Computational Semantics and Pragmatics

Autumn 2012

Raquel Fernández
Institute for Logic, Language & Computation
University of Amsterdam

Where we are

Up to know we have been concerned with compositional semantics

- Montague-style formal semantics (nothing new about this)
- how we can automate the process of semantic composition
  - in particular, using functional programming

Out next topic, starting today, is lexical semantics

- what do words mean?
- how can we represent the meaning of words?
- what do psychologists and cognitive scientists tell us about word meaning and concepts?
- how can we disambiguate between words that have different senses?

The Lexicon

Broadly speaking, the lexicon is the vocabulary of a natural language (a speaker's vocabulary).
$\hookrightarrow$ what does it consist of, how is it organised?

- In compositional theories, the basic expressions are assumed to be word senses not words:
  - *'bank'*$_1$: the slope of land adjoining a body of water
  - *'bank'*$_2$: a business establishment in which money is kept
- In its simplest form, the lexicon is thus an inventory of word senses mapped to word forms or lemmas

$\Rightarrow$ enumerative view of the lexicon

Problems of the Enumerative View

Some problems we could mention:

- what words mean is not made explicit
- relations between senses are not captured
- regularities between forms and range of possible senses are not captured

Decompositional Theories

A classic approach to defining word meaning is to decompose it into a set of *semantic primitives* or *features*

- dates back to Leibniz, and was very popular in the 70's/80's within the tradition of Generativist Semantics initiated by Katz & Fodor (1963)

hen	+chicken	+adult	+female
rooster	+chicken	+adult	-female
chick	+chicken	-adult	

$$\textit{kill}(x, y) \Leftrightarrow \text{CAUSE}(x, \text{BECOME}(\text{NOT}(\text{ALIVE}(y))))$$

- for instance, Wilkes advocated for a set of semantic primitives that could be used for machine translation

Katz & Fodor (1963) The Structure of a Semantic Theory. *Language*, 39(2):170–210.

Wilks, Y. (1972) *Grammar, Meaning and the Machine Analysis of Language*. London and Boston: Routledge.

Wierzbicka A. (1996) Semantics: Primes and Universals. Oxford University Press

Problems. . .

- Is it at all possible to define a finite, universal set of semantic primitives that can characterise all word senses?
  - What is the ontological status of these primitives?
- Is it at all possible to define a set of necessary and sufficient applicability conditions for word senses?
  - This has been fiercely criticised by cognitive psychologists like Rosch, who point to typicality effects
- Due to these problems, most current computational linguistics work does not use semantic primitives.
  - although work on semantic role labelling relies on a set of thematic roles (AGENT, THEME, INSTRUMENT, GOAL. . .)

Dictionary definitions

- If primitives don't work, how can we define what a word sense means?
- Some examples from the *American Heritage Dictionary* (from Martin & Juraftsky, 2009):

right	*adj.* located nearer the right hand esp. being on the right when facing the same direction as the observer.
left	*adj.* located nearer to this side of the body than the right.
red	*n.* the color of blood or a ruby.
blood	*n.* the red liquid that circulates in the heart, arteries and veins of animals.

- dictionary definitions are circular, but they are still useful to users who have a sufficient grasp of most words
- they define a sense through its relationship with other senses
- this approach is widely used in computational work – it underlies online databases like WordNet.

The Relational Approach

Relational theories of lexical meaning attempt to capture how lexical items are logically related to each other.

- Formulation of *meaning postulates* (introduced by Carnap (1956)) that describe analytical truths about word senses.
- Characterisation of word senses in terms of the inferences they license.

$$\begin{array}{rll}
\text{raven} & \forall x.Raven(x) \rightarrow Black(x) & \approx \quad Raven \subset Black \\
\text{dolphin} & \forall x.Dolphin(x) \leftrightarrow Mammal(x) \wedge Can(x, Swim(x)) \wedge ... & \\
\text{seek} & \forall x \forall y.Seek(x,y) \leftrightarrow Try(x, Find(x,y))) & \\
\text{kill} & \forall x \forall y.Kill(x,y) \leftrightarrow Cause(x, Become(y, \neg Alive(y))) &
\end{array}$$

- Under this view, the sense of an expression is considered to be the set of its lexical entailments
- The lexical entailments of a word W in a sentence S are all the entailments of S that are exclusively due to W.

Rudolf Carnap (1956) *Meaning and Necessity*, University of Chicago Press.

Main Semantic Relations

- Synonymy: relation of semantic identity (or near identity) between senses, e.g. *'aurora/dawn/sunrise'*, *'whore/prostitute'*, *'big/large'*
- Antonymy: relation of semantic oppositeness between senses, e.g. *'tall/short'*, *'dead/alive'*, *'up/down'*
- Hyponymy and Hypernymy: relation of semantic inclusion that holds between a more general term such as *'bird'* and a more specific term such as *'seagull'*
- Meronymy: part-whole relation between senses, e.g. *'elbow/arm'*, *'keyboard/computer'*

Semantic Relations & Lexical Entailment

Some semantic relations can be characterised in terms of lexical entailment.

