24.903 Week #1 – 2022-01-31 + 2022-02-02

Kai von Fintel

The genius of language

Language provides the means to construct a specific signal (noise or gesture) for any meaning we want to convey. As Wilhelm von Humboldt said, language allows "infinite use through finite means".

The most amazing sentence ever:

(1) A former meerkat expert at London Zoo has been ordered to pay compensation to a monkey handler she attacked with a wine glass in a love spat over a llama-keeper. (Associated Press, Oct. 14, 2015)

But also: we can communicate with minimal means. In the right context, a one word utterance (*Dude!*) can speak volumes (see https://www.dailymotion.com/video/x3ghyd).

To understand how we convey meanings, we need to do two things:

- study the compositional engine of language
- study the ways in which context can help create meaning

Starting out

(2) Aïda tabib
Aïda doctor

'Aïda is a doctor'

The *object language* here is Maltese (I found the example in Stassen 2003). The meta data (morphemic gloss, approximate translation) are given in the *meta language*, here: English, soon to be supplemented with technical concepts and notations.

Another common terminology is that the object language is *mentioned* (a bit like being "quoted") in our discussions and our linguistic writing, and that the meta language is being *used*.

We will often look at English as the object language, because it is the shared language of this classroom and thus immediately accessible as a data source. But the best work in linguistics, and thus also in semantics, takes a wide perspective, looking at multiple languages and seeking previously understudied sources of data.

The meaning of the whole sentence

To know what the sentence means you need to know under what conditions it is true (and under what conditions it is false).

In fact, the MOST CERTAIN PRINCIPLE of semantics is:

(3) If we have two sentences A and B, and A is true and B is false, then A and B do not mean the same. (Cresswell 1982: p. 69)

We think of the meaning of a sentence as something that sorts possible worlds into two kinds: the worlds of which the sentence is true and the worlds of which the sentence is false.

We model this by saying that the meaning of a sentence is a function that maps any world into one of two values. One of the values is the one that we map a world to if the sentence is true. The other value is the one for the case where the sentence is false.

Any two things will do, as long as we designate one as the one that gets assigned to worlds where the sentence is true. The term for this is "truth-values". And really anything will do. We could choose Venus for one and Mars for the other. And then we could choose one of them (Venus?) as the designated value. Another proposal (Hodges 2001: p. 5):

> even the most scrupulous sceptic could follow the literature if he defined the truth-value of all true sentences to be his left big toe and that of false sentences to be his right

One very useful choice is to use the numbers 0 and 1, a decision that is essentially due to the mathematician and logician George Boole (1815–1864). This is the convention we will adopt.

So, the meaning of the sentence in (2) is the function from possible worlds to truth-values that maps any world w to 1 if Aïda is a doctor in w and to 0 if Aïda is not a doctor in w.

We can call this function a "sorting" function since it divides the worlds into two sorts: the ones where Aïda is a doctor and the ones where she is not.

Another term that is used is that this function is the "characteristic function" of the set of worlds where Aïda is a doctor. It maps exactly those worlds to 1 and all others to 0.

Technically, there is a one-to-one relation between sets and their characteristic functions. In our system, functions are all over the place, but we will often enjoy the freedom of thinking in terms of the sets those functions characterize as well.

By the way, we now know how to state the meaning of the most amazing sentence in (1):

(4) The meaning of *A former meerkat expert at London Zoo has been ordered to pay compensation to a monkey handler she attacked with a wine glass in a love spat over a llama-keeper* is the function that maps any world w to 1 if w is a world in which a former meerkat expert at London Zoo has been ordered to pay compensation to a monkey handler she attacked with a wine glass in a love spat over a llama-keeper, otherwise w is mapped to 0.

Again: marvel at the fact that we have a specific signal that conveys precisely that rather bizarre information!

