The Devil and Noah Webster

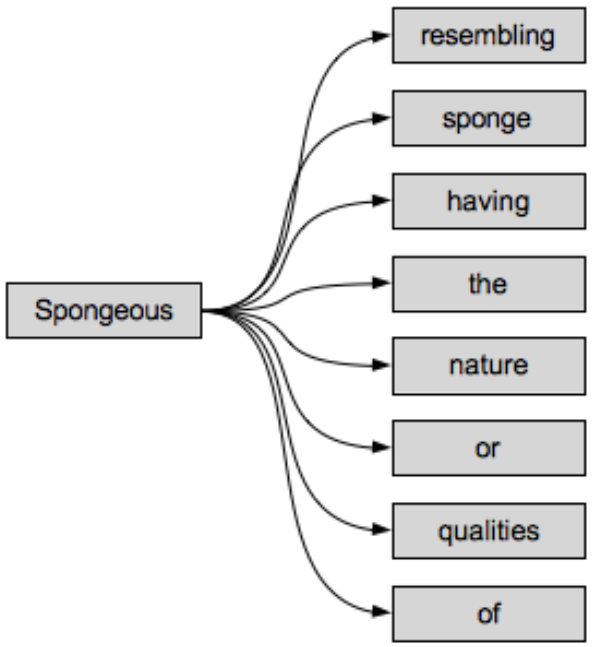
(first published 5/4/2014 on Riddles With the Sphinx)

Dictionaries as Networks

Zoë and I are in the process of building what amounts to a conceptual dictionary for Sphinx. As I gear up to discuss this, I want to report on a little mini-study I did awhile ago, on Webster's Unabridged Dictionary. WUD was published in 1864, and has around 114,000 entries. The idea of dictionaries arose in the East; in the West, they are less than 400 years old, and Webster's project was among the first major attempts.

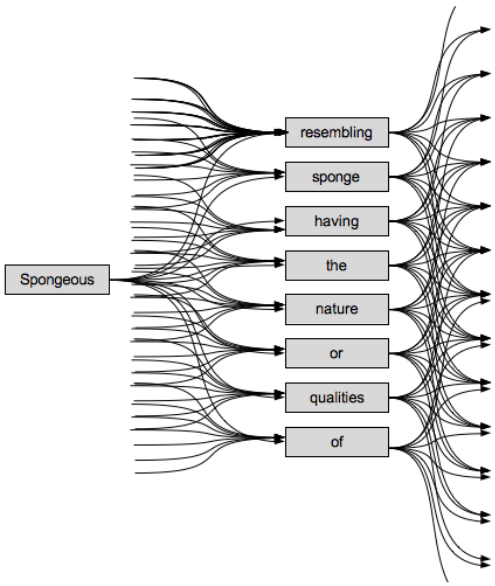
It establishes a general pattern that it can't quite live up to: that all words should be formally defined, without recursion. (By this, of course, we understand that Noah Webster was aiming to avoid *immediate* recursion, since if all words are defined, they are necessarily defined in terms of other words, which is ultimately recursive—much more on that thought in a moment.) It is not quite accurate to say that every word that appears in WUD is defined in its own entry, but this is nearly the case, and there is no particular reason why it should not be the case, given a bit more work.

We can visualize each definition in WUD as a node in a directed network, pointing downstream to as many other nodes as the (unique) words used in the definition. This gives us something like this:



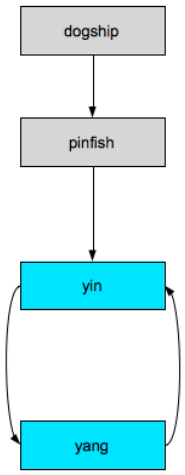
The word “spongeous”, as it happens, does not appear in the definitions of any other words. It shares this all-important feature with 40% of the words in WUD. Many of these are esoteric or spurious constructions (hydrargyrate, rigorism, or my own favorite, *dogship* (which is the proper form of address for dogs, of course)) while others are words that have some degree of day-to-day utility (receivership, certifier, indicatory). We can observe, however, that almost all of these words are constructions built off of some more common word: respectively sponge, mercury, rigor, dog, receiver, certify, and indicate. This 40%, then, forms a category that we might call *terminal concepts*—we can visualize them as the outermost edge of our directed network: nothing is upstream from there.

