Seth Cable
Spring 2009

Towards an 'Intensional' Semantics, Part 1: The Motivation for Intensions and How to Formalize Them [1]

In This Handout:

- The inadequacies of a purely extensional semantics
- Towards a solution: intensions?
- Formalizing the notion of an 'intension'

1. The Inadequacies of a Purely Extensional Semantics

(1) **The (Surprising) Adequacy of a Purely Extensional Semantics**

As you saw through much of last semester (and were reminded in the review notes):
A *purely extensional* semantics is sufficient to obtain a truth-conditional semantics for a significant portion of natural language structures.

(2) **'Purely' Extensional Semantics?**

a. The interpretation function "**[[ ]]**" is (always) a function from natural language expressions to their extensions in the (actual) world.

b. The extension of a complex phrase is (always) derived by computing the *extensions* (and *only* the extensions) of its component parts.

(3) **Some Consequences of a Purely Extensional Semantics**

a. No 'Counterfactual' Language

(Informal Observation) In our purely extensional system, all predicates are of type <et> or <eet>. Thus, all sentences will be T or F depending *only* upon properties and relations that hold *in the actual world*.

- [[smokes]] = [λx. x smokes] (yields T *iff* x smokes in the actual world)
- [[Obama smokes]] = T *iff* Obama smokes (in the actual world)

b. Semantic Arguments are Always Extensions

(Assuming that all semantic composition is function application:) If a lexical item L semantically composes with a larger structure XP, then the extension of L is a function that that takes the *extension* of XP as its argument (*cf.* (2a,b))

[1] These notes are based upon material in Heim & Kratzer (1998; Chapter 12) and von Fintel (2007; Chapter 1).

(4) **An Empirical Problem: The Existence of Counterfactual Language**

(Informal Observation) Contrary to prediction (3a), there *are* sentences of natural language whose truth or falsity depends on properties or relations holding in *(purely) hypothetical worlds*.

a. Obama *might* smoke.
b. Michelle *believes* that Obama smokes.
c. *If* Joe smoked, then Obama *wouldn't* be alone.

Nothing in our purely extensional 'semantic toolkit' seems like it will provide a decent analysis of these kinds of structures.

But, the problem is even more acute than this…

(5) **Another Empirical Problem: Semantic Arguments that *Can't* be Extensions**

Consider the verb "believe"; from sentences like the following, it seems to have a meaning that combines with the meaning of a sentence (its complement clause).

a. Hannity believes [that Obama smokes].

In our (purely) extensional semantics, the 'meaning' of a sentence is a truth value.

b. **[[** Obama smokes **]]** = T

So, if we were to analyze the verb "believes" in our (purely) extensional semantics, we would have to view it as a function of type <t <e t>>
But, now consider the fact that the extension of (5a) [in the actual world] is T

c. **[[** Hannity believes [that Obama smokes] **]]** = T

Thus, the extension of "believes" must (qua function) contain <T, <Hannity, T>>.
But, now consider that the extension of the following sentence is *also* T:

d. **[[** Obama is a natural-born citizen **]]** = T

Thus, our extensional semantics for "believe" would entail/predict that:

e. **[[** Hannity believes [that Obama is a natural-born citizen] **]]** = T

Epic Fail: Our extensional semantics for "believe" makes the *obviously false* prediction that **if X believes one true/false sentence, then X believes *all* true/false sentences!**

But this obviously false prediction is a necessary consequence of two core assumptions of our purely extensional semantic system:
(i) The semantic value of a structure is (always) its extension
(ii) The extension of a sentence is its truth value.

(6) **Conclusion: Our (Purely) Extensional Semantics is Not Enough**

For words like "believe", their meaning does *not* combine with the *extension* of their sentential complement (*cf.* purely extensional 'logical operators' like "and").

a. In this structural context, our 'semantic valuation' function "**[[ ]]**" has to provide something *other* than the extension of the complement clause.

b. For sentences containing the verb "believes", their extension is not determined purely by computing the *extensions* of their component parts.

