24.910 Topics in Linguistic Theory: Propositional Attitudes
Spring 2009

1. Some Notes on the Reading

[Ch. 1-2 of von Fintel & Heim's *Intensional Semantics*]

❖ **"intensional" / "intension"**

➢ Spelled with an "s" – presumably on analogy with "extension" and to distinguish it from "intention"

❖ **"Assignment functions"**

These are used (among other things) to capture the interpretation of pronouns, and shouldn't be relevant to us.

➢ You can ignore reference to "assignment functions" and the parameter "g."

➢ Read $[\![\alpha]\!]^{w,g}$ as $[\![\alpha]\!]^{w}$
(In most of the cases in the reading, "g" is an idle wheel anyway.)

❖ **Lambda notation**

➢ Quick and dirty version: This is a way of writing a function:

λ _____ . _____
   domain variable   output of function

➢ For example: instead of writing $f(x) = x^2$, we would write
$f = \lambda x.x^2$
or $f = \lambda x \in R.x^2$ [specifying that the domain is real numbers]
[Note: this also means that $\lambda x.x^2$ is simply the name of a function]

➢ One benefit: making it easier to write functions that take other functions as arguments (very common in semantics).

➢ For example: instead of writing $f(g(x)) = g(x) + 1$, we would write
$f = \lambda g . \lambda x . g(x) + 1$
or $f = \lambda g \in R \times R . \lambda x \in R . g(x) + 1$

❖ **Specifying types**

➢ e = individuals ["<u>e</u>ntities"]
t = truth values
s = worlds ["<u>s</u>ituations" / "<u>s</u>tates of affairs"]

➢ x_e = x (indicating that x is a variable ranging over individuals)

➢ $\langle\alpha, \beta\rangle$ = function from things of type α to things of type β. For example:

➢ For example: <e,t> = function from individuals to truth values (1-place predicate)
<s,t> = function from worlds to truth values (proposition)

❖ **Set-talk and function talk**

➢ Many semantic values are functions from something to <u>truth values</u>.

➢ Functions of this kind have a natural mapping with sets:
The set characterizing f = {x: f(x)=1}

➢ Thus $\langle \alpha, t \rangle$ can also be thought of as the type of a set of objects of type α

This relationship is so familiar to semanticists that we often go back and forth between set-talk and function talk without making this explicit.

1.1.   Selected Exercises

❖ **Exercise 1.1 (p. 6)**

➢ **⟦A famous detective lives at 221B Baker Street⟧^{w7}**

[Note: The subject takes the VP as an argument instead of vice versa because the subject is a quantifier; this makes the exercise needlessly complex for our purposes, but may help give some practice with the system assumed in the reading]

= ⟦A famous detective⟧^{w7} (⟦lives at 221B Baker Street⟧^{w7})

➢ **⟦lives at 221B Baker Street⟧^{w7} = ⟦lives at⟧^{w7} (⟦221B Baker Street⟧^{w7})**

[Substituting in lexical entries from (13c) and (15c) in Ch. 1 of von Fintel & Heim]

= [λx . λy . y lives at x in w_7] (221B Baker St.)

[Substituting 221B Baker St. for x]

= [λy . y lives at 221B Baker Street in w_7]

➢ **⟦famous detective⟧^{w7}**

[Using a rule of "predicate modification" introduced in H&K. This rule basically says that multiple modifiers can be put together semantically using set intersection.]

= ⟦famous⟧^{w7} ∩ ⟦detective⟧^{w7}

= λy . y is famous and y is a detective in w_7
In sets: {y: y is famous and y is a detective in w_7}[1]

➢ **⟦A famous detective⟧^{w7} = ⟦a⟧^{w7} (⟦famous detective⟧^{w7})**

[lexical entry for *a/some* from (14d)]

= [λf . λg . $\exists x$: f(x) = 1 and g(x) = 1] (**[λy . y is famous and y is a detective in w_7]**)

---

[1] Of course, being a famous detective (on one salient reading, at least) isn't really the same thing as being famous and being a detective – if Sherlock Holmes happened to moonlight as a waiter, we wouldn't call him a famous waiter. So *famous* should probably be treated as a non-intersective modifier (more generally, a function from sets to sets). But this will not be our problem in this course.

