Introduction to Natural Language Processing

Chapter 16

Outline

Review

Lexical Semantics

- Meaning of Words
- Lexical Relations
- WordNet
- Thematic Roles
- Selectional Restrictions
- Conceptual Dependency

Review: Meaning Representations

The approach to meaning that we'll present is based on the systematic creation of meaning representations–representations that bridge that gap from linguistic forms to knowledge of the world.

Review: Semantic Analysis

Semantic analysis is the process of taking in some linguistic form and producing a meaning representation for it.

Most methods rely in some way on a prior or concurrent syntactic analysis (parse).

We'll outline a compositional rule-to-rule approach.

Review: Compositional Semantics

At the core of most methods is the principle compositionality which states that the meaning of the whole is based on the meaning of parts.

The parts are the words and syntactic constituents specified by the grammar.

Review: Semantic Attachments

Given semantic attachments to context free grammar rules there are two approaches possible:

1. Alter a syntactic parser so that when completed constituents are created, the semantic function attached to the rule is applied and a meaning representation is created and stored with that state.
2. Let the parser run to completion and walk the resulting tree running the semantic attachments from the bottom-up.

The Lexicon

A lexicon contains lexemes.

The meaning of a lexeme (lexical semantics) is specified by its relation to the world and by its relation with other lexemes.

Lexical semantics is concerned with both relations among word meanings, and internal structure. Thus lexemes are NOT unanalyzable atomic symbols, as we have assumed so far.

Lexical Semantics

Meanings of words

Lexemes, lexicon, sense(s)

- a lexeme is an entry in the lexicon consisting of a pairing between form and sense(s)
- form: a particular orthographic and phonological form
- sense: a symbolic menaing representation (a set of related senses)
- lexicon: a finite list of lexemes

Examples

- Red, n: the color of blood or a ruby
- Blood, n: the red liquid that circulates in the heart, arteries and veins of animals
- Right, adj: located nearer the right hand esp. being on the right when facing the same direction as the observer

Do dictionaries give us definitions?

Lexical Relations I: Homonomy

What is homonomy?

- A *bank* holds investments in a custodial account.
- Agriculture is burgeoning on the east *bank*.

What about *bat*?

Variants

- homophones - "read" and "red"
- homographs - "bass" and "bass"

Applications

- spelling correction
- speech recognition
- text to speech

Lexical Relations II: Polysemy

What is polysemy?

- They rarely *serve* red meat.
- He *served* as U.S. ambassador.
- He might have *served* his time in prison.

Homonymy: distinct and unrelated meanings, possibly with different etymology (multiple lexemes)

Polysemy: single lexeme with multiple RELATED meanings

Example: an "idea bank"

Word Sense Disambiguation

For any given lexeme, how can its sense be reliably distinguished?

Are these the same sense of *serve*?

- Which of those flights *serve* breakfast?
- Does USAIR *serve* Pittsburgh?

Zeugma:

- ? Does USAIR *serve* breakfast and Pittsburgh?

What about

- Diane *went* to NYC.
- Diane *went* to William and Mary.

Lex. Rel. III: Metaphor, Metonymy

What is metaphor?

- That doesn't *scare* Digital.

What is metonymy?

- GM *killed* the Fiero.

Extension of existing sense to a new meaning.

Lexical Relations IV: Synonomy

What is synonomy? Substitutability.

- How *big* is that plane?
- How *large* is that plane?

Compare:

- A *big* fat apple
- ?A *large* fat apple
- A *big* sister
- ?A *large* sister

What about *home* / *house*?

Influences on substitutability:

- subtle shades of meaning differences
- polysemy
- register
- collocational constraints

Lexical Relations V: Hyponomy

What is hyponomy?

General: <u>hyponym</u>

Specific: <u>hypernym</u>

- Example: "car" is a hyponym of "vehicle" and "vehicle" is a hypernym of "car."
- Test: "That is a car" implies "That is a vehicle"

What is ontology?

What is taxonomy?

What is object hierarchy?

Semantic Networks

Used to represent relationships between words

Example: WordNet - created by George Miller's team at Princeton

- http://www.cogsci.princeton.edu/~wn

WordNet (1.6)

WordNet is the most widely used hierarchically organized lexical database for English – Fellbaum (1998).

