Computational Semantics and Pragmatics

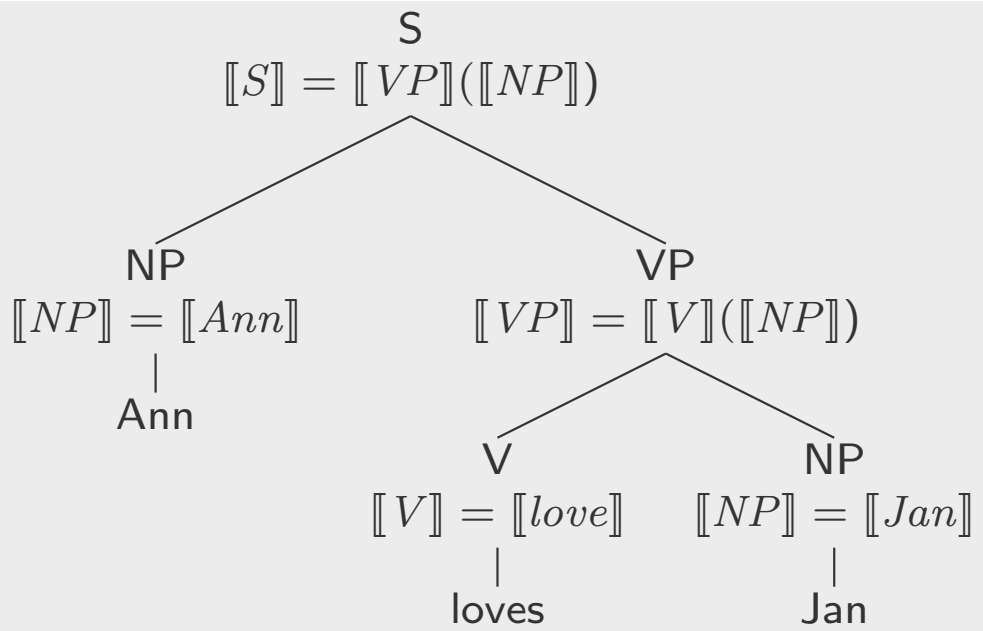
Autumn 2013

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

Formal Semantics

- Contemporary formal semantics is based on the work of Montague
  - *English as a Formal Language* (1970)
  - *Universal Grammar* (1970)
  - *The Proper Treatment of Quantification in Ordinary English* (1974)
- Focus on compositional semantics $\approx$ the computation of propositional meaning at the sentence level.

$[\![Ann]\!] = a$
$[\![Jan]\!] = j$
$[\![love]\!] = \lambda x \lambda y.Love(x, y)$



Formal Semantics: Key Components

- Meaning representations: formal language, e.g. FOL
- Principle of compositionality
  - meaning of non-sentential components: lambda expressions
  - semantic composition: functional application
- Syntax-driven semantic composition
  - input: a parse tree (given a grammar and a parser)
  - output: a logical formula
    - whose truth can be evaluated on a model (of a situation/the world)
    - which can be used for reasoning

[MoL course *Structures for Semantics*]

Computational Formal Semantics

Computational counterpart of formal semantics: automatic computation of semantic representations.

Existing implementations / frameworks:

- Patrick Blackburn & Johan Bos
  Representation and Inference for a Natural Language
  http://www.blackburnbos.org Prolog implementation
- C&C Tools: efficient and robust tools for large-scale NLP tasks
  CCG + Boxer (Johan Bos's DRT-based semantic analyzer)
  http://svn.ask.it.usyd.edu.au/trac/candc
- Jan van Eijk & Christina Unger
  Computational Semantics with Functional Programming
  http://www.computational-semantics.eu Haskell implementation
- Steven Bird, Ewan Klein, and Edward Loper
  Natural Language Processing with Python: Analyzing Text with the Natural Language Toolkit http://nltk.org (chapter 10)

Compositional Semantics

Formal (compositional) semantics very often employs a rather crude notion of lexical meaning:

$$\begin{array}{lll} [\![dolphin]\!] = \{x \mid x \text{ is a dolphin}\} & f : D \rightarrow \{1, 0\} & \langle e, t \rangle \\ [\![envy]\!] = \{\langle x, y \rangle \mid x \text{ envies } y\} & f : D \rightarrow (D \rightarrow \{1, 0\}) & \langle e, \langle e, t \rangle \rangle \end{array}$$

- Focus of compositional semantics: how the truth-conditional meaning of sentences is compositionally built from the semantic value of individual expressions.
- Focus on function words – the *glue* required for composition.
- Content words are considered "basic expressions" – not a lot of emphasis is put on them.

$\Rightarrow$ In contrast, lexical semantics is concerned with word meaning.

