mileage once we introduce intensional operators which are capable of shifting the world parameter. We mentioned that there are a number of devices for modal displacement. As advertised, for now, we will just focus on a very particular one: the expression *in the world of Sherlock Holmes*. We will assume, as seems reasonable, that this expression is a sentence-modifier both syntactically and semantically.

STEP 6: A SYNCATEGOREMATIC ENTRY. We begin with a heuristic step. We want to derive something like the following truth-conditions for our sentence:

(19) $[\![$in the world of Sherlock Holmes,
a famous detective lives at 221B Baker Street$]\!]^{w,g} = 1$
iff the world w' as it is described in the Sherlock Holmes stories is such that there exists a famous detective in w' who lives at 221B Baker Street in w'.

We would get this if in general we had this rule for *in the world of Sherlock Holmes*:

(20) For any sentence ϕ, any world w, and any assignment g:
$[\![$in the world of Sherlock Holmes $\phi]\!]^{w,g} = 1$
iff the world w' as it is described in the Sherlock Holmes stories is such that $[\![\phi]\!]^{w',g} = 1$.

This is a so-called SYNCATEGOREMATIC treatment of the meaning of this expression. Instead of giving an explicit semantic value to the expression, we specify what effect it has on the meaning of a complex expression that contains it. In (20), we do not compute the meaning for *in the world of Sherlock Holmes,* ϕ from the combination of the meanings of its parts, since *in the world of Sherlock Holmes* is not given a separate meaning, but in effect triggers a special composition principle. This format is very common in modal logic systems, which usually give a syncategorematic semantics for the two modal operators (the necessity operator $\Box$ and the possibility operator $\Diamond$). When one only has a few closed class expressions to deal with that may shift the world parameter, employing syncategorematic entries is a reasonable strategy. But we are facing a multitude of displacement devices. So, we will need to make our system more modular.

The diamond $\Diamond$ symbol for possibility is due to C.I. Lewis, first introduced in C. I. Lewis & Langford (1932), but he made no use of a symbol for the dual combination $\neg\Diamond\neg$. The dual symbol $\Box$ was later devised by F.B. Fitch and first appeared in print in 1946 in a paper by his doctoral student Barcan (1946). See footnote 425 of Hughes & M.J. Cresswell (1968). Another notation one finds is L for necessity and M for possibility, the latter from the German *möglich* 'possible'.

So, we want to give *in the world of Sherlock Holmes* its own meaning and combine that meaning with that of its prejacent by a general composition principle. The Fregean slogan we adopted says that all composition is function application (modulo the need for λ-abstraction and the possible need for predicate modification).[6] So, what we will want to do is to make (19) be the result of functional application. But we can immediately see that it cannot be the result of our usual rule of functional application, since that would feed to *in the world of Sherlock Holmes* the semantic value of *a famous detective lives in 221B Baker Street* in w, which would be a particular truth-value, 1 if a famous detective lives at 221B

6 See Heim & Kratzer, Section 4.3, pp. 63–72 for a reminder about the status of predicate modification.

Baker Street in w and 0 if there doesn't. And whatever the semantics of *in the world of Sherlock Holmes* is, it is certainly *not* a truth-functional operator.

So, we need to feed something else to *in the world of Sherlock Holmes*. At the same time, we want the operator to be able to shift the evaluation world of its prejacent. Can we do this?

Exercise 1.2: How would you show that *in the world of Sherlock Holmes* is not a truth-functional operator? □

Step 7: Intensions. We will define a richer notion of semantic value, the intension of an expression. This will be a function from possible worlds to the extension of the expression in that world. The intension of a sentence can be applied to any world and give the truth-value of the sentence in that world. Intensional operators take the intension of their prejacent as their argument, that is we will feed the intension of the embedded sentence to the shifting operator. The operator will use that intension and apply it to the world it wants the evaluation to happen in. Voilà.

Now let's spell that account out. Our system actually provides us with two kinds of meanings. For any expression α, we have $[\![\alpha]\!]^{w,g}$, the semantic value of α in w, also known as the extension of α in w. But we can also calculate $\lambda w.[\![\alpha]\!]^{w,g}$, the function that assigns to any world w the extension of α in that world. This is usually called the intension of α. We will sometimes use an abbreviatory notation[7] for the intension of α:

As before in H&K, we make no claim that the semantic values that are attributed to expressions in our framework fully capture what is informally meant by "meaning". But certainly, intensions come closer to "meaning" than extensions.

(21) $$[\![\alpha]\!]^{g}_{¢} := \lambda w.[\![\alpha]\!]^{w,g}.$$

It should be immediately obvious that since the definition of intension abstracts over the evaluation world, intensions are not world-dependent.[8,9]

Note that strictly speaking, it now makes no sense anymore to speak of "*the* semantic value" of an expression α. What we have is a semantic system that allows us to calculate extensions (for a given possible world w) as well as intensions for all (interpretable) expressions. We will see that when α occurs in a particular bigger tree, it will always be determinate which of the two "semantic values" of α is the one that enters into the compositional semantics. So, that

7 The notation with the subscripted cent-sign comes from Montague Grammar. See e.g. Dowty, Wall & Peters (1981: 147).

8 Since intensions are by definition not dependent on the choice of a particular world, it makes no sense to put a world-superscript on the intension-brackets. So don't ever write "$[\![\ldots]\!]^{w,g}_{¢}$"; we'll treat that as undefined nonsense.

9 The definition here is simplified, in that it glosses over the fact that some expressions, in particular those that contain presupposition triggers, may fail to have an extension in certain worlds. In such a case, the intension has no extension to map such a world to. Therefore, the intension will have to be a partial function. So, the official, more "pedantic", definition will have to be as follows: $[\![\alpha]\!]^{g}_{¢} := \lambda w\colon\ \alpha \in \mathrm{dom}([\![]\!]^{w,g}).[\![\alpha]\!]^{w,g}$.

one — whichever one it is, the extension or the intension of α — might then be called "*the* semantic value of α in the tree β".

The Port-Royal logicians distinguished EXTENSION from COMPREHENSION. Leibniz preferred the term INTENSION rather than COMPREHENSION. The notion probably goes back even further. See Spencer (1971) for some notes on this. The possible worlds interpretation is due to Carnap (1947).

It should be noted that the terminology of EXTENSION VS. INTENSION is time-honored but that the possible worlds interpretation thereof is more recent. The technical notion we are using is certainly less rich a notion of meaning than traditionally assumed.[10]

STEP 8: SEMANTIC TYPES AND SEMANTIC DOMAINS. If we want to be able to feed the intensions to lexical items like *in the world of Sherlock Holmes*, we need to have the appropriate types in our system.

Recall that W is the set of all possible worlds. And recall that D is the set of all POSSIBLE INDIVIDUALS and thus contains all individuals existing in the actual world *plus* all individuals existing in any of the merely possible worlds.

We now expand the set of semantic types, to add intensions. Intensions are functions from possible worlds to all kinds of extensions. So, basically we want to add for any kind of extension we have in our system, a corresponding kind of intension, a function from possible worlds to that kind of extension.

We add a new clause, (22c), to the definition of semantic types:

(22) SEMANTIC TYPES
  a. e and t are semantic types.
  b. If σ and τ are semantic types, then $\langle \sigma, \tau \rangle$ is a semantic type.
  c. If σ is a semantic type, then $\langle s, \sigma \rangle$ is a semantic type.
  d. Nothing else is a semantic type.

We also add a fourth clause to the previous definition of semantic domains:

(23) SEMANTIC DOMAINS
  a. $D_e = D$, the set of all possible individuals
  b. $D_t = \{0, 1\}$, the set of truth-values
  c. If σ and τ are semantic types, then $D_{\langle \sigma, \tau \rangle}$ is the set of all functions from D_σ to D_τ.
  d. INTENSIONS: If σ is a type, then $D_{\langle s, \sigma \rangle}$ is the set of all functions from W to D_σ.

Clause (d) is the addition to our previous system of types. The functions of the schematic type $\langle s, \ldots \rangle$ are intensions.[11] Here are some examples of intensions:

---

10 For example, Frege's "modes of presentation" are not obviously captured by this possible worlds implementation of extension/intension.

11 Note a curious feature of this set-up: there is no type s and no associated domain. This corresponds to the assumption that there are no expressions of English that take as their extension a possible world, that is, there are no pronouns or names referring to possible worlds. We will

- The intensions of sentences are of type $\langle s, t\rangle$, functions from possible worlds to truth values. These are usually called PROPOSITIONS. Note that if the function is total, then we can see the sentence as picking out a set of possible worlds, those in which the sentence is true. More often than not, however, propositions will be PARTIAL functions from worlds to truth-values, that is functions that fail to map certain possible worlds into either truth-value. This will be the case when the sentence contains a presupposition trigger, such as *the*. The famous sentence *The King of France is bald* has an intension that (at least in the analysis sketched in Heim & Kratzer) is undefined for any world where there fails to be a unique King of France.
- The intensions of one-place predicates are of type $\langle s, \langle e, t\rangle\rangle$, functions from worlds to set of individuals. These are usually called PROPERTIES.
- The intensions of expressions of type e are of type $\langle s, e\rangle$, functions from worlds to individuals. These are usually called INDIVIDUAL CONCEPTS.

STEP 9: A LEXICAL ENTRY FOR A SHIFTER. We are ready to formulate the lexical entry for *in the world of Sherlock Holmes*:[12]

(24) $[\![\text{in the world of Sherlock Holmes}]\!]^{w,g} =$
$\lambda p_{\langle s,t\rangle}$. the world w' as it is described in the Sherlock Holmes stories is such that $p(w') = 1$.

That is, *in the world of Sherlock Holmes* expects as its argument a function of type $\langle s, t\rangle$, a proposition. It yields the truth-value 1 iff the proposition is true in the world as it is described in the Sherlock Holmes stories.

All that's left to do now is to provide *in the world of Sherlock Holmes* with a proposition as its argument. This is the job of a new composition principle.

STEP 10: INTENSIONAL FUNCTIONAL APPLICATION. We add the new rule of Intensional Functional Application.

(25) INTENSIONAL FUNCTIONAL APPLICATION (IFA)
If α is a branching node and $\{\beta, \gamma\}$ the set of its daughters, then, for any world w and assignment g: if $[\![\beta]\!]^{w,g}$ is a function whose domain contains $[\![\gamma]\!]^{g}_{\cent}$, then $[\![\alpha]\!]^{w,g} = [\![\beta]\!]^{w,g}([\![\gamma]\!]^{g}_{\cent})$.

---

actually question this assumption in a later chapter. For now, we will stay with this more conventional set-up.

12 This is not yet the final semantics, see Section 1.3 for complications. One complication we will not even start to discuss is that obviously it is not a necessity that there are Sherlock Holmes stories in the first place and that the use of this operator *presupposes* that they exist; so a more fully explicit semantics would need to build in that presuppositional component. Also, note again the condensed notation: "$\lambda p_{\langle s,t\rangle}. \ldots$" stands for the fully official "$\lambda p\colon\ p \in D_{\langle s,t\rangle}. \ldots$".

This is the crucial move. It makes space for expressions that want to take the intension of their sister as their argument and do stuff to it. Now, everything is in place. Given (24), the semantic argument of *in the world of Sherlock Holmes* will not be a truth-value but a proposition. And thus, *in the world of Sherlock Holmes* will be able to check the truth-value of its prejacent in various possible worlds. To see in practice that we have all we need, please do the following exercise.

EXERCISE 1.3: Calculate the conditions under which an utterance in a given possible world w_7 of the sentence *in the world of the Sherlock Holmes stories, a famous detective lives at 221B Baker Street* is true. □

EXERCISE 1.4: What in our system prevents us from computing the extension of *Watson is slow*, for example, by applying the intension of *slow* to the extension of *Watson*? What in our system prevents us from computing the extension of *Watson is slow* by applying the intension of *slow* to the intension of *Watson*? □

Please think about this exercise before looking at Section 1.4, which explores this issue.

EXERCISE 1.5: What is wrong with the following equation:

(26) $(\lambda x.\ x \text{ is slow in } w)$ (Watson) = Watson is slow in w.

[Hint: there is nothing wrong with the following:

(27) $(\lambda x.\ x \text{ is slow in } w)$ (Watson) = 1 iff Watson is slow in w.] □

1.3 Comments and Complications

1.3.1 Intensions All the Way?

We have seen that to adequately deal with expressions like *in the world of Sherlock Holmes*, we need an intensional semantics, one that gives us access to the extensions of expressions across the multitude of possible worlds. At the same time, we have kept the semantics for items like *and*, *every*, and *a* unchanged and extensional. This is not the only way one can set up an intensional semantics. The following exercise demonstrates this.

