What have we been doing here?

In general, theoretical linguistics is looking to explore our **knowledge of language**. There are many things that we seem to **know** but not know we know. What is the nature of that knowledge? How can we characterize it?

Here, we have been specifically investigating the **properties of meaning**. (We’ve been focusing on English, but most holds true in any language, properly translated).

Considering **sentences**, we have an intuition that a sentence can be **true** or **false**. If we know what a sentence **means**, we know **under what conditions it would be true**.

We’ve taken some strides toward **modeling** our knowledge of truth conditions by trying to create an **algorithm** for computing whether a sentence is true or false, one that is faithful to our intuitions about the **connections between statements**, including knowing what other statements would follow from a statement (entailments).

The **computational system** we use to construct these truth value judgments must be **compositional**—the meaning of the whole must be derived from the meaning of the component parts and the way in which they are put together. We know this because **language is creative**, we can understand novel sentences (e.g., we know what the world would have to look like for **every elephant kicked two fish** to be true).

Our knowledge of the **truth conditions** of a sentence is not all there is. It interacts with our knowledge of the **lexical meaning** of the words (such that we know that if **John was killed**, then we know that **John died**, for example), our **knowledge of the world** (such that we know that **The rock fears heights** is odd because fear requires a conscious experiencer, and **rocks** are not conscious), our **knowledge of the context** (including preceding utterances, the “common ground” of interlocutors, the speaker, the things being pointed to), our **assumptions about interlocutors** (including the principles of cooperative discourse).

To formalize the concept of **truth conditions**, we consider the **ways things could be**. There is a truth of the matter about the way things are, but we don’t necessarily know all of the facts. We know some of the facts, but the way things are could be any of the ways things could be that are consistent with those facts. We refer to the different ways things could be as **possible worlds**. We are in the **actual world**, but those of us who are not omniscient don’t know which of the possible worlds that actually is.
The meaning of a statement includes the specification of its **truth conditions**—that is to say, **those possible worlds in which the statement is true**.

So, **given a possible world, we can tell whether the sentence is true or false**. We can think of a statement as represented by a **set of possible worlds** (a subset of all of the possible worlds, the worlds in which the statement is true). Alternatively, we can think of it as a **function** that, given a possible world, will return **true** or **false** (**true** if the statement is true in that world). Either way, same information.

For two statements A and B, if any possible world in which A is true, B is also true, then A **entails** B. That is, if A is true, B is necessarily also true. If *John was killed* then *John died*. Note that it doesn’t go the other way… if *John died*, it isn’t necessarily the case that *John was killed*. There are possible worlds in which *John died* where *John was killed* isn’t true.

Other dimensions are relevant as well. Time, for example. *John runs* might be true in a given world only from time $t_0$ to time $t_1$, but true in a different world from time $t_0$ to time $t_2$, $t_2 > t_1$.

To find out if a sentence is **true** then, we need to know both the **evaluation world** and **evaluation time**.
How do we model this knowledge?

We start with **individuals**. These are things we can refer to with **proper names** (like *John*) or with **descriptions** (*the dog*).

Individuals have **properties**. For example, some individuals are **boring**, some individuals **smoke**, some individuals are **pencils**, some individuals are **yellow**. Not every individual has every property — for a given individual *x*, *x* is either a **pencil** or it is not. We can think of **predicates** like *pencil* as being a set of **those individuals for which the predicate holds** (kind of like what we said above about possible worlds and statements).

**Smoke**, too, can be considered to be a set including all and only those individuals that smoke. Or it can be thought of as a **function** that, given an individual *x*, will return **true** or **false** depending on whether *x* is in the set of those who **smoke**. Like a slot machine: you put an individual in the slot and either get *T* or *F*.

Consider: John smokes. But it need not have been that way. There are possible worlds in which John doesn’t smoke.

**In a given world**, **smoke** represents a set of individuals who smoke in that world. This is the **extension** of the predicate **smoke**. However, when we take into account that different individuals smoke in different worlds, we have a predicate that is true or false of individuals in a given world. Given a world and given an individual, smoke is either true of that individual in that world, or false. This (taking into account the different possible worlds) is the **intension** of the predicate **smoke**.

Knowledge of meaning is not all we know about language. We also know how sentences are put together: **syntax**. We have not studied syntax seriously here, but it is relevant in part because of **compositionality** — the meaning of the whole is derived from the meanings of its components and the way in which the components are put together. *John kicked Bill* and *Bill kicked John* mean different things, despite having all of the same components. How things are put together makes a difference.