Lexical Semantics

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 its simplest form, the lexicon is an inventory of words, mapping word forms to word meanings (but not one-to-one):
  - synonymy: different forms mapped to one meaning
  - homonymy/polysemy: one form mapped to different meanings
- word form $\approx$ lemmas; word meanings $\approx$ word senses
- In compositional theories, the basic expressions are assumed to be disambiguated lemmas:
  - *'bank'*$_1$: the slope of land adjoining a body of water
  - *'bank'*$_2$: a business establishment in which money is kept

$\Rightarrow$ enumerative view of the lexicon

Key Issues

- How can we model word senses?
- How can we model and deal with lexical ambiguity?

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	

- 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?
  - Fiercely criticised by cognitive psychologists (as we will see)
- 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 Jurafsky & Martin 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

Logical 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.

raven	$\forall x.Raven(x) \rightarrow Black(x)$	$\approx$	$Raven \subset Black$
dolphin	$\forall x.Dolphin(x) \leftrightarrow Mammal(x) \wedge Swim(x) \wedge ...$		
seek	$\forall x \forall y.Seek(x,y) \leftrightarrow Try(x, Find(x,y)))$		
kill	$\forall x \forall y.Kill(x,y) \leftrightarrow Cause(x, Become(y, \neg Alive(y)))$		

- 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.

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.)

Key Issues

- How can we model word senses?
- How can we model and deal with lexical ambiguity?

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 fild.
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 with Sense Enumeration

Several authors have argued 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.

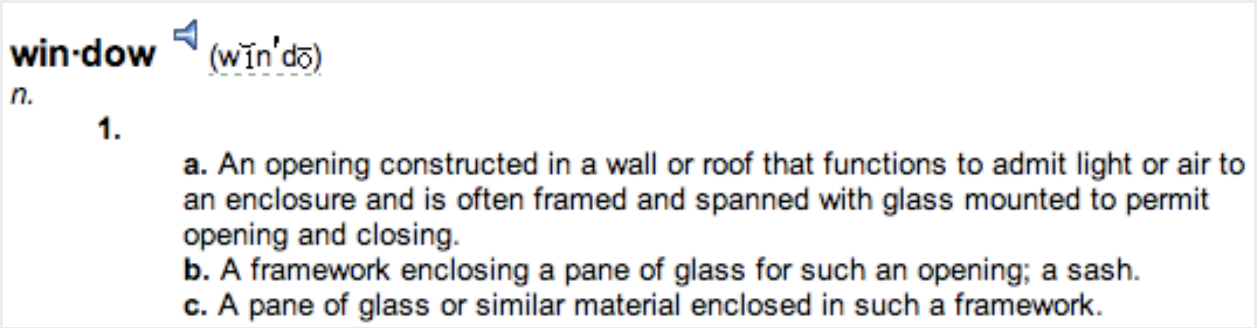
The different senses seem to be somehow present simultaneously, with one of them being *focussed* in a particular context.

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

Sense Overlap

Further more, 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

James 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

Which component of the noun is modified in each case?

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

Generative Devices

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

- 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 coerce into an event type by exploiting its TELIC (reading) and AGENTIVE (writing) roles.

- Selective Binding: a predicate selects a particular quale within the noun it takes as argument.

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

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

$\rightsquigarrow$ the composition may result in a new sense for the predicate; very productive.

Other non-Enumerative Approaches

It is worth noting that there is a different trend of approaches to polysemy that are more inherently pragmatic:

- core semantic meaning encoded in the lexicon
- the interpretation of a word in use is a consequence of general pragmatic principles of cooperative behaviour
  - maximize relevance, minimize effort
- Main references:

McCawley (1978) Conversational implicature and the lexicon, in Cole (ed.) *Syntax and Semantics 9: Pragmatics*, New York: Academic Press.

Wilson (2003) Relevance and lexical pragmatics, *Italian Journal of Linguistics*, 15: 273–91.

Blutner (2004) Pragmatics and the lexicon, in Horn & Ward (eds) *Handbook of Pragmatics*, Oxford: Blackwell.

Huang (2009) Neo-Gricean Pragmatics and the Lexicon, *International Review of Pragmatics*, 118–153.

These approaches have not had an impact in computational linguistics to date and we will not cover them.

[MoL course *Pragmatics and the Lexicon* by Henk Zeevat]

To Do

- Tomorrow we will also discuss the work of Lapata on polysemy

Maria Lapata. 2001. A Corpus-based Account of Regular Polysemy: The Case of Context-sensitive Adjectives. In Proceedings of NAACL, 63–70. Pittsburgh, PA.

Maria Lapata (2001) *The Acquisition and Modeling of Lexical Knowledge: A Corpus-based Investigation of Systematic Polysemy*. PhD Thesis, University of Edinburgh.

- If you don't know what **WordNet** it, find out by tomorrow:
  - what kind of words (part of speech) are included in WordNet?
  - how is WordNet organised: what are *synsets*?
  - what semantic relations are represented within WordNet?

- Install NLTK and go over section 2.5 (WordNet) of the NLTK book `nltk.org/book`
  - not necessarily by tomorrow, but soon.