EXERCISE 1.6: Consider the following "intensional" meaning for *and*:

(28) $[\![\text{and}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \lambda q_{\langle s,t\rangle}.\ p(w) = q(w) = 1.$

With this semantics, *and* would operate on the intensions of the two conjoined sentences. In any possible world w, the complex sentence will be true iff the component propositions are both true of that world.

Compute the truth-conditions of the sentence *In the world of Sherlock Holmes, Holmes is quick and Watson is slow* both with the extensional meaning for *and* given earlier and the intensional meaning given here. Is there any difference in the results? □

There are then at least two ways one could develop an intensional system.

(i) We could "generalize to the worst case" and make the semantics deliver intensions as *the* semantic value of an expression. Such systems are common in the literature (see D. Lewis 1970, Max Cresswell 1973).

(ii) We could maintain much of the extensional semantics we have developed so far and extend it conservatively so as to account for non-extensional contexts.

We have chosen to pursue (ii) over (i), because it allows us to keep the semantics of extensional expressions simpler. The philosophy we follow is that we will only move to the intensional sub-machinery when triggered by an expression that creates a non-extensional context. As the exercise just showed, this is more a matter of taste than a deep scientific decision.

1.3.2 Why Talk about Other Worlds?

Why would natural language bother having such elaborate mechanisms to talk about other possible worlds? While having devices for spatial and temporal displacement (talking about Hamburg or what happened yesterday) seems eminently reasonable, talking about worlds other than the actual world seems only suitable for poets and the like. So, why?

The solution to this puzzle lies in a fact that our current semantics of the shifter *in the world of Sherlock Holmes* does not yet accurately capture: modal sentences have empirical content, they make CONTINGENT claims, claims that are true or false depending on the circumstances in the actual world.

Our example sentence *In the world of Sherlock Holmes, a famous detective lives at 221B Baker Street* is true in this world but it could easily have been false. There is no reason why Sir Arthur Conan Doyle could not have decided to locate Holmes' abode on Abbey Road.

To see that our semantics does not yet capture this fact, notice that in the semantics we gave for *in the world of Sherlock Holmes*:

(29) $[\![\text{in the world of Sherlock Holmes}]\!]^{w,g} =$
$\lambda p_{\langle s,t\rangle}$. the world w' as it is described in the Sherlock Holmes stories is such that $p(w') = 1$.

there is no occurrence of w on the right hand side. This means that the truth-conditions for sentences with this shifter are world-independent. In other words, they are predicted to make non-contingent claims that are either true no-matter-what or false no-matter-what. This needs to be fixed.

The fix is obvious: what matters to the truth of our sentence is the content of the Sherlock Holmes stories as they are in the evaluation world. So, we need the following semantics for our shifter:

(30) $[\![\text{in the world of Sherlock Holmes}]\!]^{w,g} =$
$\lambda p_{\langle s,t \rangle}$. the world w' as it is described in the Sherlock Holmes stories *in* w is such that $p(w') = 1$.

We see now that sentences with this shifter do make a claim about the evaluation world: namely, that the Sherlock Holmes stories as they are in the evaluation world describe a world in which such-and-such is true. So, what is happening is that although it appears at first as if modal statements concern other possible worlds and thus couldn't really be very informative, they actually only talk about certain possible worlds, those that stand in some relation to what is going on at the ground level in the actual world. As a crude analogy, consider:

(31) My grandmother is sick.

At one level this is a claim about my grandmother. But it is also a claim about me: namely that I have a grandmother who is sick. Thus it is with modal statements. They talk about possible worlds that stand in a certain relation to the actual world and thus they make claims about the actual world, albeit slightly indirectly.

1.3.3 The Worlds of Sherlock Holmes

So far, we have played along with colloquial usage in talking of *the* world of Sherlock Holmes. But it is important to realize that this is sloppy talk. D. Lewis (1978) writes:

> [I]t will not do to follow ordinary language to the extent of supposing that we can somehow single out a single one of the worlds [as the one described by the stories]. Is the world of Sherlock Holmes a world where Holmes has an even or an odd number of hairs on his head at the moment when he first meets Watson? What is Inspector Lestrade's blood type? It is absurd to suppose that these questions about the world of Sherlock Holmes have answers. The best explanation of that is that the world*s* of Sherlock Holmes are plural, and the questions have different answers at different ones.

The usual move at this point is to talk about the set of worlds "compatible with the (content of) Sherlock Holmes stories in w". We imagine that we ask of each possible world whether what is going on in it is compatible with the stories as they were written in our world. Worlds where Holmes lives on Abbey Road are not compatible. Some worlds where he lives at 221B Baker Street are compatible (again not all, because in some such worlds he is not a famous detective but an obscure violinist). Among the worlds compatible with the stories are ones where he has an even number of hairs on his head at the moment when he first meets Watson and there are others where he has an odd number of hairs at that moment.

What the operator *in the world of Sherlock Holmes* expresses is that its complement is true throughout the worlds compatible with the stories. In other words, the operator *universally quantifies* over the compatible worlds. Our next iteration of the semantics for the operator is therefore this:

(32) $[\![\text{in the world of Sherlock Holmes}]\!]^{w,g} =$
$\lambda p_{\langle s,t\rangle}.\ \forall w'$ compatible with the Sherlock Holmes stories *in* w:
$p(w') = 1.$

At a very abstract level, the way we parse sentences of the form *in the world of Sherlock Holmes,* ϕ is that both components, the *in*-phrase and the prejacent, determine sets of possible worlds and that the set of possible worlds representing the content of the fiction mentioned in the *in*-phrase is a subset of the set of possible worlds determined by the prejacent. We will see the same rough structure of relating sets of possible worlds in other intensional constructions.

This is where we will leave things. There is more to be said about fiction operators like *in the world of Sherlock Holmes*, but we will just refer to you to the relevant literature. In particular, one might want to make sense of Lewis' idea that a special treatment is needed for cases where the sentence makes a claim about things that are left open by the fiction (no truth-value, perhaps?). One also needs to figure out how to deal with cases where the fiction is internally inconsistent. In any case, for our purposes we're done with this kind of operator.

1.3.4 What's next

With the basic framework of intensional semantics in place, we can now look at a succession of intensional operators. In particular, we will explore the semantics of propositional attitude predicates such as *believe* or *want*, modal auxiliaries such as *must* or *might*, and conditional sentences. In each case, we will see that they shift the evaluation of their prejacents to a certain set of possible worlds and that which worlds we are taken to depends on certain facts about the original evaluation world. We will look at the non-trivial issues that arise when several intensional operators interact (modals under attitudes, modals in the consequent of a conditional, etc.). We will also see that constituents of the prejacent can sometimes be evaluated with respect to a world that is not the world that the intensional operator is taking us to (so-called *de re* readings). Further, we will move from worlds to times and explore the semantics of tense and aspect. And, for the intrepid, this can all come together by exploring how tense and aspect interact with attitudes, modality, and conditionals.

1.4 *Issues with an informal meta-language

Starred sections are optional on a first pass.

Exercise 1.5 asks what is wrong with writing something like[13]

(33) (λx. x is slow in *w*) (Watson) = Watson is slow in *w*.

Think about it. On the left hand side of the "=" sign is a meta-language expression consisting of a λ-expression (so some kind of function) applied to an individual (contributed by the meta-language name "Watson"). The function is a function from individuals to truth-values that will deliver the truth-value 1 iff the individual is slow in world *w*. So, what we have on the left hand side is the result of a function from individuals to truth-values applied to an individual. In other words, on the left hand side we have a truth-value, namely the truth-value 1 if Watson is slow in *w* and the truth-value 0 if Watson is not slow in *w*.

Now, what do we have on the right hand side of the "="? We have the meta-language sentence "Watson is slow in *w*". That is not nor does it contribute a truth-value. It is a statement of fact. Truth-values are not the same as statements of fact.

The proper thing to do is to write

(34) (λx. x is slow in *w*) (Watson) = 1 iff Watson is slow in *w*.

There are actually two ways to parse the statement in (34), both legitimate it appears.

On one parse, the major connective is the meta-language expression "iff". On its left hand side is a meta-language statement (that applying the function to the individual Watson gives the truth-value 1) and on the right hand side of the "iff" we have another meta-language statement (that Watson is slow in *w*). So, the whole thing says that these two statements are equivalent: (i) that function applied to that individual gives us the truth-value 1, and (ii) that Watson is slow in *w*.

The other parse is perhaps more conspicuously represented as follows:

(35) $(\lambda x.\ x \text{ is slow in } w)\ (\text{Watson}) = \begin{cases} 1 \text{ if Watson is slow in } w \\ 0 \text{ if Watson is not slow in } w \end{cases}$

Here, the "=" sign is the major connective. The left hand side is a meta-language expression that resolves to a truth-value and the right hand side as well contributes a truth-value: 1 if such and such and 0 if such and such.

Is this weird? It turns out that natural language, not just our semi-formal meta-language, has conditionals that seem very similar: *I fear [the consequences if we fail]*. See Lasersohn 1996 for some discussion.

H& K, of course, introduced a convention that allowed meta-language statements to be used in a place where a truth-value was expected (p.37, (9)):

Read "[$\lambda\alpha$: ϕ. γ]" as either (i) or (ii), whichever makes sense.

13 Thanks to Magda Kaufmann, Angelika Kratzer, and Ede Zimmermann for discussion on the issues explored in this section.

(i) "the function which maps every α such that ϕ to γ"
(ii) "the function which maps every α such that γ to 1, if γ, and to 0 otherwise"

Since it never makes sense to map anything to a meta-language statement, no ambiguity will ever arise.

So, one might want to extend this leeway and use it in the case of (33) as well. We could say that in general, meta-language statements supply truth-values wherever that makes sense. In that case, (33) is just shorthand for (34).

This is the approach of von Stechow 1991.

Alternatively, one can introduce a new notation that indicates that a meta-language statement is being used to contribute a truth-value:

This is the approach Ede Zimmermann (pc) advocates and has been using in his classes.

(36) $$\vdash \alpha \dashv = \begin{cases} 1 \text{ if } \alpha \\ 0 \text{ if otherwise} \end{cases}$$

Lastly, one could abandon the H& K informal meta-language approach altogether and introduce a rigidly formalized meta-language.

These lecture notes will proceed to follow H& K's approach and will not introduce any further innovations. So, (33) is illicit and only (34) is acceptable.

1.5 Supplemental Readings

There is considerable overlap between this chapter and Chapter 12 of Heim & Kratzer's textbook:

Irene Heim & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Here, we approach intensional semantics from a different angle. It would probably be beneficial if you read H& K's Chapter 12 in addition to this chapter and if you did the exercises in there.

Come to think of it, some other ancillary reading is also recommended. You may want to look at relevant chapters in other textbooks:

David R. Dowty, Robert Wall & Stanley Peters. 1981. *Introduction to Montague semantics*. Kluwer. [Chapters 5& 6].

L. T. F. Gamut. 1991. *Logic, language, and meaning*. Chicago University Press. [Volume II: Intensional Logic and Logical Grammar].

Gennaro Chierchia & Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics (2nd edition)*. MIT Press. [Chapter 5: Intensionality].

Thomas Ede Zimmermann & Wolfgang Sternefeld. 2013. *Introduction to semantics: An essential guide to the composition of meaning*. de Gruyter Mouton.

The *Stanford encyclopedia of philosophy* is always a good resource. Here's the entry on possible worlds:

Christopher Menzel. 2016. Possible worlds. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Spring 2016. http://plato.stanford.edu/entries/possible-worlds/.

A couple of influential philosophical works on the metaphysics and uses of possible worlds:

Saul Kripke. 1980. *Naming and necessity*. Oxford: Blackwell.

David Lewis. 1986. *On the plurality of worlds*. Oxford: Blackwell.

An interesting paper on the origins of the modern possible worlds semantics for modal logic:

B. Jack Copeland. 2002. The genesis of possible worlds semantics. *Journal of Philosophical Logic* 31(2). 99–137. http://dx.doi.org/10.1023/A:1015273407895.

A personal history of formal semantics:

Barbara H. Partee. 2005. Reflections of a formal semanticist as of Feb 2005. ms. (longer version of introductory essay in 2004 book). http://people.umass.edu/partee/docs/BHP_Essay_Feb05.pdf.