We have a lot of knowledge about how sentences are put together — this is a very interesting area to explore on its own, how languages differ — and *can* differ — from one another in human language. Fascinating stuff, take CAS LX 522 Syntax I next semester for more.
Here, we’re pursuing a hypothesis that the **syntax** and **semantics** are tightly connected. Sentences have a **structure**. This **structure** is used by the semantic interpretation to determine, for example, the difference between *John kicked Bill* and *Bill kicked John*.

The model of grammar:

![Grammar Diagram]

We created some small toy models of a syntax using **phrase structure rules** that say things like “Sentences are made of an NP and a TP”, “NPs are made of either 1) a proper name, 2) a Determiner and a common noun,” “TPs are made of a Tense and a VP”, “VPs are made of either 1) an intransitive verb, 2) a transitive verb and an NP, 3) a sentence-embedding verb and a CP”, “CPs are made of a complementizer and a sentence”. Then we have a **lexicon** that tells us what the intransitive verbs, the transitive verbs, the sentence-embedding verbs, the proper nouns, the common nouns, the determiners, etc., are. Together, these **generate** a sentence structure (“DS”) like the following.

(1)

![Sentence Structure Diagram]
Phrase structure rules (building the initial representation, DS “deep structure”)

<table>
<thead>
<tr>
<th>Rule</th>
<th>Left</th>
<th>Transformation</th>
<th>Right</th>
</tr>
</thead>
<tbody>
<tr>
<td>CP □ C S</td>
<td></td>
<td>NegP □ Neg VP</td>
<td>NP □ Det N_C</td>
</tr>
<tr>
<td>S □ NP TP</td>
<td></td>
<td>VP □ V_t NP</td>
<td>N_C □ Adj N_C</td>
</tr>
<tr>
<td>TP □ T VP</td>
<td></td>
<td>VP □ V_c CP</td>
<td>NP □ N_p</td>
</tr>
<tr>
<td>TP □ T NegP</td>
<td></td>
<td>VP □ V_i</td>
<td></td>
</tr>
</tbody>
</table>

Lexicon

Det □ the, a, every, no
C □ that, Ø
N_C □ book, fish, man, woman
V_i □ is boring, is hungry, smokes
V_t □ likes, hates
V_c □ believe, think, say
Neg □ not

Adj □ happy
N_p □ Pavarotti, Loren, Bond, he_n, she_n, it_n, him_n, her_n, himself_n, herself_n, itself_n
T □ might, can, must, should, PAST, PRES, FUT

After the phrase structure rules and lexicon build the DS, then certain syntactic transformations may occur. Again, not something we studied here, although an example of this is the transformation of John should buy what to What should John buy? in a question as below. (You will not need to know how to do this, just observe how it might work for the moment, to help see what DS and SS are)

(2) DS

After the phrase structure rules and lexicon build the DS, then certain syntactic transformations may occur. Again, not something we studied here, although an example of this is the transformation of John should buy what to What should John buy? in a question as below. (You will not need to know how to do this, just observe how it might work for the moment, to help see what DS and SS are)
After SS (Surface Structure), the semantics gets the structure and determines the meaning by combining the meanings of the components and the way they’re put together.

We define our semantics (for this fragment) in several parts. We start with the **universe of individuals**, which we call $U$. These are the things that proper nouns can refer to, that properties can hold of. These are individuals, or “entities”. They are things of the “entity” type—**type $<e>$**, we might say.

We also define a **valuation function** that we call $V$. **This is what tells us the meanings of the individual words.** Starting with the individuals, we say that $V(John)$, for example, is the individual (drawn from $U$) that we refer to when we use the proper name $John$. The valuation function is also responsible for telling us the **extension** of different **predicates**, like the intransitive verbs and adjectives. $V^{w,t}(smokes)$ is the set of individuals (drawn from $U$) that are smokers in possible world $w$ at time $t$, and this is the set we refer to when we use the verb *smokes*. Similarly for transitive verbs like *like*, $V^{w,t}(likes)$ is the set of **pairs** of individuals (each member of the pair drawn from $U$) such that the first member of the pair likes the second member of the pair, and to which we refer when we use the world *likes*.

Together $U$ and $V$ constitute the **model** of the world (it lists the individuals and their properties and relations), so we sometimes write $M$ to stand in for both (that is, $M = <U,V>$).