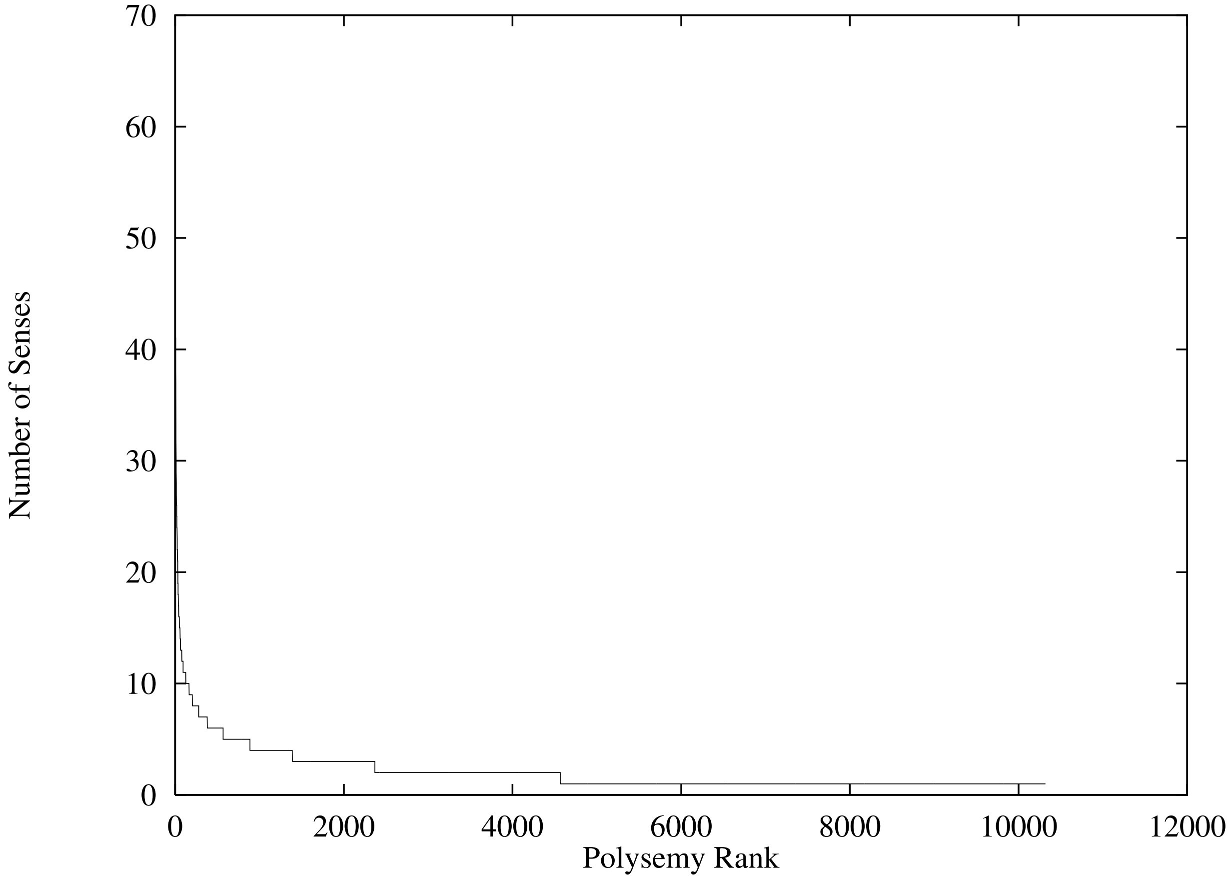
Category	Unique Forms	Number of Senses
Noun	94474	116317
Verb	10319	22066
Adjective	20170	29881
Adverb	4546	5677

Format of WordNet Entries

WordNet sense entries consist of a set of synonyms, a dictionary-style definition (or gloss), and some examples of uses.

The noun "bass" has 8 senses in WordNet.
1. bass - (the lowest part of the musical range)
2. bass, bass part - (the lowest part in polyphonic music)
3. bass, basso - (an adult male singer with the lowest voice)
4. sea bass, bass - (flesh of lean-fleshed saltwater fish of the family Serranidae)
5. freshwater bass, bass - (any of various North American lean-fleshed freshwater fishes especially of the genus Micropterus)
6. bass, bass voice, basso - (the lowest adult male singing voice)
7. bass - (the member with the lowest range of a family of musical instruments)
8. bass - (nontechnical name for any of numerous edible marine and freshwater spiny-finned fishes)