A must read for students who plan to go on to becoming specialists in semantics, together with a handbook article putting it in perspective:

Richard Montague. 1973. The proper treatment of quantification in ordinary English. In Jaako Hintikka, Julius Moravcsik & Patrick Suppes (eds.), *Approaches to natural language*, 221–242. Reprinted in Portner & Partee (2002), pp. 17–34. Dordrecht: Reidel. http://www.blackwellpublishing.com/content/BPL_Images/Content_store/Sample_chapter/9780631215417/Portner.pdf.

Barbara H. Partee & Herman L.W. Hendriks. 1997. Montague grammar. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic and language*, 5–91. Elsevier.

To learn more about discourse about fiction, read Lewis:

David Lewis. 1978. Truth in fiction. *American Philosophical Quarterly* 15(1). Reprinted with postscripts in D. Lewis (1983), pp. 261–280, 37–46. http://www.jstor.org/stable/20009693.

Recent reconsiderations:

Andrea Bonomi & Sandro Zucchi. 2003. A pragmatic framework for truth in fiction. *Dialectica* 57(2). Preprint http://filosofia.dipafilo.unimi.it/~bonomi/Pragmatic.pdf, 103–120. http://dx.doi.org/10.1111/j.1746-8361.2003.tb00259.x.

Richard Hanley. 2004. As good as it gets: Lewis on truth in fiction. *Australasian Journal of Philosophy* 82(1). 112–128. http://dx.doi.org/10.1080/713659790.

Diane Proudfoot. 2006. Possible worlds semantics and fiction. 35(1). 9–40. http://dx.doi.org/10.1007/s10992-005-9005-8.

Inconsistencies in fictions and elsewhere are discussed in:

Achille Varzi. 1997. Inconsistency without contradiction. *Notre Dame Journal of Formal Logic* 38(4). 621–638. http://dx.doi.org/10.1305/ndjfl/1039540773.

David Lewis. 1982. Logic for equivocators. *Noûs* 16(3). Reprinted in D. Lewis (1998: pp. 97–110), 431–441. http://dx.doi.org/10.2307/2216219.

Some other interesting work on stories and pictures and their content:

Jeff Ross. 1997. *The semantics of media* (Studies in Linguistics and Philosophy (SLAP) 64). Dordrecht: Kluwer.

Sandro Zucchi. 2001. Tense in fiction. In Carlo Cecchetto, Gennaro Chierchia & Maria Teresa Guasti (eds.), *Semantic interfaces: Reference, anaphora and aspect*, 320–355. CSLI Publications. http://tinyurl.com/5ulwxwg.

Ben Blumson. 2009. Pictures, perspective and possibility. *Philosophical Studies*. http://dx.doi.org/10.1007/s11098-009-9337-2.

Astonishingly, Lewis' doctrine of the reality of the plurality of possible worlds is being paralleled (pun absolutely intended) by theoretical physicists in a number of ways. There is a controversial "many worlds" interpretation of quantum mechanics, for example. Other terms found are the "multiverse" and "parallel universes". See for starters, Kai's blog entry on a popular book on the issue, http://kaivonfintel.org/many-worlds/, MIT physics professor Max Tegmark's page on the topic, http://space.mit.edu/home/tegmark/crazy.html, and a Fresh Air interview with physicist Brian Greene, who just wrote a book called *The Hidden Reality: Parallel Universes and the Deep Laws of the Cosmos*: http://www.npr.org/2011/01/24/132932268/a-physicist-explains-why-parallel-universes-may-exist.

— This page intentionally left blank —

Chapter Two

Propositional Attitudes

We start with the basic possible worlds semantics for propositional attitude ascriptions. We talk briefly about the formal properties of accessibility relations.

2.1 Hintikka's Idea

Expressions like *believe*, *know*, *doubt*, *expect*, *regret*, and so on are usually said to describe PROPOSITIONAL ATTITUDES, expressing relations between individuals (the attitude holder) and propositions (intensions of sentences).

According to Hintikka (1969), the term PROPOSITIONAL ATTITUDE goes back to Russell (1940).

The simple idea is that *George believes that Henry is a spy* claims that George believes of the proposition that Henry is a spy that it is true. Note that for the attitude ascription to be true it does not have to hold that Henry is actually a spy. But where — in which world(s) — does Henry have to be a spy for it be true that George believes that Henry is a spy? We might want to be inspired by the colloquial phrase "in the world according to George" and say that *George believes that Henry is a spy* is true iff in the world according to George's beliefs, Henry is a spy. We immediately recall from Chapter 1 that we need to fix this idea up by making space for multiple worlds compatible with George's beliefs and by tying the truth-conditions to contingent facts about the evaluation world. That is, what George believes is different in different possible worlds.

Of course, the possible worlds semantics for propositional attitudes was in place long before the extension to fiction contexts was proposed. Our discussion here has inverted the historical sequence for pedagogical purposes.

The following lexical entry thus offers itself:

(1) $[\![\text{believe}]\!]^{w,g} =$
$$\lambda p_{\langle s,t\rangle}.\ \lambda x.\ \forall w' \text{ compatible with } x\text{'s beliefs in } w\text{: } p(w') = 1.$$

It is important to realize the modesty of this semantics: we are not trying to figure out what belief systems are and particularly not what their internal workings are. That is the job of psychologists (and philosophers of mind, perhaps). For our semantics, we treat the belief system as a black box that determines for each possible world whether it considers it possible that it is the world it is located in.

What is going on in this semantics? We conceive of George's beliefs as a state of his mind about whose internal structure we will remain agnostic, a matter left to other cognitive scientists. What we require of it is that it embody opinions about what the world he is located in looks like. In other words, if his beliefs are confronted with a particular possible world w', they will determine whether that world may or may not be the world as they think it is. What we are asking of George's mental state is whether any state of affairs, any event, anything in w' is in contradiction with anything that George believes. If not, then w' is compatible with George's beliefs. For all George believes, w' may well be the world where he lives. Many worlds will pass this criterion, just consider as one factor that George is unlikely to have any precise opinions about the number of leaves on the tree in front of my house. George's belief system determines a set of worlds compatible with his beliefs: those worlds that are viable candidates for being the actual world, as far as his belief system is concerned.

Now, George believes a proposition iff that proposition is true in all of the worlds compatible with his beliefs. If there is just one world compatible with his beliefs where the proposition is not true, that means that he considers it possible that the proposition is not true. In such a case, we can't say that he believes the proposition. Here is the same story in the words of Hintikka (1969), the source for this semantics for propositional attitudes:


Jaakko Hintikka

> My basic assumption (slightly simplified) is that an attribution of any propositional attitude to the person in question involves a division of all the possible worlds (...) into two classes: into those possible worlds which are in accordance with the attitude in question and into those which are incompatible with it. The meaning of the division in the case of such attitudes as knowledge, belief, memory, perception, hope, wish, striving, desire, etc. is clear enough. For instance, if what we are speaking of are (say) a's memories, then these possible worlds are all the possible worlds compatible with everything he remembers. [...]
>
> How are these informal observations to be incorporated into a more explicit semantical theory? According to what I have said, understanding attributions of the propositional attitude in question (...) means being able to make a distinction between two kinds of possible worlds, according to whether they are compatible with the relevant attitudes of the person in question. The semantical counterpart to this is of course a function which to a given individual person assigns a set of possible worlds.

> However, a minor complication is in order here. Of course, the person in question may himself have different attitudes in the different worlds we are considering. Hence this function in effect becomes a relation which to a given individual and to a given possible world μ associates a number of possible worlds which we shall call the ALTERNATIVES to μ. The relation will be called the alternativeness relation. (For different propositional attitudes, we have to consider different alternativeness relations.)

EXERCISE 2.1: Let's adopt Hintikka's idea that we can use a function that maps x and w into the set of worlds w' compatible with what x believes in w. Call this function $\mathcal{B}$. That is,

(2) $\mathcal{B} = \lambda x.\ \lambda w.\ \{w'\colon\ w' \text{ is compatible with what } x \text{ believes in } w\}.$

Using this notation, our lexical entry for *believe* would look as follows:

(3) $[\![\text{believe}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \lambda x.\ \mathcal{B}(x)(w) \subseteq p.$

We are here indulging in the usual sloppiness in treating p both as a function from worlds to truth-values and as the set characterized by that function.

Here now are two "alternatives" for the semantics of *believe*:

(4) ATTEMPT 1 (VERY WRONG)
$[\![\text{believe}]\!]^{w,g} = \lambda p \in D_{\langle s,t\rangle}.\big[\lambda x \in D.\ p = \mathcal{B}(x)(w)\big].$

(5) ATTEMPT 2 (ALSO VERY WRONG)
$[\![\text{believe}]\!]^{w,g} = \lambda p \in D_{\langle s,t\rangle}.\big[\lambda x \in D.\ p \cap \mathcal{B}(x)(w) \neq \emptyset\big].$

Explain why these do not adequately capture the meaning of *believe*. □

EXERCISE 2.2: Follow-up: The semantics in (5) would have made *believe* into an existential quantifier of sorts: it would say that *some* of the worlds compatible with what the subject believes are such-and-such. You have argued (successfully, of course) that such an analysis is wrong for *believe*. But *are* there attitude predicates with such an "existential" meaning? Discuss some candidates. If you can't find any candidates that survive scrutiny, can you speculate why there might be no existential attitude predicates? [Warning: this is underexplored territory!]□

$\mathcal{BS}$ is meant to stand for '**b**elief **s**tate', not for what you might have thought!

We can also think of belief states as being represented by a function $\mathcal{BS}$, which maps an individual and a world into a set of propositions: those that the individual believes. From there, we could calculate the set of worlds compatible with an individual x's beliefs in world w by retrieving the set of those possible worlds in which all of the propositions in $\mathcal{BS}(x)(w)$ are true: $\{w' : \forall p \in \mathcal{BS}(x)(w) : p(w') = 1\}$, which in set talk is simply the big intersection of all the propositions in the set: $\cap\mathcal{BS}(x)(w)$. Our lexical entry then would be:

(6) $[\![\text{believe}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \lambda x.\ \cap\,\mathcal{BS}(x)(w) \subseteq p.$

Exercise 2.3: Imagine that our individual x forms a new opinion. Imagine that we model this by adding a new proposition p to the pool of opinions. So, $\mathcal{BS}(x)(w)$ now contains one further element. There are now more opinions. What happens to the set of worlds compatible with x's beliefs? Does it get bigger or smaller? Is the new set a subset or superset of the previous set of compatible worlds? □

2.2 Accessibility Relations

Another way of reformulating Hintikka's semantics for propositional attitudes is via the notion of an accessibility relation. We talk of a world w' being accessible from w. Each attitude can be associated with such an accessibility relation. For example, we can introduce the relation $w\mathcal{R}^{\mathcal{B}}_{a}w'$ which holds iff w' is compatible with a's belief state in w. We have then yet another equivalent way of specifying the lexical entry for *believe*:

(7) $[\![\text{believe}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \lambda x.\ \forall w' : w\mathcal{R}^{\mathcal{B}}_{x}w' \rightarrow p(w') = 1.$

It is profitable to think of different attitudes (belief, knowledge, hope, regret, memory, ...) as corresponding to different accessibility relations. Recall now that the linguistic study of determiners benefitted quite a bit from an investigation of the formal properties of the relations between sets of individuals that determiners express. We can do the same thing here and ask about the formal properties of the accessibility relation associated with belief versus the one associated with knowledge, etc. The obvious properties to think about are reflexivity, transitivity, and symmetry.

Kirill Shklovsky (in class) asked why we call reflexivity, transitivity, and symmetry "formal" properties of relations. The idea is that certain properties are "formal" or "logical", while others are more substantial. So, the fact that the relation "have the same birthday as" is symmetric seems a more formal fact about it than the fact that the relation holds between my daughter and my brother-in-law. Nevertheless, one of the most common ways of characterizing formal/logical notions (permutation-invariance, if you're curious) does not in fact make symmetry etc. a formal/logical notion. So, while intuitively these do seem to be formal/logical properties, we do not know how to substantiate that intuition. See MacFarlane (2005) for discussion.

2.2.1 Reflexivity

A relation is reflexive iff for any object in the domain of the relation we know that the relation holds between that object and itself. Which accessibility relations are reflexive? Take knowledge:

(8) $w\mathcal{R}^{\mathcal{K}}_{x}w'$ iff w' is compatible with what x knows in w.

We are asking whether for any given possible world w, we know that $\mathcal{R}_x^{\mathcal{K}}$ holds between w and w itself. It will hold if w is a world that is compatible with what we know in w. And clearly that must be so. Take our body of knowledge in w. The concept of knowledge crucially contains the concept of truth: what we know must be true. So if in w we know that something is the case then it must be the case in w. So, w must be compatible with all we know in w. $\mathcal{R}_x^{\mathcal{K}}$ is reflexive.

