

In the Beginning was ...

The Word ??

The Sentence ??

Maybe even something bigger??

Democritus vs Aristotle

- Democritus: the founder of atomic theory
 - Matter was not infinitely divisible
 - Smallest parts: atoms
 - Too small to be visible
 - All bodies, including ours, are composed of a huge number of atoms
- Aristotle: first great anti-atomist
 - What counted were whole, integral/integrated complexes
 - Such complexes, including our own bodies, had to be understood in terms of the various kinds of causal relations into which they entered with other complex entities
 - One such kind of relation was that of the purpose or natural end/goal toward which a kind of entity tended or for which it was intended

Words vs. Larger, More Complex Structures

- A language isn't just a collection of words – its vocabulary
- Is it a *collection* of sentences?
- Is it to be presented via a grammar that yields all (and only) its sentences?
- Never mind a language: how about an actual text?
 - Just a bunch of words?
 - A sequence of words?
 - A sequence of sentences?
- Don't texts have a purpose (or purposes?)
- Really: Didn't the author of the text have a purpose(s) in mind when she wrote it?

What is it to Understand a Text?

- Surely it is, at the very least to understand its constituent sentences.
- SO, what is it to understand a sentence?
- Focus on purely declarative texts – texts meant to communicate (something like) factual information
 - Determine its truth/falsity ? Nah, way too demanding and often just plain inappropriate
 - Determine its truth-conditions: what the world would have to be like if it were true.
 - Calculate some of its entailments: what would follow if it (and it alone???) were true.
 - Which of the (infinitely?) many entailments?
 - Take appropriate action?
 - Really? Given what purposes/preferences/desires, etc. And again, given just *its* assumed truth?
 - Figure out what would be an appropriate action for an agent with ... standard beliefs? And some specified goal?
 - Be able to translate it accurately into another language?
 - This can't be right, as it rules out us monolinguals
 - And what is an accurate translation anyway?
- Note: some of these aren't so silly if we take as the proper unit of understanding, not an individual sentence but indeed the whole text.

Extent of Unit of Analysis

Three Levels of Semantics

1. Lexical Semantics

The meanings (= senses) of words

2. Sentential/compositional Semantics

How meanings of sentences are formed from meanings of words plus (syntactic) structure

3. Discourse-level semantics

How meanings of texts (multi-sentence texts) are formed from the meanings of the constituent sentences plus ... other facts about the context

Let's Start with Words

- Even if we don't ultimately adopt some version of "linguistic atomism"
- It certainly seems that to understand a text we at least have to understand (some of) the words
 - Given lexical ambiguity: that is to understand how the words in the text are being used
- And to do this is to understand what contribution to the meanings of the sentences in the text those words are making.

Senses of Words

- Unit of word meaning is a (word) sense
 - Usually represented explicitly in a “dictionary definition”
- Most (non-rare) words have multiple senses/meanings
 - A **bank** is a financial intermediary that creates credit by lending money to a borrower, thereby creating a corresponding deposit on the bank’s balance sheet.
 - In geography a **bank** generally refers to the land alongside a body of water

Some Terminology

- A **lexeme** is a pair of a word (**lemma** or **citation form**) and a sense or meaning
 - *bank* is the lemma/citation form for *banks*
 - *<bank, financial institution>*
- Specific surface forms (*bank, banks*) are **word-forms**
- But what about that second element in the lexeme?
- What form should representations of word senses (word-meanings) take??
 - In the example, it takes the form of ... more words!
 - A noun phrase “financial institution”
 - Which happens to be a phrase for a kind of thing that banks are

Homonymy

- Homonyms are lexemes that share a word-form
 - Orthographic, phonological or both
 - Homographs (*bank* and *bank*, *bass* and *bass*)
 - Homophones (*write* and *right*, *piece* and *peace*)
- But have unrelated meanings
 - <*bank*, financial institution> and <*bank*, riverside>
 - <*bass*, fish> and <*bass*, musical instrument>

Homonymy as Trouble!

- Speech Recognition (Speech-to-Text)
- Text-to-Speech
- Machine Translation
- Information Retrieval (written or spoken queries)

Lexical Relations: Related Senses

- (a) The *bank* was constructed in 1875 out of local limestone.
- (b) I withdrew the money from the *bank*.
- Is the word being used with the same sense in (a) and (b) ?
- Or are there two related senses:
 - Sense 1: “The building belonging to a financial institution”
 - Sense 2: “A financial institution”
- (a) Heavy snow caused the roof of the *school* to collapse.
- (b) The *school* hired more teachers this year than ever before.