Thus the core assumptions of our (purely) extensional semantics in (2) are <u>wrong</u>!
(Extensions just aren't enough for a complete semantics of human language…)

> **Side-Note:**
> Some philosophers (*e.g.* Donald Davidson) have spent their career fighting against the conclusion above. Thus, for at least *some* folks, the conclusion in (6) is controversial…

---

2. Towards a Solution: Intensions?

(7) **A Recap: What Do We Need?**

a. From examining the intuitive content of words like "believe", "might", *etc.*, it appears that we must extend our theory of 'semantic/logical types' to include objects that would permit 'talk about purely hypothetical worlds/situations'.

b. From examining (in detail) how the compositional semantics for "believe" must operate, it appears that we must extend our semantic theory so that the interpretation function "**[[ ]]**" sometimes yields objects other than extensions.

So, how we get what we need? Well, let's take a break and recall the notion of an 'intension':

(8) **The Intension of a Structure X (Informally Put)**

The 'intension' of X is the *formula/description/concept*, which determines the extension of X (in the actual world).

a. "The president"
   (i) *Extension* = (referent) Barack Obama
   (ii) *Intension* = (concept) 'The person who holds the presidency'

b. "The president smokes"
   (i) *Extension* = (truth value) T
   (ii) *Intension* = (truth conditions/proposition) The president smokes.

Now, consider the following line of thought…

(9) **One Line of Thought…**

a. Sentences with the same *extension* (truth value) can nevertheless have two different *intensions* (truth conditions)

- "Obama smokes" is T *iff* **Obama smokes**
- "Obama is a natural-born citizen" is T *iff* **Obama is a natural-born citizen.**

b. Thus, if a word like "believe" took as argument the *intension* of its complement clause (rather than its extension), we could avoid the epically false prediction that **if X believes one true/false sentence, then X believes *all* true/false sentences!**

- Since the intension of "Obama smokes" is distinct from the intension of "Obama is a natural-born citizen", Hannity can stand in the 'believes'-relation to the former, but not the latter!

c. *Consequently*, if we (somehow) allow that the semantic valuation function "**[[ ]]**" yields the *intension* of the complement of "believes", we can avoid the empirical problem in (5)!!!

(10) **Some Independent, Conceptual Motivation**

Question: What kind of relation does the verb "believes" represent?

a. Not a Relation Between an Entity and a *Sentence* <e, <e, t>>!

The following seems true: "Qin Shi Huang thought terracotta was nice."
But, what kind of possible relation could Qin Shi Huang have to the *English sentence* "terracotta was nice". He predated the possibility of the sentence by at least a millennium!

b. Not a Relation Between an Entity and Truth Value <t <e, t>>!
(see reasoning above in (5))

c. A Relation Between an Entity and a Proposition / Truth Conditions?

Seems to match to our informal pre-theoretic notions…
Even though Qin Shi Huang never uttered or assented to the *English* sentence "terracotta was nice," he *did* utter/assent to an *Archaic Chinese* sentence that expressed the same 'proposition' (had the same truth-conditions).

So, even at a very pre-theoretic, informal level, it seems profitable to analyze "believe" as a relation between an individual and the *intension* of a sentence (*i.e.*, truth conditions or 'propositions').

(11) **The Plan**

Let's pursue (flesh-out and formalize) this idea that semantic values can sometimes be *intensions*.

a. The resulting theory will avoid the (acute) empirical problem in (5).

b. **Moreover**, we will see over subsequent weeks that it also resolves the more general empirical issue in (4).

The resulting 'intensional' system will provide the tools necessary to analyze those sentences that seem to describe purely 'hypothetical' situations and relations…

---

3. Formalizing the Notion of an Intension

So, following the plan in (11), we want to have a fully fleshed-out, formalized semantic system that manipulates *intensions*.