When we interpret the LF structure, we can start out with the lexical items and the values returned by $V$, but we also need rules that allow us to derive the meaning of their combinations as well. We have a small number of very general rules that allow us to do this. The simplest is **Pass-up**, which says that the “combination” of one thing has the same meaning as its single element. When we combine two things, we use **Functional**
Application. A prerequisite for Functional Application is that when we combine two
things, one must be a function that takes the other as its argument. For example the
intransitive verb *smokes* is something which, given an individual, is either true or false
(depending on whether the individual smokes). We can think of this type of function as
being \(<e,t>\), a function that takes an individual (type \(<e>\)) and returns either true or false
(type \(<t>\)). A function need not return something of type \(<t>\), however. For example, *likes*
is a transitive verb—it takes an individual and returns a function, and the function it returns
is a function that takes an individual and returns true or false. *Likes* is thus type \(<e,<e,t>>\).
If the LF structure calls for the combination of two predicates (two functions, neither able
to take the other as its argument), we use Predicate Modification, which combines the
two predicates to form a combined predicate, true of individuals for which both of the
component predicates are true (e.g., *yellow pencil*).

These are very simple and very general rules of combination. We hope that in the end, this
is all there will be to the semantic interpretation of LF. In our own fragment, we added
some further rules about the interpretation of quantifiers (since we did not define a
meaning for, e.g., *every*). We could however.

There is a class of special noun phrases, the pronouns, which refer to an individual like a
proper name does, except which individual you are referring to when you say *He smokes*
depends on who you are pointing at. Whereas all other proper names refer to the
individual \(V\) returns for that proper name (e.g., \(V(John)\) is the individual known as “John”),
pronouns have an index (written as a subscript) and the individual that the pronoun refers
to is the individual being “pointed to”. The assignment function \(g\) is used to keep track of
who is being pointed to. If John is being pointed to with index 1, then \(he_1\) will refer to the
individual known as “John”. When we have a pronoun with an index as a lexical item, it
refers to the individual \(g\) returns when given the index. E.g., \(he_1 = g(1)\).

There is a set of elements in the sentence that are scope bearing. These are elements that
cannot be interpreted where they appear in the sentence structure at SS. In order to make
semantic sense of the structure (and more importantly to match our intuitions about the
meaning), the SS needs to be adjusted in order to make it interpretable. This adjustment
takes the form of transformations to the SS that put it in a form that the semantics can
interpret it. After the transformations are finished, we have the LF (“Logical Form”)
structure, and our semantics interprets the LF, translating the LF into a metalanguage
statement of the truth conditions we called \(lf\) (also “logical form”, but this time in the
metalanguage).

The scope bearing elements are quantificational NPs (like *every boy, no fish*) as well as
negation and tense.
Thinking about what each of these add to the meaning, we see that they each say something about a whole sentence. **Negation** of a true sentence is **false**, and negation of a false sentence is **true**. **Past** is true of a sentence if the sentence is true at a time in the past. **Quantificational NPs** are true of a sentence if, for example, for every $x$ we can pick that meets the restrictive condition (e.g., for *every dog*, we consider every individual $x$ we can pick that the property *dog* holds of), the sentence is true if we use $x$ in place of the NP in the sentence (e.g., *Every dog swims* is true if for every individual $x$ we can pick, if $x$ is a dog, $x$ swims).

The rules of **QR** (“Quantifier raising”), **Neg-raising**, and **Tense-raising** each take the scope bearing element they apply to and attach them higher in the sentence (to S) so that the sister (in the LF tree) of the new position of the scope bearing element is an S. Then, the meanings we intuitively assign to quantifiers, negation, or tense can be easily stated formally.

Scope bearing elements (particularly quantifiers) interact with each other as well. **QR** does not specify anything about which quantifier must move first; as a result we can derive, for the sentence *A student bought every book*, either ‘for each book $y$, there is a student $x$ such that $x$ bought $y$’ (a different student per book) or ‘there is a student $x$ such that for each book $y$, $x$ bought $y$’ (same student bought all books). This matches our intuition, and gives us a way to understand why our intuition is the way it is.

There is a special class of “Tense” items called **modals**, including *must*, and *can*. If we claim that *John is tall*, then we are claiming something about the actual world—we are claiming that the actual world can be found among those possible worlds in which John is tall. If we use a modal, however, we are claiming something about more worlds than just the actual world—**modals are used to make claims about a set of possible worlds**.