If we remove all of the terminal concepts from the dictionary, the remaining *WUD-1* will still "work" in the sense that all the words used in the dictionary could themselves be defined. We don't need the word "spongeous" to define "sponge", or to refer to "sponge" in other word definitions.



What if we go through this culling process again? There are new terminal concepts in *WUD-1*, which we can eliminate to produce *WUD-2*. In fact, at this stage we notice that there is a large group of words (about 28% of *WUD-0*) which are only ever used in the definitions of a small handful of other words. The typical candidates here are taxonomic clades and related terminology like acanthopterygious or pinfish. If you are going to discuss porgies, you are eventually going to need the word "pinfish", but it is quite possible to go through one's life without ever discussing porgies.

Then we repeat the process a few more times, and we eventually reach an irreducible core network, *WUD-n*. *WUD-n* contains the set of all words in WUD that cannot be defined without referencing *WUD-n*, an odd sort of recursion. In graph theory terms, it is "strongly connected". In the diagram below, this core network is "yin" and "yang", both of which are ultimately required to define *any word in the network, including themselves*.



Strictly from the point of view of graph theory, we can imagine a dictionary with a core network of only one node (undefined, or self-defining), or a dictionary with multiple core networks that are not connected to one another. But this is manifestly not the case with any dictionary based on real-world language. In fact, *WUD-n* is a very large and densely interlinked network, comprising about 32% of *WUD-0,* or roughly 36,000 words—which is at the high end of the vocabulary for a native English speaker.

The Nature of the Core

The WUD core network includes a number of fairly esoteric terms that happened to get dragged even though they may not seem like "core concepts" in our language. A common mechanism for this is that ostensive definitions—which are increasingly common near the core—tend to include examples, which necessarily jump across categories. For instance, a friend asked me to show how "democracy" could ultimately be required for defining "zinc". In my Apple desktop dictionary, this takes ten steps:

zinc → corrosion → material → goats →
domesticated → milk → rich → China →
capital → government → democracy

This sequence utilizes two shortcuts that exploit ostensive rather than formal definitions. The definition for *material* contains the example phrase "*goats can eat more or less any plant material"* and the definition for *rich* contains the phrase "*China's rich and diverse mammalian fauna."*

We might argue that this is cheating, since we can certainly imagine strictly formal definitions for "material" and "rich". But this is a misleading sentiment. In the first place, I have no doubt that we could connect "zinc" and "democracy" by a series of downstream links using only formal definitions; it would just take a little longer. In the second place, the reliance on ostensive (and recursive!) definitions seems to grow as we go further into the core. Just try defining "of" without using an example or the word "of". There is a real sort of semantic entropy at work here, which cannot be shrugged off.

The more important observation is that ostensive definitions (and similar constructions) tend to *enlarge* the core, by placing otherwise peripheral concepts downstream of core concepts. It is, for instance, probably possible to create a decent English dictionary in which "goat" is not in the core, though neither Webster nor Apple managed to do so.

The core has a very strict horizon: every word in the core is downstream of every word in the dictionary, including itself. But it is not at all homogenous over smaller scales. The core network words in WUD are used in an average of 200 other definitions apiece (compared to the overall average of 64 for *WUD-0*). But some of them are vastly more central than others. Understandably, the most frequently used words in the core are articles, conjunctions, pronouns, and the like: *The, a, of, to, or, and, in*. It is revealing to look at the most common verbs and nouns, though. Making some assumptions about the meaning of certain polysemes, we have the following list for verbs:

to be	**65k**
to see	**27k**
to have	**16k**

Note that *to be* and *to have* are both auxiliary verbs as well as fundamental concepts, so they are being double-counted. *To see* occurs almost entirely in the imperative form of "See X", comparable to "cf.", which occurs 14 thousand times. In other words, the subject of the verb is the reader, making it a fairly unusual type of definition (and again, one that seems to be more typical of the core).