… so, let's get started…

(12) **PROBLEM**

a. Thus far, our concept of an 'intension' has been a rather informal one (*cf.* (8)).

b. If we want a formal system that 'manipulates' intensions, we need some kind of a formal model of what an 'intension' is.

c. Thus, we need to have a way of modeling the notion of an 'intension' in our formal machinery (*i.e.* using set-theoretic concepts and our lambda notation).

(13) **Towards A Formal Model of 'Intensions': The Basic Idea**

a. The Core Property of an 'Intension' (*cf.* (8))
For any structure X, the 'intension' of X determines the extension of X (in the actual world).

Thus, if you pair the 'intension' of X with the actual world, that yields the extension of X.

b. The Formal Insight
Thus, we can think of the 'intension of X' as a kind of *function!*
- It's a function which, if you give it a universe (particularly the actual universe), it gives you some object back - namely, an object identifiable as the extension of X (in that universe).

So, let's try to flesh out in more detail the basic idea in (13b)!

3.1 Step 1: The Ontology of 'Possible Worlds'

The 'basic idea' in (13b) is that intensions are functions from 'universes' to things (extensions).

Thus, the semantic system we are aiming to create is one where we have functions whose domain is a set of 'universes'…

Thus, we must somehow add a set of 'universes' to the overall inventory of 'things' that the meanings of natural language structures make reference to…

(14) **Some Novel Terminology**

Instead of the term 'universes', we'll adopt the (centuries-old, but initially confusing) term 'possible worlds'

- Thus, intensions will be functions from 'possible worlds' to things.
- The actual world is considered one of an infinite set of 'possible worlds'

(15) **A Picture of the 'Metaphysics' this Semantics Assumes**

a. <u>The Actual World</u>
The 'actual world' is the sum total of all facts - past, present and future - until the end of time. It encompasses such facts as:
- Obama is president in 2009
- Seth Cable lived at 67 Harvard Avenue in 2003
- Julius Caesar was stabbed exactly 23 times.
- 10,000 years ago, there were 239,653 California condors in existence.
- 5,673 years from now, on a planet orbiting a pulsar 328 light years away, a stone will fall on its side.

b. <u>The Contingency of the Actual World</u>
Many of these 'facts' making up the actual world *could have been otherwise*.
- Obama could have decided *not* to run for president.
- I could have stayed in the grad dorms until 2007.
- One of the Roman senators could have missed.
- 10,000 years ago, there could have been only 239,652 California condors.

c. <u>The HUGE Metaphysical Leap</u>
For every fact about the actual world that *could have been otherwise*, there actually exists an 'alternative reality' where that fact **is** otherwise.
(That's what "X could have been otherwise" *means*: there is an alternative reality where X does not hold.)
- Our actual universe is one of a set of (very real) alternative universes. The whole set of these universes could be called a 'multi-verse' or a 'bulk'.
- The 'multi-verse/bulk' will be represented by the set **W** (set of all possible worlds.)
- The actual world (but one member of W) can be represented as '$\boldsymbol{w_0}$'

3.2 Step 2: Formalizing the 'World-Dependency' of Extensions

Now that we've explicitly added to the "ontology" of our semantic theory the set W of *all possible worlds*, we can begin to explicitly formalize the (informal) notion that the 'extension' of a given structure depends upon the state of the world…

To begin, let us reflect on the following, key property of our semantic valuation function "**[[ ]]**"

(16) **The Value of "[[ ]]" Depends Upon the State of the Real World**

If the semantic valuation function "**[[ ]]**" takes a sentence as argument, it gives back T or F *depending upon the state of the (actual) world.*
Thus, this function can serve to derive truth-conditional statements like the following:

a. [[Obama smokes]] = T iff Obama smokes.

However, these kinds of truth conditional statements are always implicitly about *truth in the actual world*. Thus, we might offer the following as a more explicit restatement of the truth-conditions our purely extensional semantic derives:

b. [[Obama smokes]] = T (in the real world) iff Obama smokes (in the real world)

Finally, given our 'special notation' for the real world in (15c), we can rewrite this as:

c. [[Obama smokes]] = T in w_0 *iff* Obama smokes in w_0

Thus, we find that our purely extensional semantics derives generalizations of the form in (16c). However, when we consider the truth of sentences *across* possible worlds, it's apparent that (16c) is simply a limited consequence of a much broader generalization.