Sense Distribution for WordNet Verbs



Lexical (N) Relations in WordNet

Relation	Definition	Example
Hypernym	From concepts to superordinates	*breakfast* → *meal*
Hyponym	From concepts to subtypes	*meal* → *lunch*
Has-Member	From groups to their members	*faculty* → *professor*
Member-Of	From members to their groups	*copilot* → *crew*
Has-Part	From wholes to parts	*table* → *leg*
Part-Of	From parts to wholes	*course* → *meal*
Antonym	Opposites	*leader* → *follower*

Verb Relations in WordNet

Relation	Definition	Example
Hypernym	From events to superordinate events	*fly* $\rightarrow$ *travel*
Troponym	From events to their subtypes	*walk* $\rightarrow$ *stroll*
Entails	From events to the events they entail	*snore* $\rightarrow$ *sleep*
Antonym	Opposites	*increase* $\Longleftrightarrow$ *decrease*

Adj. and Adv. Relations in WordNet

Relation	Definition	Example
Antonym	Opposite	*heavy* $\Longleftrightarrow$ *light*
Adverb	Opposite	*quickly* $\Longleftrightarrow$ *slowly*

Synsets in WordNet

WordNet is organized around the notion of synset.

{ chump, fish, fool, gull, mark, patsy, fall guy, sucker, schlemiel, shlemiel, soft touch, mug }

Important: It is this exact synset that makes up one of the sense for each of the entries listed in the synset.

Theoretically, each synset can be viewed as a concept in a taxonomy – like the concepts described in Chapter 14.

Hyponomy in WordNet

```
Sense 3
bass, basso --
(an adult male singer with the lowest voice)
=> singer, vocalist
   => musician, instrumentalist, player
      => performer, performing artist
         => entertainer
            => person, individual, someone...
               => life form, organism, being...
                  => entity, something
               => causal agent, cause, causal agency
                  => entity, something

Sense 7
bass --
(the member with the lowest range of a family of
musical instruments)
=> musical instrument
   => instrument
      => device
         => instrumentality, instrumentation
            => artifact, artefact
               => object, physical object
                  => entity, something
```

Outline

Lexical Semantics Papers

Review Meaning of Words

- Lexicons, Lexemes, Forms, Senses
- Lexical Relations
- WordNet

Today

- WordNet demo
- Thematic Roles
- Selectional Restrictions
- Conceptual Dependency
- Framenet demo

Other Lexical Resources

The AI encyclopedia: CYC

Machine Readable Dictionaries: e.g., "stock" from the LDOCE

- 0100 a supply (of something) for use: a good stock of food
- 0200 goods for sale: Some of the stock is being taken without being paid for
- ...
- 1200 a liquid made from the juices of meat, bones, etc.,, used in cooking

Roget's Thesaurus

Limits of Hand-Encoded Resources

Manual resource construction is very costly

The coverage problem, e.g., words from the LOB corpus are not in the Oxford Dictionary

- proper noun: Chateau-Chalon
- foreign word: perestroika
- code: R101
- non-standard English: Havin'
- hyphen-omitted: bedclothes
- technical vocabulary: normoglycaemia

Internal Structure of Words

What are the meaning components underlying word sense?

Recall that a predicate-argument structure was the foundation of our meaning representation.

We assumed that certain classes of lexemes tend to contribute the predicate and the predicate argument structure, while others contribute the arguments.

The meaning representations associated with lexemes have internal structure and it is these structures, together with a grammar, that determine the relations among lexemes in a grammatical sentence.

Thematic Roles (theta-roles)

What is a thematic role?

$$\exists w, x, y, z Giving(x) \wedge Giver(w, x) \wedge Givee(z, x) \wedge Given(y, x)$$

$$\exists w, x, z Breaking(x) \wedge Breaker(w, x) \wedge Broken(z, x)$$

Generic Thematic Roles

Thematic Role	Definition
AGENT	The volitional causer of an event
EXPERIENCER	The experiencer of an event
FORCE	The non-volitional causer of the event
THEME	The participant most directly affected by an event
RESULT	The end product of an event
CONTENT	The proposition or content of a propositional event
INSTRUMENT	An instrument used in an event
BENEFICIARY	The beneficiary of an event
SOURCE	The origin of the object of a transfer event
GOAL	The destination of an object of a transfer event

Examples of Thematic Roles

Thematic Role	Example
AGENT	*The waiter* spilled the soup.
EXPERIENCER	*John* has a headache.
FORCE	*The wind* blows debris from the mall into our yards.
THEME	Only after Benjamin Franklin broke *the ice*...
RESULT	The French government has built a *regulation-size baseball diamond*...
CONTENT	Mona asked *"You met Mary Ann at a supermarket"?*
INSTRUMENT	He turned to poaching catfish, stunning them *with a shocking device*...
BENEFICIARY	Whenever Ann Callahan makes hotel reservations *for her boss*...
SOURCE	I flew in *from Boston*.
GOAL	I drove *to Portland*.