What sentences are good for

Thinking of sentences as sorting worlds into "true" worlds and "false" worlds gives us a handle on what we do when we utter sentences. Here's a first stab:

(5) ASSERTION OF A SENTENCE
When an utterer U asserts a sentence S, they are thereby proposing that the addressee A update their beliefs so as to exclude any world w such that $[\![S]\!]^w = 0$.

There's more to be done

Now, having made a determination about the meaning of the whole sentence is only part of the job. We need to explain how this meaning arises *compositionally* from the meaning of the parts of the sentence and the structure of the sentence.

(6) THE PRINCIPLE OF COMPOSITIONALITY
The meaning of any linguistic expression is determined by the meaning of the parts of that expression and its structure.

Without this, we couldn't produce a specific signal for a given arbitrary piece of information we want to convey.

The meaning of *Aïda*

We will make the assumption that the meaning of a name like *Aïda* is simply the individual that bears that name. This is also called the "reference" of the name, the "denotation" of the name, or the "extension" of the name.

(7) The meaning of the name *Aïda* is the individual named Aïda.

Note that we can't easily put the individual into this handout (they wouldn't fit to the page even if we used superglue). In its stead, we **use** the metalanguage phrase *the individual named "Aïda"*. We should keep in mind that the meaning of the name is in fact the individual not the phrase.

The meaning of *tabib*

Not a bad idea: the meaning of *tabib* is the set of individuals that are doctors, or more simply said: the set of doctors.

Again, the idea that we will work with seems a bit more complicated (but is really equivalent): the meaning of *tabib* is the characteristic function of the set of doctors. This means that when applied to any individual, it gives the designated "true" result (1 in our setup) for doctors and another result (0 in our case) for non-doctors. So, the meaning of *tabib* is a sorting function, again. While the whole sentence sorts worlds, the predicate *tabib* sorts individuals. A whole lotta sorting going on.

How do worlds enter the composition?

But actually, what we have now can't quite be right: there's no mention of the world in what we have said about the meanings of the two parts of the sentence and somewhere there needs to be something "world-dependent" that results in a claim about worlds!

We will assume that the world-dependence is not really in the name. *Aïda* picks out the same individual in all worlds (at least those where Aïda exists). The individual may have different properties in the different worlds (including what their name is), but it's the same individual. (This is something that can be debated and one can set up a different system, but we won't go into that in this class.)

But the predicate *tabib* is world-dependent. We may be in a world where Aïda is a doctor (she's an element of the set of doctors in our world), but she might have become a racecar driver, so there's some worlds where she's not a doctor.

So, the story is a bit more complicated. The meaning of *tabib* is a function that takes two arguments: a world and an individual, and it yields 1 if the individual is a doctor in the world and 0 if not.

This means that we need a way to deal with functions that take two arguments. To be honest, this is getting too hard to do just in prose. We need a formally explicit and perspicuous notation.

A start on notation

We will use the following notation for the meaning (or: semantic value) of any expression α: $[\![\alpha]\!]$. There's an interesting article on the history of this notation: Rabern 2016.

As we've said, many meanings in our system will be functions. We will make a distinction between how to "feed" arguments to these functions, depending on whether the argument is the meaning of a fellow constituent in the expression being evaluated, or whether the argument comes "from the outside". By the latter, we mean that we are using the meaning to apply to, say, the world we're trying to describe. Worlds, in our system, aren't meanings of constituents in linguistic expressions, so they will be treated differently in the notation.

So, we write:

(8) For any world w, $[\![\text{Aïda tabib}]\!]^w = 1$ if Aïda is a doctor in w, 0 otherwise.

As you can see, the meaning of *Aïda tabib* is written as $[\![\text{Aïda tabib}]\!]$ and we write its argument world w as a superscript on the right-hand $]\!]$.

We said that $[\![\text{tabib}]\!]$ is a function that takes both a world and an individual and then gives us 1 if that individual is a doctor in the world and 0 otherwise. In our sentence, the individual that is fed to the function as its argument is the meaning of *Aïda*, so its the meaning of a fellow linguistic constituent. We write this as follows:

(9) For any world w, $[\![\text{tabib}]\!]^w$ (Aïda) $= 1$ if Aïda is a doctor in w, 0 otherwise.