- Synonymy (assuming there are true synonyms)
  - Two expressions A and B are synonymous if and only if they have the same lexical entailments
  - or $\forall x[A(x) \leftrightarrow B(x)]$
- Hyponymy and Hypernymy
  - A is a hyponym of B iff the lexical entailments of B are a proper subset of the lexical entailments of A

For instance, *to devour* is an hyponym of *to eat* (it is a special way of eating):

X devours Y	X eats Y
(e.g. they devoured the cake)	→ X does something
→ X eats Y	→ Y disappears
→ X acts quickly	→ X causes Y to disappear
...	...

  - So if A is a hyponym of B (and hence B is a hypernym of A) then $\forall x[A(x) \rightarrow B(x)]$ (or $\forall x \forall y[A(x,y) \rightarrow B(x,y)]$ for two arguments, etc.)

WordNet

WordNet is a lexical database created to deal with tasks that require knowledge of lexical semantics. It can be searched online at `http://wordnetweb.princeton.edu/perl/webwn`

- what kind of words (part of speech) are included in WordNet?
- how is WordNet organised?
- what are *synsets*?
- what semantic relation are covered?
- ...

WordNet: An Example

Noun

- S: (n) **port** (a place (seaport or airport) where people and merchandise can enter or leave a country)
- S: (n) **port**, port wine (sweet dark-red dessert wine originally from Portugal)
- S: (n) **port**, embrasure, porthole (an opening (in a wall or ship or armored vehicle) for firing through)
- S: (n) larboard, **port** (the left side of a ship or aircraft to someone who is aboard and facing the bow or nose)
- S: (n) interface, **port** ((computer science) computer circuit consisting of the hardware and associated circuitry that links one device with another (especially a computer and a hard disk drive or other peripherals))

Verb

- S: (v) **port** (put or turn on the left side, of a ship) *"port the helm"*
- S: (v) **port** (bring to port) *"the captain ported the ship at night"*
- S: (v) **port** (land at or reach a port) *"The ship finally ported"*
- S: (v) **port** (turn or go to the port or left side, of a ship) *"The big ship was slowly porting"*
- S: (v) **port** (carry, bear, convey, or bring) *"The small canoe could be ported easily"*
- S: (v) **port** (carry or hold with both hands diagonally across the body, especially of weapons) *"port a rifle"*
- S: (v) **port** (drink port) *"We were porting all in the club after dinner"*
- S: (v) **port** (modify (software) for use on a different machine or platform)

Adjective

- S: (adj) **port**, larboard (located on the left side of a ship or aircraft)

WordNet: An Example

Noun

- S: (n) **port** (a place (seaport or airport) where people and merchandise can enter or leave a country)
- S: (n) **port**, port wine (sweet dark-red dessert wine originally from Portugal)
  - *direct hypernym* / ***inherited hypernym*** / *sister term*
    - S: (n) fortified wine (wine to which alcohol (usually grape brandy) has been added)
      - S: (n) wine, vino (fermented juice (of grapes especially))
        - S: (n) alcohol, alcoholic drink, alcoholic beverage, intoxicant, inebriant (a liquor or brew containing alcohol as the active agent) *"alcohol (or drink) ruined him"*
          - S: (n) beverage, drink, drinkable, potable (any liquid suitable for drinking) *"may I take your beverage order?"*
            - S: (n) food, nutrient (any substance that can be metabolized by an animal to give energy and build tissue)
              - S: (n) substance (a particular kind or species of matter with uniform properties) *"shigella is one of the most toxic substances known to man"*
                - S: (n) matter (that which has mass and occupies space) *"physicists study both the nature of matter and the forces which govern it"*
                  - S: (n) physical entity (an entity that has physical existence)
                    - S: (n) entity (that which is perceived or known or

Lexical Ambiguity: One Form, Several Senses

Not all ambiguous forms exhibit the same kind of ambiguities:

- Homonymy or *contrastive ambiguity*: accidental ambiguity between unrelated senses; one sense invalidates the other:

(1) a. Mary walked along the bank of the river.
b. ABN-AMRO is the richest bank in the city.

(2) a. Nadia's plane taxied to the terminal.
b. The central data storage device is served by multiple terminals. c. He disliked the angular planes of his cheeks and jaw.

- Polysemy or *complementary ambiguity*: ambiguity between semantically related senses that overlap:

(3) a. The bank raised its interest rates yesterday.
b. The store is next to the newly constructed bank.