Then we have:

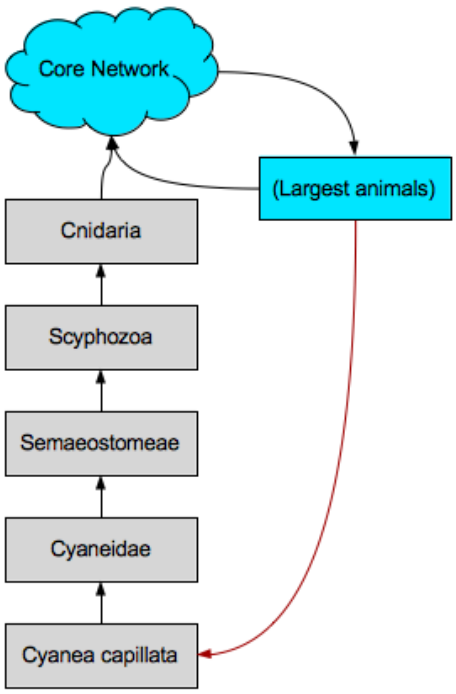
to call	**9.5k**	**to resemble**	**1.6k**
to pertain	**6k**	**to do**	**1.4k**
to act	**5k**	**to consist**	**1.3k**
to make	**4k**	**to obtain**	**1.3k**
to say	**2.3k**	**to work**	**1.2k**
to apply	**1.9k**	**to ally**	**1.2k**
to find	**1.8k**	**to cut**	**1.1k**
to cause	**1.8k**	**to form**	**1.1k**
to give	**1.7k**	**to produce**	**1.1k**
to take	**1.6k**	**to oppose**	**1k**

Several of these verbs--pertain, resemble, consist, oppose--are noteworthy because they do not appear that frequently in everyday English. “Pertain”, in particular, is the fifth most common verb to appear in the definitions, and yet it is virtual unused outside of that context. Arguably, “see also” and “pertain” are meta-concepts, used only when defining other concepts.

I can't provide empirical evidence for this (yet), but I have the distinct sense that the deep core network for human dictionaries contains many words that have to do with the human body, and with very basic trans-culturally obvious things like the moon, the ocean, etc. (Creating a comparable deep core network for Sphinx is a major desiderata of mine at the moment.)

From the perspective of this analysis, one of the things we see people doing when they learn is drawing new concepts into the core. For instance, if you tell me that *Cyanea capillata* is a species in the Cyaneidae, this is sort of a detached fact; it is the very minimal degree of understanding. If you then tell me that Cyaneidae are in the Semaeostomeae, which are in the Scyphozoa, I'm not really any better off. If you walk away at that point, I might easily forget the whole thing: that kind of understanding isn't worth whatever is limited in my memory-space.

But, if you tell me that Scyphozoans are in Cnidaria, then I can say “Oh! Jellyfish! You're talking about some kind of jellyfish, I know what those are.” This is a second level of understanding: I can reach the core from this concept by going upstream. But I still have no real way to manipulate the idea of *C. capillata* by itself, and once again, I am apt to forget about it.



At this point I probably ask you for more information about *C. capillata,* and I am told that they are the largest known jellyfish. This puts them downstream of a core concept (“largest animals”). But it also takes me to a third and crucial level of understanding: I can now reach *C. capillata* and the other clade concepts from the core, by strictly downstream links. This means, in fact, that *C. capillata* is now part of the core, and can be reached by a series of downstream links from any concept that I understand at least at level two.

This is very far from the entirety of what we mean by “understanding”, but I think it puts us on the right track. And I think it explains why learners tend to insist on hearing examples before we feel like we are beginning to understand something. We need to be able to approach concepts from all possible directions.

A possible metric emerges: from some random point in the core (or some particular non-random point, like “*i_myself”), how many basically different downstream paths are there to X? That number has got to be a very good approximation of how well one understands X.

And this means, by-the-by, that Webster's Unabridged Dictionary doesn't “really understand” two-thirds of the English language.