We talk here about knowledge entailing (or even presupposing) truth but we do not mean to say that knowledge simply equals true belief. Professors Socrates and Gettier and their exegetes have further considerations.

Now, if an attitude X corresponds to a reflexive accessibility relation, then we can conclude from α Xs *that* p being true in w that p is true in w. This property of an attitude predicate is often called VERIDICALITY. It is to be distinguished from FACTIVITY, which is a property of attitudes which *presuppose* – rather than (merely) entail – the truth of their complement.

In modal logic notation: $\Box p \rightarrow p$. This pattern is sometimes called **T** or **M**, as is the corresponding system of modal logic.

If we consider the relation $\mathcal{R}_x^{\mathcal{B}}$ pairing with a world w those worlds w' which are compatible with what x *believes* in w, we no longer have reflexivity: belief is not a veridical attitude. It is easy to have false beliefs, which means that the actual world is not in fact compatible with one's beliefs, which contradicts reflexivity. And many other attitudes as well do not involve veridicality/reflexivity: what we hope may not come true, what we remember may not be what actually happened, etc.

The difference between *believe* and *know* in natural discourse is quite delicate, especially when one considers first person uses (*I believe the earth is flat* vs. *I know the earth is flat*).

In modal logic, the correspondence between formal properties of the accessibility relation and the validity of inference patterns is well-studied. What we have just seen is that reflexivity of the accessibility relation corresponds to the validity of $\Box p \rightarrow p$. Other properties correspond to other characteristic patterns. Let's see this for transitivity and symmetry.

2.2.2 *Transitivity

Transitivity of the accessibility relation corresponds to the inference $\Box p \rightarrow \Box\Box p$. The pattern seems not obviously wrong for knowledge: if one knows that p, doesn't one thereby know that one knows that p? But before we comment on that, let's establish the formal correspondence between transitivity and that inference pattern. This needs to go in both directions.

In the literature on epistemic modal logic, the pattern is known as the KK THESIS or POSITIVE INTROSPECTION. In general modal logic, it is the characteristic axiom **4** of the modal logic system **S4**, which is a system that adds **4** to the previous axiom **M/T**. Thus, **S4** is the logic of accessibility relations that are both reflexive and transitive.

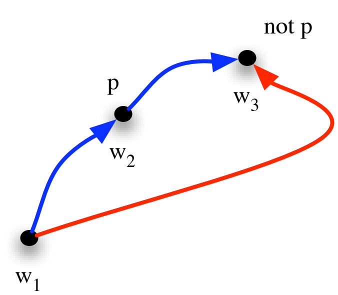

Figure 2.1: Transitivity

What does it take for the pattern to be valid? Assume that $\Box$p holds for an arbitrary world w, i.e. that p is true in all worlds w' accessible from w. Now, the inference is to the fact that p again holds in any world w'' accessible from any of those worlds w' accessible from w. But what would prevent p from being false in some w'' accessible from some w' accessible from w? That could only be prevented from happening if we knew that w'' itself is accessible from w as well, because then we would know from the premiss that p is true in it (since p is true in *all* worlds accessible from w). Ah, but w'' (some world accessible from a world w' accessible from w) is only guaranteed to be accessible from w if the accessibility relation is transitive (if w' is accessible from w and w'' is accessible from w', then transitivity ensures that w'' is accessible from w). This reasoning has shown that validity of the pattern requires transitivity. The other half of proving the correspondence is to show that transitivity entails that the pattern is valid.

The proof proceeds by reductio. Assume that the accessibility relation is transitive. Assume that (i) $\Box$p holds for some world w but that (ii) $\Box\Box$p doesn't hold in w. We will show that this situation cannot obtain. By (i), p is true in all worlds w' accessible from w. By (ii), there is some non-p world w'' accessible from some world w' accessible from w. But by transitivity of the accessibility relation, that non-p world w'' must be accessible from w. And since *all* worlds accessible from w are p worlds, w'' must be a p world, in contradiction to (ii). So, as soon as we assume transitivity, there is no way for the inference not to go through.

Now, do any of the attitudes have the transitivity property? It seems rather obvious that as soon as you believe something, you thereby believe that you believe it (and so it seems that belief involves a transitive accessibility relation). And in fact, as soon as you believe something, you believe that you *know* it. But one might shy away from saying that knowing something automatically amounts to knowing that you know it. For example, many are attracted to the idea that to know something requires that (i) that it is true, (ii) that you believe it, and (iii) that you are justified in believing it: the justified true belief analysis of knowledge. So, now couldn't it be that you know something, and thus (?) that you believe you know it, and thus that you believe that you are justified in believing it, but that you are not justified in believing that you are *justified* in believing it? After all, one's source of knowledge, one's reliable means of acquiring knowledge, might be a mechanism that one has no insight into. So, while one can implicitly trust (believe) in its reliability, and while it is in fact reliable, one might not have any means to have trustworthy beliefs about it. [Further worries about the KK Thesis are discussed by Williamson (2000).]

2.2.3 *Symmetry

What would the consequences be if the accessibility relation were symmetric? Symmetry of the accessibility relation $\mathcal{R}$ corresponds to the validity of the following principle:

(9) Brouwer's Axiom:

$$\forall p \forall w: \ w \in p \to \left[\forall w' \left[w \mathcal{R} w' \to \exists w'' \left[w' \mathcal{R} w'' \& \ w'' \in p\right]\right]\right]$$

In modal logic notation: $p \to \Box\Diamond p$, known simply as B in modal logic. The system that combines **T**/**M** with B is often called Brouwer's System (**B**), after the mathematician L.E.J. Brouwer, not because he proposed it but because it was thought that it had some connections to his doctrines.

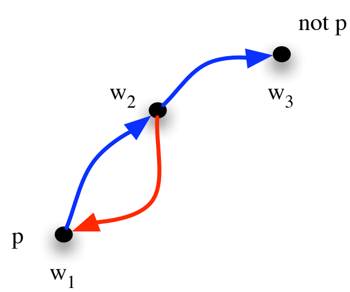

Figure 2.2: Symmetry

L.E.J. Brouwer

Here's the reasoning: Assume that R is in fact symmetric. Pick a world w in which p is true. Now, could it be that the right hand side of the inference fails to hold in w? Assume that it does fail. Then, there must be some world w' accessible from w in which $\Diamond p$ is false. In other words, from that world w' there is no accessible world w'' in which p is true. But since R is assumed to be symmetric, one of the worlds accessible from w' is w and in w, p is true, which contradicts the assumption that the inference doesn't go through. So, symmetry ensures the validity of the inference.

The other way (validity of the inference requires symmetry): the inference says that from any p world we only have worlds accessible from which there is at least one accessible p world. But imagine that p is true in w but not true in any other world. So, the only way for the conclusion of the inference to hold automatically is to have a guarantee that w (the only p world) is accessible from any world accessible from it. That is, we need to have symmetry. QED.

To see whether a particular kind of attitude is based on a symmetric accessibility relation, we can ask whether Brouwer's Axiom is intuitively valid with respect to this attitude. If it is not valid, this shows that the accessibility relation can't be symmetric. In the case of a knowledge-based accessibility relation (epistemic accessibility), one can argue that *symmetry does not hold*:[1]

> The symmetry condition would imply that if something happens to be true in the actual world, then you know that it is compatible

1 Thanks to Bob Stalnaker (pc to Kai von Fintel) for help with the following reasoning.

> with your knowledge (Brouwer's Axiom). This will be violated by any case in which your beliefs are consistent, but mistaken. Suppose that while p is in fact true, you feel certain that it is false, and so think that you know that it is false. Since you think you know this, it is compatible with your knowledge that you know it. (Since we are assuming you are consistent, you can't both believe that you know it, and know that you do not). So it is compatible with your knowledge that you know that *not* p. Equivalently[2]: you don't know that you don't know that *not* p. Equivalently: you don't know that it's compatible with your knowledge that p. But by Brouwer's Axiom, since p is true, you would have to know that it's compatible with your knowledge that p. So if Brouwer's Axiom held, there would be a contradiction. So Brouwer's Axiom doesn't hold here, which shows that epistemic accessibility is not symmetric.

Game theorists and theoretical computer scientists who traffic in logics of knowledge often assume that the accessibility relation for knowledge is an equivalence relation (reflexive, symmetric, and transitive). But this is appropriate only if one abstracts away from any error, in effect assuming that belief and knowledge coincide. One striking consequence of working with an equivalence relation as the accessibility relation for knowledge is that one predicts the principle of NEGATIVE INTROSPECTION to hold:

All one really needs to make **NI** valid is to have a EUCLIDEAN accessibility relation: any two worlds accessible from the same world are accessible from each other. It is a nice little exercise to prove this, if you have become interested in this sort of thing. Note that all reflexive and Euclidean accessibility relations are transitive and symmetric as well — another nice little thing to prove.

(10) NEGATIVE INTROSPECTION (NI)
If one doesn't know that p, then one knows that one doesn't know that p. ($\neg\Box p \rightarrow \Box\neg\Box p$).

This surely seems rather dubious: imagine that one strongly believes that p but that nevertheless p is false, then one doesn't know that p, but one doesn't seem to believe that one doesn't know that p, in fact one believes that one does know that p.

2.3 Supplemental Readings

A recent survey on attitudes:

Eric Swanson. 2011. Propositional attitudes. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of meaning*, vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft 33.2), chap. 60, 1538–1561. Berlin/Boston: de Gruyter Mouton. http://dx.doi.org/10.1515/9783110255072.1538. http://tinyurl.com/swanson-hsk.

2 This and the following step rely on the duality of necessity and possibility: q is compatible with your knowledge iff you don't know that *not* q.

Further connections between mathematical properties of accessibility relations and logical properties of various notions of necessity and possibility are studied extensively in modal logic:

G.E. Hughes & M.J. Cresswell. 1996. *A new introduction to modal logic*. London: Routledge.

James Garson. 2008. Modal logic. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/logic-modal/, especially section 7 and 8, "Modal Axioms and Conditions on Frames", "Map of the Relationships between Modal Logics".

A thorough discussion of the possible worlds theory of attitudes, and some of its potential shortcomings, can be found in Bob Stalnaker's work:

Robert Stalnaker. 1984. *Inquiry*. MIT Press.

Robert Stalnaker. 1999. *Context and content*. Oxford: Oxford University Press.

A quick and informative surveys about the notion of knowledge:

Matthias Steup. 2008. The analysis of knowledge. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Fall 2008. http://plato.stanford.edu/archives/fall2008/entries/knowledge-analysis/.

Linguistic work on attitudes has often been concerned with various co-occurrence patterns, particularly which moods (indicative or subjunctive or infinitive) occur in the complement and whether negative polarity items are licensed in the complement.

Mood licensing:

Paul Portner. 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5(2). 167–212. http://dx.doi.org/10.1023/A:1008280630142.

NPI-Licensing:

Nirit Kadmon & Fred Landman. 1993. *Any*. *Linguistics and Philosophy* 16(4). 353–422. http://dx.doi.org/10.1007/BF00985272.

Kai von Fintel. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148. http://dx.doi.org/10.1093/jos/16.2.97.

Anastasia Giannakidou. 1999. Affective dependencies. *Linguistics and Philosophy* 22(4). 367–421. http://dx.doi.org/10.1023/A:1005492130684.

There is some interesting work out of Amherst rethinking the way attitude predicates take their complements:

Angelika Kratzer. 2006. Decomposing attitude verbs. Handout from a talk honoring Anita Mittwoch on her 80th birthday at the Hebrew University of Jerusalem July 4, 2006. http://semanticsarchive.net/Archive/DcwY2JkM/attitude-verbs2006.pdf.

Keir Moulton. 2008. Clausal complementation and the *Wager*-class. *North East Linguistics Society (NELS)* 38. http://sites.google.com/site/keirmoulton/Moultonnels2008wager.pdf. http://people.umass.edu/keir/Wager.pdf.

Keir Moulton. 2009. *Natural selection and the syntax of clausal complementation*. University of Massachusetts at Amherst dissertation. http://scholarworks.umass.edu/open_access_dissertations/99/.

Tamina Stephenson in her MIT dissertation and related work explores the way attitude predicates interact with epistemic modals and taste predicates in their complements:

Tamina Stephenson. 2007a. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525. http://dx.doi.org/10.1007/s10988-008-9023-4.

Tamina Stephenson. 2007b. *Towards a theory of subjective meaning*. Massachusetts Institute of Technology dissertation. http://semanticsarchive.net/Archive/2QxMjk0O/Stephenson-2007-thesis.pdf.