(4) a. John crawled through the window.
b. The window is closed.

(5) a. The lamb is running in the field.
b. John ate lamb for dinner.

(6) a. John spilled coffee on the newspaper
b. The newspaper fired its editor.

Polysemy vs. Homonymy

In dictionaries, it is common to group polysemous senses within one lexical entry and to include a different lexical entry for each homonymous sense or group of senses.

bank[1] [bangk] Show IPA

–noun

1. a long pile or heap; mass: *a bank of earth; a bank of clouds.*
2. a slope or acclivity.
3. *Physical Geography* . the slope immediately bordering a stream course along which the water normally runs.

bank[2] [bangk] Show IPA

–noun

1. an institution for receiving, lending, exchanging, and safeguarding money and, in some cases, issuing notes and transacting other financial business.
2. the office or quarters of such an institution.

http://www.dictionary.com/

Problems of Sense Enumeration

Pustejovsky argues against this enumerative view of the lexicon:

- ambiguity seems to be an inherent feature of word meaning
- the number of senses increases with the frequency of a word
- there is no clear upper boundary for the possible set of senses
- words can take an infinite number of meanings in novel contexts
  - ∗ particularly apparent with adjectives such as *'good'* or *'fast'*, which take novel senses depending on the nominal they modify.

good [good] Show IPA ***adjective,*** bet·ter, best, ***noun, interjection, adverb***

–adjective

1. morally excellent; virtuous; righteous; pious: *a good man.*
2. satisfactory in quality, quantity, or degree: *a good teacher; good health.*
3. of high quality; excellent.
4. right; proper; fit: *It is good that you are here. His credentials are good.*
5. well-behaved: *a good child.*

...plus 36 other senses

Regular Polysemy

Another aspect not captured by enumerative approaches is that polysemous senses exhibit *systematic alternations*:

(7) Count/Mass alternations: *lamb, dear, rabbit, chicken*
   a. The lamb is running in the field.
   b. John ate lamb for dinner.

(8) Container/Content: *bottle, glass, box*
   a. Mary broke the bottle.
   b. The baby finished the bottle.

(9) Figure/Ground: *door, window, fireplace*
   a. The window is rotting.
   b. Alex crawled through the window.

(10) Product/Producer: *newspaper, Honda*
   a. The newspaper fired its editor.
   b. John spilled coffee on the newspaper.

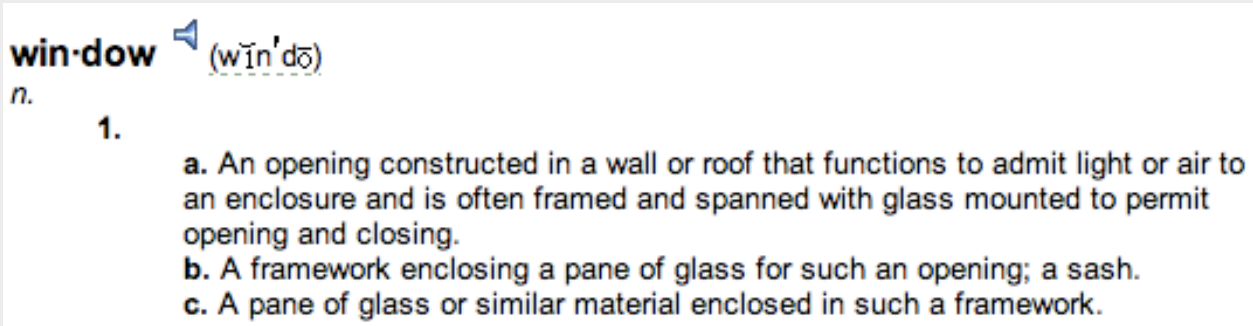
The different senses seem to be somehow present simultaneously, with one of them being *focused* in a particular context.

Apresjan, J. (1974). Regular Polysemy. Linguistics, 142:5–32.

Sense Overlap

Furthermore, Pustejovsky argues that polysemous senses are *permeable*: they may overlap, with a single use being able to denote two senses:

(11) John crawled through the broken window.



win·dow (wĭn'dō)

n.

1.

a. An opening constructed in a wall or roof that functions to admit light or air to an enclosure and is often framed and spanned with glass mounted to permit opening and closing.
b. A framework enclosing a pane of glass for such an opening; a sash.
c. A pane of glass or similar material enclosed in such a framework.

The Generative Lexicon

Pustejovsky has proposed a framework, the *Generative Lexicon*, that aims at explaining these features of polysemy.

- It reacts against enumerative models of the lexicon to propose a *generative* theory of lexical meaning.
- It proposes a method for the *decomposition* of semantic categories that can explain the *generation* of interpretations in particular contexts.
- Rather than assuming a fixed set of *semantic primitives* and defining senses with sets of features, GL assumes:
  - *structured forms or templates* common to all lexical items, and
  - a set of *compositional devices*.