Polysemy: Systematically (?) Related Senses

- Polysemy: multiple *related* senses
- E.g., organizations and buildings housing (or otherwise associated with) them:
 - *school, university, hospital, church,...*
- Authors (creators) and their works:
 - *Shakespeare* died in 1616. I read *Shakespeare* every year.

Distinguishing Multiple Senses

- How do we know when a word has more than one sense?
 - *Which flights serve breakfast?*
 - *Does America West serve Philadelphia?*
- The “zeugma” test:
 - *?Does United serve (both) breakfast and San Jose?*
- Since this sounds weird, we say that these are two different senses of *serve*.

Lexical Relations

- Synonymy
- Antonymy
- Hyponymy and Hypernymy
- Meronymy and Holonymy

For all this and more, check out WordNet

<http://wordnet.princeton.edu/>

WordNet

<http://wordnet.princeton.edu/>

Via web interface:

<http://wordnetweb.princeton.edu/perl/webwn>

- Hierarchically organized lexical database
- Machine-readable thesaurus (+ dictionary-like features)
- Versions for other languages under development
- For Italian: Istituto di Linguistica Computazionale; C.N.R., Pisa
- Or visit: <http://www.illc.uva.nl/EuroWordNet/>

Synonymy: As a Relation between Words

- Intuitively: two words are synonymous when they have the same meaning.
- So, this is a kind of identity relation and hence should be pretty simple ????
- Examples, as always in English
 - filbert/hazelnut
 - automobile/car
 - big/large (???)
 - water/H₂O (???)

WordNet Synsets

- A synset is a set of "synonyms" (words) representing a (single) sense
 - Each such sense — every member of the synset — shares the same *gloss* (the material in parens)
- “branch” (**n**)
 - *branch, subdivision, arm* (a division of some larger or more complex organization) "a branch of Congress"; "botany is a branch of biology"; "the Germanic branch of Indo-European languages"
 - *branch, leg, ramification* (a part of a forked or branching shape) "he broke off one of the branches"
 - *branch* (a stream or river connected to a larger one)

Synonymy: As a Relation between Senses, not Words

- *big* and *large*
 - *Daniele was like a **big** brother to David.*
 - *Daniele was like a **large** brother to David.*
- *big* has a sense that means: being older or more grown up
- *large* lacks this sense
- To be synonyms must words share all their senses/meanings in common?
- Surely not; for if so, there are probably *no* synonyms in any natural language.

Synonymy (cont.)

- More formally: $w1$ and $w2$ are synonymous – really they share a sense in common -- when they can be substituted, one for the other, *in sentences in which they are used with that sense*, without changing ... *the (propositional) meaning of the sentences (?)*
 - “car” is one-syllable word
 - “automobile” is a one-syllable word
 - Use vs mention
- But how about: Giovanni believes that water boils at 100C. Giovanni doesn't believe that H₂O boils is 100C.
 - Giovanni might not know/believe that water is H₂O?
- And what is our model of propositional meaning?
 - Truth-conditions?
 - And how are these represented?

Synonymy as an Equivalence Relation

- $w1$ is synonymous with $w1$ (*Reflexivity*)
- If $w1$ is synonymous with $w2$, then $w2$ is synonymous with $w1$ (*Symmetry*)
- If $w1$ is synonymous with $w2$ and $w2$ with $w3$, then $w1$ is synonymous with $w3$. (*Transitivity*)

Are There any Real Synonyms??

(and Does it Matter?)

- Why do we care about synonymy?
- In MT, we seem to need to find for a word in the Source Language a word in the Target Language that has the same meaning
- Or, to be more precise: for w_s , *used in the text (with sense s)*, we need a word w_t , *with that same sense s*
- Really??? Is that what we need?
- Isn't it enough to find a word or words in the target language that are really similar in meaning to w_s ??
- *Also: how would we determine synonymy computationally???*
 - *We can't really use the intersubstitutability *salve veritate* criterion*

Now for A Historical Digression

- Remember my little comically simplified history of approaches to AI??
- AI as a unique mode of “Synthetic Cognitive Psychology”
 - late '60's – very early '80's
- AI as Applied Logic (typically first-order logic)
 - McCarthy ('56--); others early '80's – mid '90's
- AI as Applied Probability Theory & Statistics
 - late -80's -- today