[substituting **bolded** argument for f]

$= \lambda g \,.\, \exists x\colon$ **[$\lambda y \,.\, y$ is famous and y is a detective in w_7]**$(x) = 1$ and $g(x) = 1$

$= \lambda g \,.\, \exists x\colon x$ is famous and x is a detective in w_7 and $g(x) = 1$

Back to the main computation:

➢ **$[\![$A famous detective lives at 221B Baker Street$]\!]^{w_7}$**

$= [\![$A famous detective$]\!]^{w_7}$ ($[\![$lives at 221B Baker Street$]\!]^{w_7}$)

[from above]

$= [\lambda g \,.\, \exists x\colon x$ is famous and x is a detective in w_7 and $g(x) = 1]$
(**[$\lambda y \,.\, y$ lives at 221B Baker Street in w_7]**)

[substituting **bolded** argument for g]

$= \exists x\colon x$ is famous and x is a detective in w_7 and **[$\lambda y \,.\, y$ lives at 221B Baker Street in w_7]** $(x) = 1$

$= \exists x\colon x$ is famous and x is a detective in w_7 and x lives at 221B Baker Street in w_7

2. Introducing Intensional Semantics

We can set up a lot of the mechanics of an intensional semantics using a simplistic but useful example of a fictional world.

2.1.   Preliminaries

(1)   A famous detective lives at 221B Baker Street.

➢ Extension of (1) at a given world w:

$[\![$a famous detective lives at 221B Baker Street$]\!]^{w}$
$= 1$ iff a famous detective lives at 221B Baker St. in w

➢ Intension of (1):

$\lambda w' \,.\,$ a famous detective lives at 221B Baker St. in w'

[We won't need to deal with the internal composition of (1).]

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

➢ Assume *in the world of S.H.* is a sentential modifier

➢ We can give a semantics for this modifier using an intensional system.

For convenience, let's pretend that:

➢ $w_9 =$ the world depicted in Sir Arthur Conan Doyle's stories about Sherlock Holmes

➢ $w_6 =$ the actual world

2.2.   Version I: "The world of Sherlock Holmes"

❖ **Informal idea**

The modifier signals that the sentence is to be evaluated at world w_9 (the world of Sherlock Holmes) rather than at the actual world

❖ **More formally:**

(3) $[\![$In the world of Sherlock Holmes, $\phi]\!]^w = [\![\phi]\!]^{w9}$

[i.e., *in the world of Sherlock Holmes, ϕ* is true in any world w iff ϕ is true in w_9.]

More compositionally:

(4) $[\![$In the world of Sherlock Holmes$]\!]^w = \lambda p_{<s,t>} . p(w_9)$

Evaluating (2) at world w_6:

(5) $[\![$In the world of Sherlock Holmes, a famous detective lives at 221B Baker Street$]\!]^{w6}$

[the meaning of the modifier (at the actual world) applied to the proposition that there is a famous detective at 221B Baker St.]

$= [\![$in the world of Sherlock Holmes$]\!]^{w6}$
$([\lambda w . $ a famous detective lives at 221B Baker St. in w$])$

$= [\lambda p_{<s,t>} . p(w_9)] ($**$[\lambda w .$ a famous detective lives at 221B Baker St. in w]** $)$

[substituting the argument for p]

$= $**$[\lambda w .$ a famous detective lives at 221B Baker St. in w]** (w_9)

[substitute w_9 for the variable w]

$= $ 1 iff a famous detective lives at 221B Baker St. in w_9

2.3.   Version II: The world of Sherlock Holmes as presented by Sir Arthur Conan Doyle in world w

➢ A problem with the idea in 2.2:

It's a contingent fact that Sir Arthur Conan Doyle wrote the Sherlock Holmes stories the way he did. He might have (in some <u>other</u> world) set things up so that Sherlock Holmes lived on Abbey Rd (and no detective lived on Baker St.)

➢ Improvement: Use a function from worlds to worlds:

sher (w) = the world of Sherlock Holmes as set down in w
so in particular, sher(w_6) = w_9

Perhaps:

$w_6 \rightarrow w_9$
$w_7 \rightarrow w_{12}$
$w_8 \rightarrow w_{11}$
...

Implementing this:

(6) $[\![\text{In the world of Sherlock Holmes, } \phi]\!]^w = [\![\phi]\!]^{sher(w)}$

Again, more compositionally:

(7) $[\![\text{In the world of Sherlock Holmes, } \phi]\!]^w = \lambda p_{<s,t>} . \, p(sher(w))$

= $\lambda p_{<s,t>}$. the world w' as it is described in the Sherlock Holmes stories <u>as written in w</u> is such that p(w') = 1

2.4.   Version III: The set of worlds compatible with what is presented by Sir Arthur Conan in world w

➢ A problem with the idea in 2.3: Sir Arthur Conan Doyle didn't actually make every aspect of his fictional world explicit. For example, we don't know whether Holmes had an odd or even number of hairs on his head the day he met Watson (and in some sense Sir Arthur Conan Doyle doesn't know either!)

So:

➢ It makes more sense to talk about a <u>set of worlds</u> compatible with what is written in the Sherlock Holmes stories.