Jon Gajewski in his MIT dissertation and subsequent work explores the distribution of the NEG-RAISING property among attitude predicates and traces it back to presuppositional components of the meaning of the predicates:

Jon Gajewski. 2005. *Neg-raising: Polarity and presupposition*. Massachusetts Institute of Technology dissertation. http://dx.doi.org/1721.1/33696.

Jon Gajewski. 2007. Neg-raising and polarity. *Linguistics and Philosophy*. http://dx.doi.org/10.1007/s10988-007-9020-z.

Interesting work has also been done on presupposition projection in attitude contexts:

Nicholas Asher. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10(2). 125–197. http://dx.doi.org/10.1007/BF00584317.

Irene Heim. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Bart Geurts. 1998. Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21(6). 545–601. http://dx.doi.org/10.1023/A:1005481821597.

Chapter Three

Modality

We turn to modal auxiliaries and related constructions. We see more context-dependency. We still quantify over possible worlds.

3.1 The Quantificational Theory of Modality

We will now be looking at modal auxiliaries like *may, must, can, have to*, etc. Most of what we say here should carry over straightforwardly to modal adverbs like *maybe, possibly, certainly*, etc. We will make certain syntactic assumptions, which make our work easier but which leave aside many questions that at some point deserve to be addressed.

3.1.1 Syntactic Assumptions

We will assume, at least for the time being, that a modal like *may* is a RAISING predicate (rather than a CONTROL predicate), i.e., its subject is not its own argument, but has been moved from the subject-position of its infinitival complement. So, we are dealing with the following kind of structure:

The issue of raising vs. control will probably be taken up later. If you are eager to get started on it and other questions of the morphosyntax of modals, read the handout from an LSA class Sabine and Kai taught a few years ago: http://web.mit.edu/fintel/lsa220-class-2-handout.pdf.

(1) a. Ann may be smart.
b. [Ann [λ_1 [may [t_1 be smart]]]]

Actually, we will be working here with the even simpler structure below, in which the subject has been reconstructed to its lowest trace position. (E.g., these could be generated by deleting all but the lowest copy in the movement chain.) We will be able to prove that movement of a name or pronoun never affects truth-conditions, so at any rate the interpretation of the structure in (1b) would be the same as that of (2). As a matter of convenience, then, we will take the reconstructed structures, which allow us to abstract away from the (here irrelevant) mechanics of variable binding.

We will talk about reconstruction in more detail later.

(2) may [Ann be smart]

So, for now at least, we are assuming that modals are expressions that take a full sentence as their semantic argument.[1] Now then, what do modals mean?

3.1.2 Quantification over Possible Worlds

The basic idea of the possible worlds semantics for modal expressions is that they are quantifiers over possible worlds. Toy lexical entries for *must* and *may*, for example, would look like this:

This idea goes back a long time. It was famously held by Leibniz, but there are precedents in the medieval literature; see **knuuttila:2003:modality-medieval**. See Copeland (2002) for the modern history of the possible worlds analysis of modal expressions.

(3) $[\![\text{must}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \forall w'\colon\ p(w') = 1.$

(4) $[\![\text{may}]\!]^{w,g} = \lambda p_{\langle s,t\rangle}.\ \exists w'\colon\ p(w') = 1.$

This analysis is too crude (in particular, notice that it would make modal sentences non-contingent — there is no occurrence of the evaluation world on the right hand side!). But it does already have some desirable consequences that we will seek to preserve through all subsequent refinements. It correctly predicts a number of intuitive judgments about the logical relations between *must* and *may* and among various combinations of these items and negations. To start with some elementary facts, we feel that *must* ϕ entails *may* ϕ, but not vice versa:

1 We will assume that even though *Ann be smart* is a non-finite sentence, this will not have any effect on its semantic type, which is that of a sentence, which in turn means that its semantic value is a truth-value. This is hopefully independent of the (interesting) fact that *Ann be smart* on its own cannot be used to make a truth-evaluable assertion.

(5) You must stay.
Therefore, you may stay. VALID

(6) You may stay.
Therefore, you must stay. INVALID

(7) a. You may stay, but it is not the case that you must stay.[2]
b. You may stay, but you don't have to stay. CONSISTENT

We judge *must* ϕ incompatible with its "inner negation" *must* [*not* ϕ], but find *may* ϕ and *may* [*not* ϕ] entirely compatible:

(8) You must stay, and/but also, you must leave. (leave = not stay).
CONTRADICTORY

(9) You may stay, but also, you may leave.
CONSISTENT

We also judge that in each pair below, the (a)-sentence and the (b)-sentences say the same thing.

(10) a. You must stay.
b. It is not the case that you may leave.
You aren't allowed to leave.
(You may not leave.)[3]
(You can't leave.)

(11) a. You may stay.
b. It is not the case that you must leave.
You don't have to leave.
You don't need to leave.
(You needn't leave.)

---

2 The somewhat stilted *it is not the case*-construction is used in (a) to make certain that negation takes scope over *must.* When modal auxiliaries and negation are together in the auxiliary complex of the same clause, their relative scope seems not to be transparently encoded in the surface order; specifically, the scope order is not reliably negation $\succ$ modal. (Think about examples with *mustn't, can't, shouldn't, may not* etc. What's going on here? This is an interesting topic which we must set aside for now. See the references at the end of the chapter for relevant work.) With modal *main* verbs (such as *have to*), this complication doesn't arise; they are consistently inside the scope of clause-mate auxiliary negation. Therefore we can use (b) to (unambiguously) express the same scope order as (a), without having to resort to a biclausal structure.

3 The parenthesized variants of the (b)-sentences are pertinent here only to the extent that we can be certain that negation scopes over the modal. In these examples, apparently it does, but as we remarked above, this cannot be taken for granted in all structures of this form.

Given that *stay* and *leave* are each other's negations (i.e. $[\![\text{leave}]\!]^{w,g} = [\![\text{not stay}]\!]^{w,g}$, and $[\![\text{stay}]\!]^{w,g} = [\![\text{not leave}]\!]^{w,g}$), the LF-structures of these equivalent pairs of sentences can be seen to instantiate the following schemata:[4]

(12) a. *must* $\phi \equiv$ *not* [*may* [*not* ϕ]]
b. *must* [*not* ψ] $\equiv$ *not* [*may* ψ]

(13) a. *may* $\phi \equiv$ *not* [*must* [*not* ϕ]]
b. *may* [*not* ψ] $\equiv$ *not* [*must* ψ]

Our present analysis of *must*, *have-to*, … as universal quantifiers and of *may*, *can*, … as existential quantifiers straightforwardly predicts all of the above judgments, as you can easily prove.

More linguistic data regarding the "parallel logic" of modals and quantifiers can be found in Larry Horn's dissertation (**horn:1972:dissertation**).

(14) a. $\forall x \phi \equiv \neg \exists x \neg \phi$
b. $\forall x \neg \phi \equiv \neg \exists x \phi$

(15) a. $\exists x \phi \equiv \neg \forall x \neg \phi$
b. $\exists x \neg \phi \equiv \neg \forall x \phi$

3.2 Flavors of Modality

3.2.1 Contingency

We already said that the semantics we started with is too simple-minded. In particular, we have no dependency on the evaluation world, which would make modal statements non-contingent. This is not correct.

If one says *It may be snowing in Cambridge*, that may well be part of useful, practical advice about what to wear on your upcoming trip to Cambridge. It may be true or it may be false. The sentence seems true if said in the dead of winter when we have already heard about a Nor'Easter that is sweeping across New England. The sentence seems false if said by a clueless Australian acquaintance of ours in July.

The contingency of modal claims is not captured by our current semantics. All the *may*-sentence would claim under that semantics is that there is some possible world where it is snowing in Cambridge. And surely, once you have read Lewis' quote in Chapter 1, where he asserts the existence of possible worlds with different physical constants than we enjoy here, you must admit that there have to be such worlds even if it is July. The problem is that in our semantics, repeated here

(16) $[\![\text{may}]\!]^{w,g} = \lambda p_{\langle s,t \rangle}.\ \exists w' \colon\ p(w') = 1.$

4 In logicians' jargon, *must* and *may* behave as duals of each other. For definitions of "dual", see **barwise-cooper:1981:generalized** or Gamut (1991: vol.2,238).

there is no occurrence of w on the right hand side. This means that the truth-conditions for *may*-sentences are world-independent. In other words, they make non-contingent claims that are either true whatever or false whatever, and because of the plenitude of possible worlds they are more likely to be true than false. This needs to be fixed. But how?

Conversely, the plenitude of possible worlds would make *must*-claims very likely false if they are not reigned in or anchored somehow.

Well, what makes *it may be snowing in Cambridge* seem true when we know about a Nor'Easter over New England? What makes it seem false when we know that it is summer in New England? The idea is that we only consider possible worlds COMPATIBLE WITH THE EVIDENCE AVAILABLE TO US. And since what evidence is available to us differs from world to world, so will the truth of a *may*-statement.

(17) $[\![\text{may}]\!]^{w,g} = \lambda p.\ \exists w'$ compatible with the evidence in w: $p(w') = 1$.[5]

(18) $[\![\text{must}]\!]^{w,g} = \lambda p.\ \forall w'$ compatible with the evidence in w: $p(w') = 1$.

Let us consider a different example:

(19) You have to be quiet.

Imagine this sentence being said based on the house rules of the particular dormitory you live in. Again, this is a sentence that could be true or could be false. Why do we feel that this is a contingent assertion? Well, the house rules can be different from one world to the next, and so we might be unsure or mistaken about what they are. In one possible world, they say that all noise must stop at 11pm, in another world they say that all noise must stop at 10pm. Suppose we know that it is 10:30 now, and that the dorm we are in has either one or the other of these two rules, but we have forgotten which. Then, for all we know, *you have to be quiet* may be true or it may be false. This suggests a lexical entry along these lines:

(20) $[\![\text{have-to}]\!]^{w,g} = \lambda p.\ \forall w'$ compatible with the rules in w: $p(w') = 1$.

Again, we are tying the modal statement about other worlds down to certain worlds that stand in a certain relation to actual world: those worlds where the rules as they are here are obeyed.

A note of caution: it is very important to realize that the worlds compatible with the rules as they are in w are those worlds where nothing happens that violates any of the w-rules. This is not at all the same as saying that the worlds compatible with the rules in w are those worlds where the same rules are in force. Usually, the rules do not care what the rules are, unless the rules contain some kind of meta-statement to the effect that the rules have to be the way they are, i.e. that the rules cannot be changed. So, in fact, a world w' in which

5 From now on, we will leave off type-specifications such as that p has to be of type $\langle s, t \rangle$, whenever it is obvious what they should be and when saving space is aesthetically called for.

nothing happens that violates the rules as they are in w but where the rules are quite different and in fact what happens violates the rules as they are in w' is nevertheless a world compatible with the rules in w. For example, imagine that the only relevant rule in w is that students go to bed before midnight. Take a world w' where a particular student goes to bed at 11:30 pm but where the rules are different and say that students have to go to bed before 11 pm. Such a world w' is compatible with the rules in w (but of course not with the rules in w').

Apparently, there are different flavors of modality, varying in what kind of facts in the evaluation world they are sensitive to. The semantics we gave for *must* and *may* above makes them talk about evidence, while the semantics we gave for *have-to* made it talk about rules. But that was just because the examples were hand-picked. In fact, in the dorm scenario we could just as well have said *You must be quiet.* And, vice versa, there is nothing wrong with using *it has to be snowing in Cambridge* based on the evidence we have. In fact, many modal expressions seem to be multiply ambiguous.

Traditional descriptions of modals often distinguish a number of "readings": epistemic, deontic, ability, circumstantial, dynamic, ... (Beyond "epistemic" and "deontic," there is a great deal of terminological variety. Sometimes all non-epistemic readings are grouped together under the term root modality.) Here are some initial illustrations.

(21) Epistemic Modality

A: Where is John?
B: I don't know. He *may* be at home.

(22) Deontic Modality

A: Am I allowed to stay over at Janet's house?
B: No, but you *may* bring her here for dinner.

(23) Circumstantial/Dynamic Modality

A: I will plant the rhododendron here.
B: That's not a good idea. It *can* grow very tall.

How are *may* and *can* interpreted in each of these examples? What do the interpretations have in common, and where do they differ?