(17) **Truth of Sentences Across Possible Worlds**

a. First, note that there are possible worlds other than w_0 where Obama smokes. In these other possible worlds, some other facts are different (maybe water is purple) but the fact remains that Obama smokes.

b. Intuitively, in these *other* possible worlds, the sentence "Obama smokes" is T.

c. Next, note that there are possible worlds other than w_0 where Obama *doesn't* smoke. Intuitively, in *these* possible worlds, the sentence "Obama smokes" is F.

d. Thus, we arrive at this generalization (of which (16c) is a particular consequence):

e. **Let w be *any* possible world in W:**
[[Obama smokes]] = T in w iff Obama smokes in w.

New Sub-Goal:
Let us augment our semantic theory so that it derives not simply the (limited) set of statements in (16c), but the more general, trans-world statements in (17e).

(18) **Sub-Step 1: New Notation**

Let's first introduce the following, more compact notation for the "left hand" part of the targeted generalization in (17e).

a. **[[X]]**^{w} = the extension of X at world w.

Thus, we can rewrite our targeted generalization as the following:

b. Let w be *any* possible world in W:
[[Obama smokes]]^{w} = T iff Obama smokes in w.

Side-Note:
In the notation "**[[X]]**^{w}", the possible world w paired with "**[[ ]]**" is *the evaluation world*.

(19) **Sub-Step 2: New Lexical Entries**

Using the new notation in (18a), let us re-write our lexical entries for "smokes" and "Obama" accordingly.

a. [[smokes]]^{w} = λx. x smokes in w.
At any possible world w, the extension of "smokes" is:
the function from entities x to truth-values which yields T iff x smokes in w.

Side-Note:
Following the lexical entry in (19a), the actual <et> function that "smokes" has as its extension will depend upon the 'world' that the word "smokes" is being 'evaluated at'.
- In worlds where Obama smokes, the extension in (19a) will yield T when applied to Obama.
- In worlds where Obama *doesn't* smoke, (19a) will yield F when applied to Obama.
Thus, following (19a), the extension of "smokes" will be different at different possible worlds!

b. [[Obama]]^{w} = Obama
At any possible world w, the extension of "Obama" is Obama

Side-Note:
Following the lexical entry in (19b), the entity that "Obama" has as its extension is the *same* in all possible worlds.
- The term for such a word (same extension in all possible worlds) is a *rigid designator.*
The treatment of proper names as *rigid designators* could be taken as just a 'simplifying assumption'. ***However, there are some serious philosophical arguments in support of it.***

New Sub-Goal Achieved!
With lexical entries like those in (19a,b), our semantic theory can now derive statements of the more general form in (17e)!!

(20) **Derivation of Statement (17e)**

Let w be any possible world…

a. $[[\text{ Obama smokes }]]^w = T$ *iff* (by FA)
b. $[[\text{ smokes }]]^w([[\text{Obama}]]^w) = T$ *iff* (by (19b))
c. $[[\text{ smokes }]]^w(\text{Obama}) = T$ *iff* (by (19a))
d. $[\lambda x.\ x \text{ smokes in } w](\text{Obama}) = T$ *iff* (by LC)
e. Obama smokes in w.

Side-Note:
By a completely parallel computation, we can derive the statement in (16c):

$[[\text{ Obama smokes }]]^{w0} = T$ *iff* Obama smokes in w_0

Thus, our augmented system still derives all the truth-conditional statements of our original, purely extensional system.

We've simply extended the system so that it captures in an explicit way the notion that a given structure of English will have a different 'extension' at different possible worlds.

Thus, we've explicitly formalized the notion that the 'extension' of a phrase depends upon the state of the world that the phrase is evaluated at!