Pustejovsky (1991) The Generative Lexicon, *Computational Linguistics*, 17(4):409–441.

Pustejovsky (1995) *The Generative Lexicon*, MIT Press.

Levels of Semantic Representation in GL

1. Argument Structure: Information about arity and type of arguments for a predicate.
2. Event Structure: Information about event types for a predicate, e.g. *state*, *process*, *transition*
3. Qualia Structure: A representation of the defining attributes of an entity, e.g. its constitutive parts, purpose and function, mode of creation, etc.
   * constitutive / formal / telic / agentive
4. Inheritance Structure: Information about how a word is related to other items in the lexicon. Interface with conceptual structure.

Qualia Structure

- The set of properties or events that best describes what words mean $\rightsquigarrow$ inspired by Aristotle's *modes of explanation or aitiae*.
- All categories express a qualia structure, but not all lexical items carry a value for each qualia role. Very important for nominals.
  - Constitutive: information about the constituent parts of an object.
    'house' [CONST = windows, rooms...] *'hand'* [CONST = part-of-body...]
  - Formal: distinctive features of objects
    'house' [FORMAL = building, size x, shape y...]
  - Telic: purpose and function
    'house' [TELIC = living-in] *'novel'*= [TELIC = reading]
  - Agentive: factors involved in the origin or creation of entities
    'house' [AGENT = building] *'novel'*= [AGENT = writing]

Qualia Structure and Composition

According to Generative Lexicon framework, the Qualia Structure is what allows us to how words enter into compositional processes that generate creative senses.

wooden	arrow	CONST
large	arrow	FORMAL
useful	arrow	TELIC
carved	arrow	AGENT

Can you *shine* the lamp over here?	TELIC
Mary *hung* the lamp in the kitchen.	FORMAL
John *assembled* the lamp.	AGENTIVE

Generative Devices

GL aims at capturing the means by which words can assume a potentially infinite number of senses in context.

1) Type Coercion: a predicate converts its argument to the right type, exploiting the qualia structure of its argument.

(12) a. John began to read a book.
b. John began a book.

$\rightsquigarrow$ the type of *'book'* is coerced into an event type by exploiting its TELIC (reading) and AGENTIVE (writing) roles.

2) Co-composition: the complement co-specifies the verb; the composition of qualia structures results in a derived sense of the verb.

(13) a. John baked the potato / the cake.
b. Mary painted the wall / a picture .

$\rightsquigarrow$ *change of state* vs. *creation* sense due to the AGENTIVE role of *'cake'* (baking-act) and *'picture'* (painting-act).

Generative Devices

3) Selective Binding: a predicate (typically an adjective) selects a particular quale within the noun it modifies. The composition results in a new sense for the predicate.

(14) a. a fast boat / a fast train / a fast typist
b. a good knife / a good teacher / a good child

$\rightsquigarrow$ very productive and apparently non-compositional.

Pustejovsky (1991) The Generative Lexicon, *Computational Linguistics*, 17(4):409–441.

Pustejovsky (1995) *The Generative Lexicon*, MIT Press.

Summing Up

We have seen three approaches to word meaning and the organisation of the lexicon

- decompositional approaches based on semantic primitives
- relational approaches that focus on the relationship amongst words and lexical entailments
  - WordNet is a digital database that builds on these approaches
- generative approaches (the Generative Lexicon) that are decompositional but propose mechanisms to explain ambiguity regularities and the creativity of word senses

Next, we will look into how some of these ideas have been taken up by computational linguists.

Readings for Wednesday

Please read and prepare for discussion the following papers:

- Lapata (2001) A Corpus-based Account of Regular Polysemy: The Case of Context-sensitive Adjectives, in *Proc. of NAACL*.
- Utt & Pado (2011) Ontology-based Distinction between Polysemy and Homonymy, *Proc. of the 9th International Conference on Computational Semantics*.

Links to the papers on the COSP website.

* what's the topic under investigation?
* what's the approach adopted and motivation for it?
* what's the main proposal/contribution of the paper?
* how is it evaluated and what are the results?
* ...
* search for any notions which are novel to you
* be ready to ask in class those aspects you did not understand

Relevant Seminar

Friday 23 November, 4pm, VOC-zaal, Kloveniersburgwal 48

SMART Cognitive Science Lecture by **Mark Steedman** on
Using Linguistic Knowledge in Natural Language Processing
with an introduction by Khalil Sima'an
http://smartcognitivescience.wordpress.com/

For an abstract of the talk, visit `http://smartcognitivescience.wordpress.com/2012/11/14/november-23d-mark-steedman-on-using-linguistic-knowledge-in-natural-language-processing/`