In all three examples, the modal makes an existentially quantified claim about possible worlds. This is usually called the modal force of the claim. What differs is what worlds are quantified over. In epistemic modal sentences, we quantify over worlds compatible with the available evidence. In deontic modal sentences, we quantify over worlds compatible with the rules and/or regulations. And in the circumstantial modal sentence, we quantify over the set of worlds which conform to the laws of nature (in particular, plant biology). What speaker B in (23) is saying, then, is that there are some worlds conforming to the laws of nature in which this rhododendron grows very tall. (Or is this another instance of an epistemic reading? See below for discussion of the distinction between circumstantial readings and epistemic ones.)

Angelika Kratzer

How can we account for this variety of readings? One way would be to write a host of lexical entries, basically treating this as a kind of (more or less principled) ambiguity. Another way, which is preferred by many people, is to treat this as a case of context-dependency, as argued in seminal work by **kratzer:1977:must-can**, **kratzer:1978:dissertation**, Kratzer (1981, 1991).

According to Kratzer, what a modal brings with it intrinsically is just a modal force, that is, whether it is an existential (possibility) modal or a universal (necessity) modal. What worlds it quantifies over is determined by context. In essence, the context has to supply a restriction to the quantifier. How can we implement this idea?

It is well-known that natural language quantification is in general subject to contextual restriction. See **stanley-szabo:2000:restriction** for a recent discussion.

We encountered context-dependency before when we talked about pronouns and their referential (and E-Type) readings (H&K, chapters 9–11). We treated referential pronouns as free variables, appealing to a general principle that free variables in an LF need to be supplied with values from the utterance context. If we want to describe the context-dependency of modals in a technically analogous fashion, we can think of their LF-representations as incorporating or subcategorizing for a kind of invisible pronoun, a free variable that stands for a set of possible worlds. So we posit LF-structures like this:

Warning: This account is problematic and will be refined soon.

(24) $[_{I'}\ [_{I}\ \text{must}\ p_{\langle n,\langle s,t\rangle\rangle}\ ]\ [_{VP}\ \text{you quiet}]]$

$p_{\langle n,\langle s,t\rangle\rangle}$ here is a variable over (characteristic functions of) sets of worlds, which — like all free variables — needs to receive a value from the utterance context. Possible values include: the set of worlds compatible with the speaker's current knowledge; the set of worlds in which everyone obeys all the house rules of a certain dormitory; and many others. The denotation of the modal itself now has to be of type $\langle st, \langle st, t\rangle\rangle$ rather than $\langle st, t\rangle$, thus it will be more like a quantificational determiner rather than a complete generalized quantifier. Only after the modal has been combined with its covert restrictor do we obtain a value of type $\langle st, t\rangle$.

We are using the notation for variables of types other than e introduced by Heim & Kratzer. See p. 213. An index on a variable now is an ordered pair of a natural number and a type.

Q: Can you think of overt anaphoric expressions that are arguably of the type $\langle s, t\rangle$, a proposition?

(25) a. $[\![\text{must}]\!]^{w,g} = [\![\text{have-to}]\!]^{w,g} = [\![\text{need-to}]\!]^{w,g} = \ldots =$
$\lambda p \in D_{\langle s,t\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \forall w \in W\ [p(w) = 1 \to q(w) = 1]$

in set talk: $p \subseteq q$

b. $[\![\text{may}]\!]^{w,g} = [\![\text{can}]\!]^{w,g} = [\![\text{be-allowed-to}]\!]^{w,g} = \ldots =$
$\lambda p \in D_{\langle s,t\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \exists w \in W\ [p(w) = 1\ \&\ q(w) = 1]$

in set talk: $p \cap q \neq \emptyset$

On this approach, the epistemic, deontic, etc. "readings" of individual occurrences of modal verbs come about by a combination of two separate things. The lexical semantics of the modal itself encodes just a quantificational force, a *relation* between sets of worlds. This is either the subset-relation (universal quantification; necessity) or the relation of non-disjointness (existential quantification; possibility). The covert variable next to the modal picks up a contextually salient set of worlds, and this functions as the quantifier's restrictor. The labels "epistemic", "deontic", "circumstantial" etc. group together certain conceptually natural classes of possible values for this covert restrictor.

Notice that, strictly speaking, there is not just one deontic reading (for example), but many. A speaker who utters

(26) You have to be quiet.

might mean: 'I want you to be quiet,' (i.e., you are quiet in all those worlds that conform to my preferences). Or she might mean: 'unless you are quiet, you won't succeed in what you are trying to do,' (i.e., you are quiet in all those worlds in which you succeed at your current task). Or she might mean: 'the house rules of this dormitory here demand that you be quiet,' (i.e., you are quiet in all those worlds in which the house rules aren't violated). And so on. So the label "deontic" appears to cover a whole open-ended set of imaginable "readings", and which one is intended and understood on a particular utterance occasion may depend on all sorts of things in the interlocutors' previous conversation and tacit shared assumptions. (And the same goes for the other traditional labels.)

3.2.2 Epistemic vs. Circumstantial Modality

Is it all context-dependency? Or do flavors of modality correspond to some sorts of signals in the structure of sentences? Read the following famous passage from Kratzer and think about how the two sentences with their very different modal meanings differ in structure:

Quoted from Kratzer (1991). In Kratzer (1981), the hydrangeas were *Zwetschgenbäume* 'plum trees'. The German word *Zwetschge*, by the way, is etymologically derived from the name of the city Damascus (Syria), the center of the ancient plum trade.

> Consider sentences (27) and (28):
>
> (27) Hydrangeas can grow here.
>
> (28) There might be hydrangeas growing here.
>
> The two sentences differ in meaning in a way which is illustrated by the following scenario.
>
> "Hydrangeas"
>
> Suppose I acquire a piece of land in a far away country and discover that soil and climate are very much like at home, where hydrangeas prosper everywhere. Since hydrangeas are my favorite plants, I wonder whether they would grow in this place and inquire about it. The answer is (27). In such a situation, the proposition expressed by (27) is true. It is true regardless of whether it is or isn't likely that there are already hydrangeas in the country we are considering. All that matters is climate, soil, the special properties of hydrangeas, and the like. Suppose now that the country we are in

> has never had any contacts whatsoever with Asia or America, and the vegetation is altogether different from ours. Given this evidence, my utterance of (28) would express a false proposition. What counts here is the complete evidence available. And this evidence is not compatible with the existence of hydrangeas.
>
> (27) together with our scenario illustrates the pure CIRCUMSTANTIAL reading of the modal *can*. [...]. (28) together with our scenario illustrates the epistemic reading of modals. [...] circumstantial and epistemic conversational backgrounds involve different kinds of facts. In using an epistemic modal, we are interested in what else may or must be the case in our world given all the evidence available. Using a circumstantial modal, we are interested in the necessities implied by or the possibilities opened up by certain sorts of facts. Epistemic modality is the modality of curious people like historians, detectives, and futurologists. Circumstantial modality is the modality of rational agents like gardeners, architects, and engineers. A historian asks what might have been the case, given all the available facts. An engineer asks what can be done given certain relevant facts.

Consider also the very different prominent meanings of the following two sentences, taken from Kratzer as well:

(29) a. Cathy can make a pound of cheese out of this can of milk.
b. Cathy might make a pound of cheese out of this can of milk.

Exercise 3.1: Come up with examples of epistemic, deontic, and circumstantial uses of the necessity verb *have to*. Describe the set of worlds that constitutes the understood restrictor in each of your examples. □

3.2.3 Contingency Again

We messed up. If you inspect the context-dependent meanings we have on the table now for our modals, you will see that the right hand sides again do not mention the evaluation world w. Therefore, we will again have the problem of not making contingent claims, indirectly about the actual world. This needs to be fixed. We need a semantics that is both context-dependent and contingent.

The problem, it turns out, is with the idea that the utterance context supplies a *determinate set of worlds* as the restrictor. When I understand that you meant your use of *must*, in *you must be quiet*, to quantify over the set of worlds in which the house rules of our dorm are obeyed, this does not imply that you and I have to know or agree on which set exactly this is. That depends on what the house rules in our world actually happen to say, and this may be an open question at the current stage of our conversation. What we do agree on, if I have understood your use of *must* in the way that you intended it, is just that it quantifies over *whatever set of worlds it may be* that the house rules pick out.

The technical implementation of this insight requires that we think of the context's contribution not as a set of worlds, but rather as a function which for each world it applies to picks out such a set. For example, it may be the function which, for any world w, yields the set $\{w'$: the house rules that are in force in w are obeyed in $w'\}$. If we apply this function to a world w_1, in which the house rules read "no noise after 10 pm", it will yield a set of worlds in which nobody makes noise after 10 pm. If we apply the same function to a world w_2, in which the house rules read "no noise after 11 pm", it will yield a set of worlds in which nobody makes noise after 11 pm.

Suppose, then, that the covert restrictor of a modal predicate denotes such a function, i.e., its value is of type $\langle s, st\rangle$.

(30) $[_{I'}\ [_I \text{ must } R_{\langle n,\langle s,st\rangle\rangle}\ ]\ [_{VP} \text{ you quiet}]]$

And the new lexical entries for *must* and *may* that will fit this new structure are these:

(31) a. $[\![\text{must}]\!]^{w,g} = [\![\text{have-to}]\!]^{w,g} = [\![\text{need-to}]\!]^{w,g} = \ldots =$
$\lambda R \in D_{\langle s,st\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \forall w' \in W\ [R(w)(w') = 1 \rightarrow q(w') = 1]$
in set talk: $R(w) \subseteq q$

b. $[\![\text{may}]\!]^{w,g} = [\![\text{can}]\!]^{w,g} = [\![\text{be-allowed-to}]\!]^{w,g} = \ldots =$
$\lambda R \in D_{\langle s,st\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \exists w' \in W\ [R(w)(w') = 1\ \&\ q(w') = 1]$
in set talk: $R(w) \cap q \neq \emptyset$

Let us see now how this solves the contingency problem.

(32) Let w be a world, and assume that the context supplies an assignment g such that $g(R_{\langle 17,\langle s,st\rangle\rangle}) = \lambda w.\ \lambda w'.$ the house rules in force in w are obeyed in w'

$[\![\text{must } R_{\langle 17,\langle s,st\rangle\rangle} \text{ you quiet}]\!]^{w,g} =$ (IFA)
$[\![\text{must } R_{\langle 17,\langle s,st\rangle\rangle}]\!]^{w,g}(\lambda w'\ [\![\text{you quiet}]\!]^{w'}) =$ (FA)
$[\![\text{must}]\!]^{w,g}\ ([\![R_{\langle 17,\langle s,st\rangle\rangle}]\!]^{w,g})\ (\lambda w'\ [\![\text{you quiet}]\!]^{w'}) =$ (lex. entries *you*, *quiet*)
$[\![\text{must}]\!]^{w,g}\ ([\![R_{\langle 17,\langle s,st\rangle\rangle}]\!]^{w,g})\ (\lambda w'.$ you are quiet in $w') =$ (lex. entry *must*)
$\forall w' \in W$: $[\![R_{\langle 17,\langle s,st\rangle\rangle}]\!]^{w,g}(w)(w') = 1 \rightarrow$ you are quiet in w' = (pronoun rule)
$\forall w' \in W$: $g(R_{\langle 17,\langle s,st\rangle\rangle})(w)(w') = 1 \rightarrow$ you are quiet in w' = (def. of g)
$\forall w' \in W$ [the house rules in force in w are obeyed in w'
$\rightarrow$ you are quiet in w']

As we see in the last line of (32), the truth-value of ?? depends on the evaluation world w.

EXERCISE 3.2: Describe two worlds w_1 and w_2 so that
$[\![\text{must } R_{\langle 17,\langle s,st\rangle\rangle} \text{ you quiet}]\!]^{w_1,g} = 1$ and $[\![\text{must } R_{\langle 17,\langle s,st\rangle\rangle} \text{ you quiet}]\!]^{w_2,g} = 0$. □

Exercise 3.3: In analogy to the deontic relation $g(R_{\langle 17,\langle s,st\rangle\rangle})$ defined in (32), define an appropriate relation that yields an epistemic reading for a sentence like *You may be quiet.* □

3.2.4 Iteration

Consider the following example:

(33) You might have to leave.

What does this mean? Under one natural interpretation, we learn that the speaker considers it possible that the addressee is under the obligation to leave. This seems to involve one modal embedded under a higher modal. It appears that this sentence should be true in a world w iff some world w' compatible with what the speaker knows in w is such that every world w'' in which the rules as they are in w' are followed is such that you leave in w''.

There is more to be said about which modals can embed under which other modals. See for some discussion the handout mentioned earlier: http://web.mit.edu/fintel/lsa220-class-2-handout.pdf.