3.3 Step 3: Building an 'Intension' from These Ingredients

Let us now recall the 'formal insight' we wish our system to capture:

(21) **Targeted 'Formal Insight'**
We can think of the 'intension of X' as a kind of *function*, which takes a possible world w and returns *the extension of X in w.*

Well, in this context, consider the function in (22), written variously as (22a), (22b) and (22c).

(22) **The Intension of "Obama Smokes"**

a. $\lambda w.\ [[\text{ Obama smokes }]]^w$ =

b. $\lambda w.$ Obama smokes in w =

c. *The function from possible worlds to T values, which when given a possible world w as an argument, yields T iff* $[[\text{ Obama smokes }]]^w = T$ (*iff Obama smokes in w*).

(23) **The Importance of the Function in (22)**

Following the notation we developed in Section 3.2, the function in (22) takes a particular possible world *w* as argument, and returns the extension of "Obama smokes" at *w*.

- It takes *w* and returns T *iff* (the extension of) "Obama smokes" is T at *w*.
- It takes *w* and returns F *iff* (the extension of) "Obama smokes" is F at *w*.

Thus, following our 'targeted formal insight' in (21), the function in (22) could be regarded as the *intension* of "Obama smokes".

This basic result generalizes to all structures of natural language…

(24) **The Intension of "X"**

Recall the definition of the following notation from (18a)

a. **[[X]]**^{w} = the extension of X at world w.

Thus, the function in (24b) clearly can be characterized by the prose in (24c)

b. λw. [[X]]^{w} =

c. *The function whose domain is the set of possible worlds, and when given a possible world w as argument, yields* ***the extension of X at w*** *as its value.*

Thus, following our 'targeted formal insight', the function in (24b) is identifiable as 'the intension of X'

(25) **General Conclusion**
For any structure X, the function **'λw. [[X]]^{w}'** is the ***intension*** **of X**.

(26) **Some New 'Abbreviating' Notation**

[[X]]$_{¢}$ = λw. [[X]]^{w} = *the intension of X*

(27) **Some New Terminology**

a. Proposition: *function from worlds to truth values (the intension of a sentence)*

b. Property: *function from worlds to <et> functions (the intension of a VP, NP)*

c. Individual Concept: *function from worlds to entities (the intention of a DP)*

3.4 Step 4: Extensions to Our Type Theory

We have achieved the goals stated at the beginning of this section:

We have a way of modeling the notion of an 'intension' in our formal machinery, using set-theoretic concepts and our lambda-notation.

This formalization captures the 'insight' that we can think of an 'intension' as a kind of function, one that takes a possible world and returns an extension.

We are thus well on our way to achieving our over-arching goal – stated in (11) – to flesh out a formal system in which the 'semantic value' of an expression can sometimes be its *intension*!

This ultimate goal will come to fruition in "Part 2" of these notes.

Before we close, however, there is one last piece of business, regarding our 'type theory'.

(28) **'Intensional-izing' our Theory of Types**

Recall that our 'type theory' is a kind of notation for describing the full inventory of *things* that our semantic valuation function yields as the meanings of structures in our object language.

If we are pursing the notion that our semantic value function sometimes yields *intensions* as the meanings of structures, then we are (of course) adding to this 'inventory of things', and so should modify our 'type theory' accordingly.

All we need to add to our type theory from the first handout is the following additional 'recursive'/'inductive' statement:

a. If σ is a type, then $<s , \sigma>$ is a type: $D_{<s, \sigma>}$ = the functions from W to D_σ

Thus, for every semantic type (extension) σ of our earlier extensional semantics, we also have the 'intensional type' $<s , \sigma>$ of functions from possible worlds to things of type σ.

(The statement in (28a) also gives us a lot more besides that, but that's part of the broader empirical question of 'what semantic types does natural language really make use of')

<u>Side-Note:</u>

Note that (as of yet) we are not adding to our type theory a 'type *s*' of possible world.

This is because (as far as we've seen thus far) there are no expressions of natural language that have *specific possible worlds* as their values…