Suppose we say *John must leave*. This means that in all of the possible worlds under consideration, *John leaves* is true. It isn’t actually saying anything about the actual world *per se*. Similarly, *John can leave* means that in at least one of the possible worlds under consideration, *John leaves* is true. The possible worlds under consideration, or **modal base**, come in a couple of types. There’s the **root modal base**, which are basically all of the possible worlds consistent with the facts (known by anyone) about the world. The root modality generally conveys a “future” feel because even holding the facts about the world constant, events may unfold in different ways in the future. *John can leave* with the root modality says that John is capable of leaving, that there is some future point in some possible world for which the facts now are the same in which John leaves. There’s the **deontic modal base**, which are the possible worlds that meet some **standard of**
propriety. *John can leave* with the **deontic** modality says that in some possible world in which the standard of propriety is met, John leaves—or, that it would not be improper for John to leave. The **deontic** modality has a “permission” effect; John is permitted to leave. And, there’s the **epistemic modal base**, which are the possible worlds that are consistent with what the speaker knows about the world (probably incomplete). It gives the effect of ‘as far as I know’. So, *John can leave* with the **epistemic** modality says that in some possible world that is consistent with what I know, John leaves. There may be some fact (e.g., that John is in prison) that I don’t know and that would change things, but ‘As far as I know, John is capable of leaving.’

**Dynamics of discourse**

We can move beyond the question of what a sentence means (in the sense of the conditions under which it is true) and ask about how a sentence is *used*.

Generally, a sentence is used in a **discourse** and is designed to have a certain effect.

If a proposition is **asserted**, the effect is intended to be that the proposition is added to the set of things we know about the world, to the set of facts.

We can think of discourse between two participants as being kind of a “game” in which the goal is to determine “the way things are” (or, put another way, to identify the actual world from in among the possible worlds).

Two participants enter a discourse with a certain amount of shared knowledge about the world. For example, that the sky is cloudy, that it’s Tuesday, it’s April, that Pavarotti is a singer. These form a **common ground** of facts that the interlocutors share. This already narrows down considerably the possible worlds that are still candidates for being the actual world.

If one person **asserts** a proposition, s/he is most likely suggesting that this proposition be added to the common ground, that it be accepted by the hearer as a fact to further narrow down the candidates for the actual world.

One prominent phenomenon of statements is that they can also carry **presuppositions**. A **presupposition** is a proposition that is “taken for granted” and generally must be accepted for the asserted proposition to be evaluated as either true or false. *John has stopped smoking*, for example. Part of the lexical meaning of the word *stop* is that it presupposes that prior to stopping, the condition held. If John stopped smoking, he must have previously been smoking. If John has never smoked, it is hard to evaluate whether *John
has stopped smoking is true or false—it seems to be neither. Accepting the proposition John once smoked seems to be a precondition for evaluating John has stopped smoking. To accept (or even to deny) a proposition with a presupposition is to implicitly accept the presupposition. Similarly, the dog carries the presupposition that there is a unique dog among the individuals we are currently concerned with.

Often presuppositions are part of the common ground. If they are not, but are innocuous, the hearer will accommodate the presupposition and add it to the common ground before opting to add or reject the asserted proposition as well. Accommodation allows for increased efficiency in the game to determine how things are.

Another means to increase efficiency is making the assumption that the conversation is being conducted in a cooperative way. Paul Grice formulated a number of “maxims” that spell out more specifically what it means to be cooperative, but one of the main ideas behind them all is that people choose to say what they do say (rather than something else that they could have said) for a reason, and we are entitled to make inferences based on what they chose not to say. Specifically, a speaker’s contribution is:

- relevant (RELATION),
- (only) as informative as required (QUANTITY),
- something for which the speaker has adequate evidence and does not believe to be false (QUALITY),
- unambiguous, succinctly, orderly, and not obscure (MANNER).

If someone says I have two children, you are entitled to conclude that That person has n children is false for any number n higher than two. That’s because if they had three children, it would have been a stronger statement to say I have three children—having 3 kids implies having 2, but having 2 kids does not imply having 3. To be cooperative, you say as much as you have evidence for. Since the speaker didn’t opt to say 3, we conclude that there was a reason—since the speaker should have evidence and assuming it would be relevant, it must be false that the speaker has 3 children.

The things you are “entitled to conclude” on the basis of assumptions about cooperative discourse are implicatures—sort of like implications, but weaker. It is generally possible to defeat or reinforce an implicature (“I have two children; in fact, I have three”, “I have two children, but I don’t have three”).

…to be continued…