Assume the following LF:

(34) $[_{I'}$ [might $R_{\langle 1,\langle s,st\rangle\rangle}$] $[_{VP}$ [have-to $R_{\langle 2,\langle s,st\rangle\rangle}$] $[_{IP}$ you leave]]]

Suppose w is the world for which we calculate the truth-value of the whole sentence, and the context maps R_1 to the function which maps w to the set of all those worlds compatible with what is known in w. *might* says that some of those worlds are worlds w' that make the tree below *might* true. Now assume further that the context maps R_2 to the function which assigns to any such world w' the set of all those worlds in which the rules as they are in w' are followed. *have to* says that all of those worlds are worlds w'' in which you leave.

From now on, we will omit the type-designation of variables whenever we feel confident that their type is easy to figure out from the context.

In other words, while it is not known to be the case that you have to leave, for all the speaker knows it might be the case.

Exercise 3.4: Describe values for the covert $\langle s, st\rangle$-variable that are intuitively suitable for the interpretation of the modals in the following sentences:

(35) As far as John's preferences are concerned, you *may* stay with us.

(36) According to the guidelines of the graduate school, every PhD candidate *must* take 9 credit hours outside his/her department.

(37) John *can* run a mile in 5 minutes.

(38) This *has* to be the White House.

(39) This elevator *can* carry up to 3000 pounds.

For some of the sentences, different interpretations are conceivable depending on the circumstances in which they are uttered. You may therefore have to sketch the utterance context you have in mind before describing the accessibility relation. □

EXERCISE 3.5: Collect two naturally occurring examples of modalized sentences (e.g., sentences that you overhear in conversation, or read in a newspaper or novel – not ones that are being used as examples in a linguistics or philosophy paper!), and give definitions of values for the covert $\langle s, st\rangle$-variable which account for the way in which you actually understood these sentences when you encountered them. (If the appropriate interpretation is not salient for the sentence out of context, include information about the relevant preceding text or non-linguistic background.) □

3.2.5 A technical variant of the analysis

In our account of the contingency of modalized sentences, we adopted lexical entries for the modals that gave them world-dependent extensions of type $\langle\langle s, st\rangle, \langle st, t\rangle\rangle$:

(40) (repeated from earlier):
For any $w \in W$: $[\![\text{must}]\!]^{w,g}$
$\lambda R \in D_{\langle s,st\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \forall w' \in W\ [R(w)(w') = 1 \rightarrow q(w') = 1]$
(in set talk: $\lambda R_{\langle s,st\rangle}.\ \lambda q_{\langle s,t\rangle}.\ R(w) \subseteq q$).

Unfortunately, this treatment somewhat obscures the parallel between the modals and the quantificational determiners, which have world-independent extensions of type $\langle et, \langle et, t\rangle\rangle$.

Let's explore an alternative solution to the contingency problem, which will allow us to stick with the world-independent type-$\langle st, \langle st, t\rangle\rangle$-extensions that we assumed for the modals at first:

(41) (repeated from even earlier):
$[\![\text{must}]\!]^{w,g} = \lambda p \in D_{\langle s,t\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ \forall w \in W\ [p(w) = 1 \rightarrow q(w) = 1]$
(in set talk: $\lambda p \in D_{\langle s,t\rangle}.\ \lambda q \in D_{\langle s,t\rangle}.\ p \subseteq q$).

We posit the following LF-representation:

(42) $[_{I'}\ [_{I}\ \text{must}\ [\ R_{\langle 4,\langle s,st\rangle\rangle}\ w*]]\ [_{VP}\ \text{you quiet}]]$

What is new here is that the covert restrictor is complex. The first part, $R_{\langle 4,\langle s,st\rangle\rangle}$, is (as before) a free variable of type $\langle s, st\rangle$, which gets assigned an accessibility relation by the context of utterance. The second part is a special terminal symbol which is interpreted as picking out the evaluation world:

Note that as soon as we're introducing an object language expression whose extension is a possible world, we will now have expressions of type s and should also introduce the domain of things of type s: $D_s = W$.

(43) For any $w \in W : [\![w*]\!]^{w,g} = w.$[6]

6 **dowty:1982:time-adverbs** introduced an analogous symbol to pick out the evaluation *time*. We have chosen the star-notation to allude to this precedent.

When $R_{\langle 4,\langle s,st\rangle\rangle}$ and $w*$ combine (by Functional Application), we obtain a constituent whose extension is of type $\langle s,t\rangle$ (a proposition or set of worlds). This is the same type as the extension of the free variable p in the previous proposal, hence suitable to combine with the old entry for *must* (by FA). However, while the extension of p was completely fixed by the variable assignment, and did not vary with the evaluation world, the new complex constituent's extension depends on both the assignment and the world:

(44) For any $w \in W$ and any assignment g:
$[\![R_{\langle 4,\langle s,st\rangle\rangle}\ w*]\!]^{w,g} = g(\langle 4, \langle s, st\rangle\rangle)(w)$.

As a consequence of this, the extensions of the higher nodes I and I′ will also vary with the evaluation world, and this is how we capture the fact that ?? is contingent.

Maybe this variant is more appealing. But for the rest of this chapter, we continue to assume the original analysis as presented earlier. Nevertheless, make sure you understand what we just proposed, since it may turn out to be useful.

3.3 *Kratzer's Conversational Backgrounds

Angelika Kratzer has some interesting ideas on how accessibility relations are supplied by the context. She argues that what is really floating around in a discourse is a CONVERSATIONAL BACKGROUND. Accessibility relations can be computed from conversational backgrounds (as we shall do here), or one can state the semantics of modals directly in terms of conversational backgrounds (as Kratzer does).

A conversational background is the sort of thing that is identified by phrases like *what the law provides, what we know*, etc. Take the phrase *what the law provides*. What the law provides is different from one possible world to another. And what the law provides in a particular world is a *set of propositions*. Likewise, what we know differs from world to world. And what we know in a particular world is a set of propositions. The intension of *what the law provides* is then that function which assigns to every possible world the set of propositions p such that the law provides in that world that p. Of course, that doesn't mean that p holds in that world itself: the law can be broken. And the intension of *what we know* will be that function which assigns to every possible world the set of propositions we know in that world. Quite generally, conversational backgrounds are functions of type $\langle s, \langle st, t\rangle\rangle$, functions from worlds to (characteristic functions of) sets of propositions.

Now, consider:

(45) (In view of what we know,) Brown must have murdered Smith.

The *in view of*-phrase may explicitly signal the intended conversational background. Or, if the phrase is omitted, we can just infer from other clues in the discourse that such an epistemic conversational background is intended. We will focus on the case of pure context-dependency.

How do we get from a conversational background to an accessibility relation? Take the conversational background at work in (45). It will be the following:

(46) $\lambda w.\ \lambda p.\ p$ is one of the propositions that we know in w.

This conversational background will assign to any world w the set of propositions p that in w are known by us. So we have a set of propositions. From that we can get the set of worlds in which all of the propositions in this set are true. These are the worlds that are compatible with everything we know. So, this is how we get an accessibility relation:

(47) For any conversational background f of type $\langle s, \langle st, t\rangle\rangle$, we define the corresponding accessibility relation R_f of type $\langle s, st\rangle$ as follows:
$R_f := \lambda w.\ \lambda w'.\ \forall p\ [f(w)(p) = 1 \ \rightarrow\ p(w') = 1]$.

In words, w' is f-accessible from w iff all propositions p that are assigned by f to w are true in w'.

Kratzer uses the term MODAL BASE for the conversational background that determines the set of accessible worlds. We can be sloppy and use this term for a number of interrelated concepts:

(i) the conversational background (type $\langle s, \langle st, t\rangle\rangle$),
(ii) the set of propositions assigned by the conversational background to a particular world (type $\langle st, t\rangle$),
(iii) the accessibility relation (type $\langle s, st\rangle$) determined by (i),
(iv) the set of worlds accessible from a particular world (type $\langle s, t\rangle$).

Kratzer calls a conversational background (modal base) REALISTIC iff it assigns to *any* world a set of propositions that are all true in that world. The modal base *what we know* is realistic, the modal bases *what we believe* and *what we want* are not.

What follows are some (increasingly technical exercises) on conversational backgrounds.

EXERCISE 3.6: Show that a conversational background f is realistic iff the corresponding accessibility relation R_f (defined as in (47)) is reflexive. □

EXERCISE 3.7: Let us call an accessibility relation TRIVIAL if it makes every world accessible from every world. R is TRIVIAL iff $\forall w \forall w' :\ w' \in R(w)$. What would the conversational background f have to be like for the accessibility relation R_f to be trivial in this sense? □

EXERCISE 3.8: The definition in (47) specifies, in effect, a function from $D_{\langle s,\langle st,t\rangle\rangle}$ to $D_{\langle s,st\rangle}$. It maps each function f of type $\langle s,\langle st,t\rangle\rangle$ to a unique function R_f of type $\langle s,st\rangle$. This mapping is not one-to-one, however. Different elements of $D_{\langle s,\langle st,t\rangle\rangle}$ may be mapped to the same value in $D_{\langle s,st\rangle}$.[7]

- Prove this claim. I.e., give an example of two functions f and f' in $D_{\langle s,\langle st,t\rangle\rangle}$ for which (47) determines $R_f = R_{f'}$.
- As you have just proved, if every function of type $\langle s,\langle st,t\rangle\rangle$ qualifies as a 'conversational background', then two different conversational backgrounds can collapse into the same accessibility relation. Conceivably, however, if we imposed further restrictions on conversational backgrounds (i.e., conditions by which only a proper subset of the functions in $D_{\langle s,\langle st,t\rangle\rangle}$ would qualify as conversational backgrounds), then the mapping between conversational backgrounds and accessibility relations might become one-to-one after all. In this light, consider the following potential restriction:

  (48) Every conversational background f must be "closed under entailment"; i.e., it must meet this condition:
  $\forall w.\forall p\ [\cap f(w) \subseteq p \ \rightarrow\ p \in f(w)]$.

  (In words: if the propositions in $f(w)$ taken together entail p, then p must itself be in $f(w)$.) Show that this restriction would ensure that the mapping defined in (47) will be one-to-one. □

3.4 Supplementary Readings

The most important background readings for this chapter are the following two papers by Kratzer:

---

7 In this exercise, we systematically substitute sets for their characteristic functions. I.e., we pretend that $D_{\langle s,t\rangle}$ is the power set of W (i.e., elements of $D_{\langle s,t\rangle}$ are sets of worlds), and $D_{\langle st,t\rangle}$ is the power set of $D_{\langle s,t\rangle}$ (i.e., elements of $D_{\langle st,t\rangle}$ are sets of sets of worlds). On these assumptions, the definition in (47) can take the following form:

(i) For any conversational background f of type $\langle s,\langle st,t\rangle\rangle$,
we define the corresponding accessibility relation R_f of type $\langle s,st\rangle$ as follows:
$R_f := \lambda w.\ \{w' : \ \forall p\ [p \in f(w) \ \rightarrow\ w' \in p]\}$.

The last line of this can be further abbreviated to:

(ii) $R_f := \lambda w.\ \cap f(w)$

This formulation exploits a set-theoretic notation which we have also used in condition (48) of the second part of the exercise. It is defined as follows:

(iii) If S is a set of sets, then $\cap S \ :=\ \{x : \ \forall Y\ [Y \in S \ \rightarrow\ x \in Y]\}$.

Angelika Kratzer. 1981. The notional category of modality. In Hans-Jürgen Eikmeyer & Hannes Rieser (eds.), *Words, worlds, and contexts: New approaches in word semantics* (Research in Text Theory 6), 38–74. Berlin: de Gruyter.

Angelika Kratzer. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter. https://udrive.oit.umass.edu/kratzer/kratzer-modality.pdf.

There are updated versions of Kratzer's classic papers in her volume "Modals and conditionals" (https://doi.org/10.1093/acprof:oso/9780199234684.001.0001).