So, for "John gave Mary a book", instead of

$$\exists x, y Giving(x) \land Giver(John, x) \land Givee(Mary, x) \land Given(y, x) \land ISA(y, Book)$$

we now have

$$\exists x, y Giving(x) \land Agent(John, x) \land Goal(Mary, x) \land Theme(y, x) \land ISA(y, Book)$$

Early Theories of Thematic Roles

1967-1968: "The beginning of Lexical Semantics" (Fillmore; Gruber; Jackendoff (based on Gruber))

Two fundamentally different approaches to linguistics

- Gruber/Jackendoff: account for semantics and use grammar derived to say something about syntax
- Fillmore: account for syntax and use that to describe semantics

Thematic Level

Why posit a thematic level distinct from that of syntactic subcategorization?

This takes some of the work away from the lexicon and places it on the grammar or the semantic analyzer:

1. capture similarity between different (but related) uses of same lexical item)
2. obviate need for subcategorization frames: mapping from syntax to lexical semantics

Evidence: "John flibbed Mary the google."

- Who did it?
- What was flibbed?
- Who has the google?

Thematic Roles

Semantic roles, syntactic structures, and the prototypical places that roles appear.

AGENTS are prototypically subjects

In VPS where V NP VP, the first NP is often a GOAL, the second a THEME

Selectional Restrictions

What are selectional restrictions?

Recall the "Godzilla" example, *I want to eat someplace near campus.*

Let's assume that *eat* has agent and theme roles, what can we say about their values?

Can, and should, we extend our FOPC meaning representations?

Selectional Restriction Implementation

A WordNet approach: hamburgers are edible

```
Sense 1
hamburger, beefburger --
(a fried cake of minced beef served on a bun)
=> sandwich
   => snack food
      => dish
         => nutriment, nourishment, sustenance...
            => food, nutrient
               => substance, matter
                  => object, physical object
                     => entity, something
```

What about ...

- *If you want to kill the Soviet Union, get it to try to eat Afghanistan.*
- *But it fell apart in 1931, perhaps because people realized you can't eat gold for lunch if you're hungry.*

Can we use hyponymy to discover soft constraints and give a probability distribution (Resnik)?

Primitive Decomposition

Jim killed his philodendron

Jim did something to cause his philodendron to become not alive

Schank's Primitives

Conceptual Dependency

Primitive	Definition
ATRANS	The abstract transfer of possession or control from one entity to another.
PTRANS	The physical transfer of an object from one location to another
MTRANS	The transfer of mental concepts between entities or within an entity.
MBUILD	The creation of new information within an entity.
PROPEL	The application of physical force to move an object.
MOVE	The integral movement of a body part by an animal.
INGEST	The taking in of a substance by an animal.
EXPEL	The expulsion of something from an animal.
SPEAK	The action of producing a sound.
ATTEND	The action of focusing a sense organ.

Pred. Independence vs. Dependence

Predicate-Independent

- single set of roles is chosen independent of the type of predicates involved (no reference to type of predicates)
- Schank

Predicate-Dependent

- roles identified by particular positions arguments occupy with respect to primitive predicates

Decomposition vs. Non-Decomposition

Decomposition / Compositional Approach (Schank, Jackendoff) vs. Non-decomposition / Noncomposition Approach (Fillmore)

Within compositional approaches: exhaustive (Schank) vs. nonexhaustive (Jackendoff)

Schank: Motivation

Underlying Motivation: "Strong AI"

Focus: understanding. Argues that the representation is reversible.

Rejects syntax during analysis. Allows it during generation.

Attempts to come up with well-defined system of rules and conceptualizations.

Inferences, expectation, syntax, conversational norms, real world.

Conceptual Structure (CD): Language-independent conceptual level.

Schank: Problem 1

"John caused Mary to die" vs. "John killed Mary"

Identically substitutable?

Flaw of all compositional approaches of this nature.

Schank: Problem 2

The decompositions are very complex.

- Too specific
- Why are these conceptualizations so radically distinct from the syntactic realization?
- Talks CD from NL understanding point of view -what about generation?

For Next Time

Parts of Chapter 17