➢ Recall: it's still a contingent fact that they were written the way they are

➢ So the function **sher** should now yield a <u>set of worlds</u>:
sher(w) = {w': w' is compatible with the world as depicted in the Sherlock Holmes stories as written in w}

Finally:

(8) $[\![\text{In the world of Sherlock Holmes, } \phi]\!]^w = 1$ iff $\forall w' \in sher(w), [\![\phi]\!]^{w'} = 1$

Again, more compositionally:

(9) $[\![\text{In the world of Sherlock Holmes, } \phi]\!]^w = \lambda p_{<s,t>}$. for all $w' \in sher(w), p(w') = 1$

3. Semantics of Attitude Predicates

Recall: the function **sher**:

➢ sher(w) = {w': w' is compatible with the world depicted in the Sherlock Holmes stories, as written in w}

We can define a similar function for a person's beliefs (and other attitudes):

(10) $Bel_{x, w}$ = {w': w' is compatible with what x believes in w}

[Note: We haven't said much of anything about what it means to be "compatible with what x believes" – more on this later.]

Using this function:

(11) $[\![\text{x believes } \phi]\!]^w = 1$ iff $\forall w' \in \text{Bel}_{x,w} : [\![\phi]\!]^{w'} = 1$

Breaking it down:

(12) $[\![\text{believe}]\!]^w = \lambda p_{<s,t>} . \lambda x_e \, \forall w' \underline{\in \text{Bel}_{x,w}} : p(w') = 1$
       $= \lambda p_{<s,t>} . \lambda x_e \, \forall w' : \underline{\text{w' is compatible with what x believes in w}} : p(w') = 1$

Of course, we can do something parallel for other attitudes:

(13) $[\![\text{know}]\!]^w = \lambda p_{<s,t>} . \lambda x_e \, \forall w' : \underline{\text{w' is compatible with what x knows in w}} : p(w') = 1$

(14) $[\![\text{suspect}]\!]^w$
       $= \lambda p_{<s,t>} . \lambda x_e \, \forall w' : \underline{\text{w' is compatible with what x suspects in w}} : p(w') = 1$

(15) $[\![\text{imagine}]\!]^w$
       $= \lambda p_{<s,t>} . \lambda x_e \, \forall w' : \underline{\text{w' is compatible with what x imagines in w}} : p(w') = 1$

(16) $[\![\text{want}]\!]^w = \lambda p_{<s,t>} . \lambda x_e \, \forall w' : \underline{\text{w' is compatible with what x wants in w}} : p(w') = 1$

Obviously we'll have more to say than this…

3.1.   Accessibility Relations

Q. what does it mean to be "compatible with" a person's knowledge, beliefs, etc.?

We won't really answer this, but we can say a little bit more about knowledge, belief, etc. [and thus about the lexical semantics of *know*, *believe*, etc.]

One thing that helps: hold the subject and type of attitude constant, and consider mental states as relations between worlds:

Another notation:

(17) **w R$_x$$^{\text{Bel}}$ w'**
       = w' is compatible with x's beliefs in w

When the subject and attitude type are understood, we might write **wRw'**.

Some terminology:

➢ **R**'s of this type are called <u>accessibility relations</u>.

➢ **wRw'** can be read as **w' is accessible to w** / (sometimes) **w sees w'**

We can say something more about propositional attitudes by talking about the properties of these various accessibility relations

3.1.1. Some properties of relations

❖ **Reflexivity**

R is <u>reflexive</u> iff   for all x in the domain of R, **xRx**

❖ **Transitivity**

R is <u>transitive</u> iff   whenever **xRy** and **yRz**, it's also the case that **xRz**

❖ **Symmetry**

R is <u>symmetric</u> iff whenever **xRy**, it's also the case that **yRx**

3.1.2. Accessibility Relations for *know*

➢ **Reflexive**
(This reflects the intuition that you can only <u>know</u> things that are true)

➢ **Transitive?**

Maybe… This depends on whether we want to assume that if you know something, then you know that you know it

➢ **Symmetric?**

Probably not … This depends on whether we want to assume that if something happens to be true in the actual world, then you <u>know</u> that it's compatible with your knowledge

3.1.3. Accessibility Relations for *believe*

➢ **NOT reflexive**
(because you can believe things that are false)

➢ **Transitive**
If you believe something, then you believe that you believe it

➢ **NOT Symmetric**

Something can be the case in the actual world which you do not believe to be compatible with your beliefs

3.1.4. Accessibility Relations for *want*

➢ **NOT reflexive**
(because you can want things to be the case that are not the case)

➢ **NOT Transitive**
(because presumably you can want something without <u>wanting</u> to want it)

➢ **NOT Symmetric**

Something can be the case in the actual world which is not compatible with what you <u>want yourself to want</u>

[Obviously there's a lot more to say about these relations than these three properties, but this gives us a framework]

3.2. Selected Exercises

[Possibly work through in class depending on time]

❖ **Exercise 2.1 (p. 19)**

❖ **Exercise 2.2 (p. 19)**

❖ **Exercise 2.3 (p. 20)**