Many semanticists will look at (9) and call Aïda an "argument" and w a "parameter (of evaluation)". They really both are arguments of the function, but since they are conceptually distinct and also notationally distinct, it makes sense to use two different terms. (You can look at Wikipedia for some discussion of the terms "parameter" and "argument" in various programming languages, but the use here in semantics, and in logic, is slightly different from the one in computer science.)

One more decision for now: even though we have decided to say that *Aïda* does not have a world-dependent meaning, we will treat $[\![\text{Aïda}]\!]$ as a function that takes a world parameter and then yields the individual. This will make for smoother calculations (and also makes space for possible future decisions to treat names as world-dependent after all).

(10) For any world w, $[\![\text{Aïda}]\!]^w =$ Aïda.

We see that w doesn't occur on the right hand side of the "=" sign, which means that the meaning of *Aïda* applies to the world w vacuously and always gives the same result. So, that's what it means to have a world-independent meaning in our system.

One more thing: it would get tedious to keep writing things like "= 1 if such and such, and = 0 otherwise/if not such and such". We will use an abbreviation invented by the mathematician Halmos (see https://mathworld.wolfram.com/Iff.html): *iff* stands for "if and only if". So, we can write:

(11) For any world w, $[\![\text{tabib}]\!]^w$ (Aïda) = 1 iff Aïda is a doctor in w.

Now, we have that if Aïda is not in a doctor in w, the truth-value is not 1 (because it is *only* 1 if Aïda is a doctor in w). And since the only other truth-value is 0, we know that the truth-value is 0 if Aïda is not a doctor in w.

Putting things together

We can now see how the meanings of the parts combine to give the meaning of the whole. When we ask what sorting function on worlds the sentence has as its meaning, we say: the function that maps a world w onto 1 when the meaning of *tabib* (fed w as a parameter) applied to the meaning of *Aïda* (again, fed w as a parameter, but vacuously) yields 1. That in turn will be the case if Aïda is a doctor in w.

(12) For any world w, $[\![\text{Aïda tabib}]\!]^w = 1$ iff $[\![\text{tabib}]\!]^w([\![\text{Aïda}]\!]^w) = 1$

The idea is that the meaning of the whole results from combining the meanings of the parts by *function-argument application*. Note that the world parameter is fed to both part-meanings before they can combine by application.

The general formulation is:

(13) FUNCTION APPLICATION
If a constituent α has two daughters β and γ, and for any world w, if $[\![\beta]\!]^w$ is a function whose domain contains $[\![\gamma]\!]^w$, then $[\![\alpha]\!]^w = [\![\beta]\!]^w([\![\gamma]\!]^w)$.

FUNCTION APPLICATION is to semantics what MERGE is to syntax.

We will spend next week on solidifying this system, before we expand beyond the simplest possible sentences.

References

Cresswell, M. J. 1982. The autonomy of semantics. In Stanley Peters & Esa Saarinen (eds.), *Processes, beliefs, and questions*, 69–86. Dordrecht: Springer. DOI: 10.1007/978-94-015-7668-0_2.

Hodges, Wilfrid. 2001. Elementary predicate logic. In D. M. Gabbay & F. Guenthner (eds.), *Handbook of philosophical logic*, 2nd edn., 1–129. Dordrecht: Springer. DOI: 10.1007/978-94-015-9833-0_1.

Rabern, Brian. 2016. The history of the use of ⟦.⟧-notation in natural language semantics. *Semantics and Pragmatics* 9(12). DOI: 10.3765/sp.9.12.
Stassen, Leon. 2003. *Intransitive predication*. Reprint. new as paperback (Oxford Studies in Typology and Linguistic Theory). Oxford: Clarendon Press.