A major new resource on modality is Paul Portner's book:

portner:2009:modality-book

You might also profit from other survey-ish type papers:

fintel:2005:modality

fintel-gillies:2007:ose2

swanson:2008:modality

hacquard:2009:hsk-modality

On the syntax of modals, there are only a few papers of uneven quality. Some of the more recent work is listed here. Follow up on older references from the bibliographies in these papers.

bhatt:1997:haveto

wurmbrand:1999:raising

cormack-smith:2002:modals

butler:2003:modality

The following paper explores some issues in the LF-syntax of epistemic modals:

fintel-iatridou:2003:ec

Valentine Hacquard's MIT dissertation is a rich source of cross-linguistic issues in modality, as is Fabrice Nauze's Amsterdam dissertation:

hacquard:2006:dissertation

nauze:2008:thesis

Some more recent work by Hacquard deals with deriving and correlating modal flavors with syntactic position of the modal auxiliaries:

hacquard-2010-event-relativity

hacquard-2013-modality

A recent handbook article by Hacquard on actuality entailments (involving the interaction of modality with aspect; we'll discuss aspect later in these notes):

hacquard-2016-actuality

The semantics of epistemic modals has become a hot topic recently. Here are some of the main references:

hacking:1967:possibility

teller:1972:epistemic

derose:1991:epistemic

egan-hawthorne-weatherson:2005:epistemic

egan:2007:epistemic

macfarlane:2006:might

Tamina Stephenson. 2007a. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525. http://dx.doi.org/10.1007/s10988-008-9023-4.

hawthorne:2007:danger

fintel-gillies:2008:cia-leaks

fintel-gillies:2008:mmr

A paper by Pranav Anand and Valentine Hacquard tackles what happens to epistemic modals under attitude predicates:

anand-hacquard-2013-epistemic-attitudes

Evidentiality is a topic closely related to epistemic modality. Some references:

willett:1988:evidentials

aikhenvald:2004:evidentiality

drubig:2001:epistemic

blain-dechaine:2007:evidentials

mccready-ogata:2007:evidentials

speas:2008:evidentials

fintel-gillies:2010:mss

Modals interact with disjunction and indefinites to generate so-called free choice-readings, which are a perennial puzzle. Here is just a very small set of initial references:

kamp:1973:freechoice

zimmermann:2000:fc-disjunction

schulz:2005:fcp-synthese

aloni:2007:freechoice

alonso-ovalle:2006:thesis

fox:2007:freechoice

rooij:2006:donkeys

PART II

DPS IN INTENSIONAL CONTEXTS

Bibliography

Asher, Nicholas. 1987. A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy* 10(2). 125–197. http://dx.doi.org/10.1007/BF00584317.

Barcan, Ruth C. 1946. A functional calculus of first order based on strict implication. *Journal of Symbolic Logic* 11(1). 1–16. http://dx.doi.org/10.2307/2269159.

Blumson, Ben. 2009. Pictures, perspective and possibility. *Philosophical Studies*. http://dx.doi.org/10.1007/s11098-009-9337-2.

Bonomi, Andrea & Sandro Zucchi. 2003. A pragmatic framework for truth in fiction. *Dialectica* 57(2). Preprint http://filosofia.dipafilo.unimi.it/~bonomi/Pragmatic.pdf, 103–120. http://dx.doi.org/10.1111/j.1746-8361.2003.tb00259.x.

Carnap, Rudolf. 1947. *Meaning and necessity: A study in semantics and modal logic*. Chicago: University of Chicago Press.

Chierchia, Gennaro & Sally McConnell-Ginet. 2000. *Meaning and grammar: An introduction to semantics (2nd edition)*. MIT Press.

Copeland, B. Jack. 2002. The genesis of possible worlds semantics. *Journal of Philosophical Logic* 31(2). 99–137. http://dx.doi.org/10.1023/A:1015273407895.

Cresswell, Max. 1973. *Logics and languages*. London: Methuen.

Dowty, David R., Robert Wall & Stanley Peters. 1981. *Introduction to Montague semantics*. Kluwer.

von Fintel, Kai. 1999. NPI licensing, Strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148. http://dx.doi.org/10.1093/jos/16.2.97.

Gajewski, Jon. 2005. *Neg-raising: Polarity and presupposition*. Massachusetts Institute of Technology dissertation. http://dx.doi.org/1721.1/33696.

Gajewski, Jon. 2007. Neg-raising and polarity. *Linguistics and Philosophy*. http://dx.doi.org/10.1007/s10988-007-9020-z.

Gamut, L. T. F. 1991. *Logic, language, and meaning*. Chicago University Press.

Garson, James. 2008. Modal logic. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*. http://plato.stanford.edu/entries/logic-modal/.

Geurts, Bart. 1998. Presuppositions and anaphors in attitude contexts. *Linguistics and Philosophy* 21(6). 545–601. http://dx.doi.org/10.1023/A:1005481821597.

Giannakidou, Anastasia. 1999. Affective dependencies. *Linguistics and Philosophy* 22(4). 367–421. http://dx.doi.org/10.1023/A:1005492130684.

Hanley, Richard. 2004. As good as it gets: Lewis on truth in fiction. *Australasian Journal of Philosophy* 82(1). 112–128. http://dx.doi.org/10.1080/713659790.

Heim, Irene. 1992. Presupposition projection and the semantics of attitude verbs. *Journal of Semantics* 9(3). 183–221. http://dx.doi.org/10.1093/jos/9.3.183.

Heim, Irene & Angelika Kratzer. 1998. *Semantics in generative grammar*. Oxford: Blackwell.

Hintikka, Jaako. 1969. Semantics for propositional attitudes. In J.W. Davis, D.J. Hockney & W.K. Wilson (eds.), *Philosophical logic*, 21–45. Dordrecht: Reidel.

Hockett, Charles F. 1960. The origin of speech. *Scientific American* 203. 89–96.

Hockett, Charles F. & Stuart A. Altmann. 1968. A note on design features. In Thomas A. Sebeok (ed.), *Animal communication: Techniques of study and results of research*, 61–72. Indiana University Press.

Hughes, G.E. & M.J. Cresswell. 1968. *An introduction to modal logic*. London: Methuen.

Hughes, G.E. & M.J. Cresswell. 1996. *A new introduction to modal logic*. London: Routledge.

Kadmon, Nirit & Fred Landman. 1993. *Any*. *Linguistics and Philosophy* 16(4). 353–422. http://dx.doi.org/10.1007/BF00985272.

Kratzer, Angelika. 1981. The notional category of modality. In Hans-Jürgen Eikmeyer & Hannes Rieser (eds.), *Words, worlds, and contexts: New approaches in word semantics* (Research in Text Theory 6), 38–74. Berlin: de Gruyter.

Kratzer, Angelika. 1991. Modality. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantics: An international handbook of contemporary research*, 639–650. Berlin: de Gruyter. https://udrive.oit.umass.edu/kratzer/kratzer-modality.pdf.

Kratzer, Angelika. 2006. Decomposing attitude verbs. Handout from a talk honoring Anita Mittwoch on her 80th birthday at the Hebrew University of Jerusalem July 4, 2006. http://semanticsarchive.net/Archive/DcwY2JkM/attitude-verbs2006.pdf.

Kripke, Saul. 1980. *Naming and necessity*. Oxford: Blackwell.

Lasersohn, Peter. 1996. Adnominal conditionals. *Semantics and Linguistic Theory* 6. 154–166. http://semanticsarchive.net/Archive/jRmZThkO/adnom.pdf.

Lewis, Clarence Irving & Cooper Harold Langford. 1932. *Symbolic logic*. New York: Century.

Lewis, David. 1970. General semantics. *Synthese* 22(1-2). 18–67. http://dx.doi.org/10.1007/BF00413598.

Lewis, David. 1978. Truth in fiction. *American Philosophical Quarterly* 15(1). Reprinted with postscripts in D. Lewis (1983), pp. 261–280, 37–46. http://www.jstor.org/stable/20009693.

Lewis, David. 1982. Logic for equivocators. *Noûs* 16(3). Reprinted in D. Lewis (1998: pp. 97–110), 431–441. http://dx.doi.org/10.2307/2216219.

Lewis, David. 1983. *Philosophical papers: Volume 1*. Oxford: Oxford University Press.

Lewis, David. 1986. *On the plurality of worlds*. Oxford: Blackwell.

Lewis, David. 1998. *Papers in philosophical logic*. Cambridge: Cambridge University Press.
MacFarlane, John. 2005. Logical constants. In Edward N. Zalta (ed.), *The stanford encyclopedia of philosophy*. http://plato.stanford.edu/archives/win2005/entries/logical-constants/.
Menzel, Christopher. 2016. Possible worlds. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Spring 2016. http://plato.stanford.edu/entries/possible-worlds/.
Montague, Richard. 1973. The proper treatment of quantification in ordinary English. In Jaako Hintikka, Julius Moravcsik & Patrick Suppes (eds.), *Approaches to natural language*, 221–242. Reprinted in Portner & Partee (2002), pp. 17–34. Dordrecht: Reidel. http://www.blackwellpublishing.com/content/BPL_Images/Content_store/Sample_chapter/9780631215417/Portner.pdf.
Moulton, Keir. 2008. Clausal complementation and the *Wager*-class. *North East Linguistics Society (NELS)* 38. http://sites.google.com/site/keirmoulton/Moultonnels2008wager.pdf.
Moulton, Keir. 2009. *Natural selection and the syntax of clausal complementation*. University of Massachusetts at Amherst dissertation. http://scholarworks.umass.edu/open_access_dissertations/99/.
Partee, Barbara H. 2005. Reflections of a formal semanticist as of Feb 2005. ms. (longer version of introductory essay in 2004 book). http://people.umass.edu/partee/docs/BHP_Essay_Feb05.pdf.
Partee, Barbara H. & Herman L.W. Hendriks. 1997. Montague grammar. In Johan van Benthem & Alice ter Meulen (eds.), *Handbook of logic and language*, 5–91. Elsevier.
Portner, Paul. 1997. The semantics of mood, complementation, and conversational force. *Natural Language Semantics* 5(2). 167–212. http://dx.doi.org/10.1023/A:1008280630142.
Portner, Paul & Barbara H. Partee (eds.). 2002. *Formal semantics: The essential readings*. Oxford: Blackwell.
Proudfoot, Diane. 2006. Possible worlds semantics and fiction. 35(1). 9–40. http://dx.doi.org/10.1007/s10992-005-9005-8.
Ross, Jeff. 1997. *The semantics of media* (Studies in Linguistics and Philosophy (SLAP) 64). Dordrecht: Kluwer.
Russell, Bertrand. 1940. *An inquiry into meaning and truth*. London: George Allen & Unwin.
Spencer, Mary. 1971. Why the "s" in "intension"? *Mind* 80(317). 114–115. http://dx.doi.org/10.1093/mind/LXXX.317.114.
Stalnaker, Robert. 1984. *Inquiry*. MIT Press.
Stalnaker, Robert. 1999. *Context and content*. Oxford: Oxford University Press.
von Stechow, Arnim. 1991. Intensionale semantik – eingeführt anhand der temporalität. http://www.sfs.uni-tuebingen.de/~astechow/Aufsaetze/Int.SemTemp91.pdf.

Steiner, George. 1998. *After Babel: Aspects of language and translation*. 3rd edn. Oxford University Press.

Stephenson, Tamina. 2007a. Judge dependence, epistemic modals, and predicates of personal taste. *Linguistics and Philosophy* 30(4). 487–525. http://dx.doi.org/10.1007/s10988-008-9023-4.

Stephenson, Tamina. 2007b. *Towards a theory of subjective meaning*. Massachusetts Institute of Technology dissertation. http://semanticsarchive.net/Archive/2QxMjk0O/Stephenson-2007-thesis.pdf.

Steup, Matthias. 2008. The analysis of knowledge. In Edward N. Zalta (ed.), *The Stanford encyclopedia of philosophy*, Fall 2008. http://plato.stanford.edu/archives/fall2008/entries/knowledge-analysis/.

Swanson, Eric. 2011. Propositional attitudes. In Klaus von Heusinger, Claudia Maienborn & Paul Portner (eds.), *Semantics: An international handbook of meaning*, vol. 2 (Handbücher zur Sprach- und Kommunikationswissenschaft 33.2), chap. 60, 1538–1561. Berlin/Boston: de Gruyter Mouton. http://dx.doi.org/10.1515/9783110255072.1538. http://tinyurl.com/swanson-hsk.

Varzi, Achille. 1997. Inconsistency without contradiction. *Notre Dame Journal of Formal Logic* 38(4). 621–638. http://dx.doi.org/10.1305/ndjfl/1039540773.

Williamson, Timothy. 2000. *Knowledge and its limits*. Oxford: Oxford University Press.

Zimmermann, Thomas Ede & Wolfgang Sternefeld. 2013. *Introduction to semantics: An essential guide to the composition of meaning*. de Gruyter Mouton.

Zucchi, Sandro. 2001. Tense in fiction. In Carlo Cecchetto, Gennaro Chierchia & Maria Teresa Guasti (eds.), *Semantic interfaces: Reference, anaphora and aspect*, 320–355. CSLI Publications. http://tinyurl.com/5ulwxwg.