Models of Lexical Semantics: Semantic Networks

- Models of "semantic memory"
- And of the organization of the mental lexicon
 - Collins, Quillian, Loftus, et al. 1970'S
- Model of retrieval/priming/association: spreading activation
- First, take nodes as atomic: each node is a "concept", a meaning or sense of a word
 - Typically a common noun
- Links are heterogeneous, but most important are "is-a" links, directed from sub-concept (hyponym) to super-concept (hypernym)

Frames

- Now look inside the nodes
- Informal representation of crucial features of the concept (of the word-sense)
- attribute-value or role-filler pairs, often with "(stereo-)typical" or "default" values
 - Italians:
 - Type/kind: human
 - Hair-color: brunette
 - Human Father:
 - Children: humans (>0)

Logic-based Critique of KR Languages

- What are the rules for manipulating these frames or for reasoning using semantic networks?
- E.g., what are the rules of inference that take us from one ... graph (?) to another, preserving ... truth (?)
 - Note: early work took them as models of memory and association
 - Not as a formalism for representing information about the world, but for representing how people conceptualize? -- and reason??
- To answer these questions seemed to require giving these representations a clear, unambiguous and definite interpretation – a formal semantics
- Obvious solution: "translate" these representations into a first-order language – which has a formal semantics
- Then we get, for free, a sound and complete deductive system
- But ...Various things seemed to get lost in translation.....

Language of FOL

- Relations of varying degree (-arity)
 - 1-place: no systematic distinction between common nouns, adjectives, intransitive verbs
 - All get treated as Rx
- No distinction between accidental universal truths, law-like universal truths, definitional truths
 - All get treated as truth of $(x)(Fx \rightarrow Gx)$
 - That is: in a fixed interpretation, the set of F's is a subset of the set of G's

Mis-Match

- First-order languages seem very ill-matched to capture intuitions about word-sense / lexical semantics
 - Though darned good for lots of other things!
- Match is better with languages for modal/intensional logics
- Very little interest in *these* in NLP
- Very little interest in synonymy!

Word Similarity

- Real synonymy between/among words may be hard to come by, but there surely is word similarity – and that’s probably all we need
- Surely a word is similar to itself
- Surely if $w1$ is similar to $w2$, then $w2$ is similar to $w1$
- But similarity can “diminish” or fade away, so transitivity doesn’t hold true in general
 - $w1$ is similar to $w2$ and $w2$ is similar to $w3$; but $w1$ just may not really be similar to $w3$.

What Is It for One Word to be Similar to Another?

- Good question!
- First, translate into talk of word senses
- So what is it for two have words to have similar senses?
 - If ambiguous, they might also have dis-similar senses
- Words of wisdom:
 - Wittgenstein: Don't ask for the meaning/sense of a word; ask for its *use*
 - Firth: You shall know a word by the company it keeps
 - Israel: Words of a feather cluster together
- SO, two word-senses are similar to they extent that the words with those senses are used similarly; that is, to the extent to which the contexts in which they are (with those senses!) are similar.
- But aren't contexts just made up of other words??
 - Remember, we are talking about purely textual contexts here

Antonyms

- Antonyms: words that — in at least one of their senses — have opposing or contrary meanings
 - loud/soft (pertaining to volume)
 - hot/cold (pertaining to temperature)
 - up/down (pertaining to position or direction on ...the vertical axis)
 - living/dead (binary opposition)
 - Beware vampires and zombies!
 - rise/fall
 - How about e.g., red/green or red/blue ??

Hyponyms and Hypernyms

- Hyponymy: **X** is a hyponym of **Y** if it *denotes* (??) a sub-class/sub-set of Y or if it *means/connotes* (??) a concept that is included in the Y-concept
 - asymmetric, transitive relation between senses
 - E.g. dog/mammal ; mango/fruit
- Hypernymy: the inverse relation
 - mammal/dog ; fruit/mango

Hyponyms: Instances vs. Subclasses

- Is *Earth* a hyponym of *planet*?
- It is certainly an instance
- And its singleton {Earth} is a sub-class of the class (set) of planets {Mercury, Venus, Earth, ...}
- In my lingo: No! Instances are not sub-classes

Meronyms & Holonyms

- Meronymy: X is a meronym of Y if it denotes a part of Y
 - asymmetric, transitive relation between senses
 - wheel/car
 - nose/face
- Holonymy: its inverse
 - wheel/car
 - face/nose

Meronym subtypes

- **Part** vs. **Substance** vs. **Member** meronyms
 - Part: nose/face ; porch/house
 - Substance: rubber/tire ; water/ocean
 - Member: professor/faculty ; player/team