CS460/626 : Natural Language Processing/Speech, NLP and the Web
(Lecture 21, 23– Semantics)

Pushpak Bhattacharyya
CSE Dept.,
IIT Bombay
27\textsuperscript{th} Feb and 1\textsuperscript{st} March, 2012
(lecture 22 was on UNL by Janardhan)
Layers of NLP

- Morphology
- POS tagging
- Chunking
- Parsing
- Semantics
- Discourse and Coreference

Increased Complexity Of Processing

- CRF
- HMM
- MEMM

Problem

- Semantics
- Parsing
- Part of Speech Tagging

Language

- Hindi
- English
- French

NLP Trinity

Algorithm
“Delhi is the Capital of India”

**Paradigmatic knowledge**
- Delhi instance-of city
- Capital is-a Center-of-authority
- India instance-of country

**Syntagmatic knowledge**
- aoj(capital, Delhi)
- Mod(capital, India)
“How many red blocks are there?”
(combined paradigmatic and syntagmatic)

[S:00]
{org:en}
How many red blocks are there?
{/org}
{unl}
{unl}
mod(red(icl>color>property).@entry.@present.@interrogative,
 how_many(icl>how,com>number))
aoj(red(icl>color>property).@entry.@present.@interrogative,
   block(icl>artifact>thing).@pl))
plc(red(icl>color>property).@entry.@present.@interrogative,
    there(icl>how,ant>here))
{/unl}
[/S]
Wordnet

- A lexical knowledgebase based on conceptual lookup
- Organize lexical information in terms of word meaning, rather than word form
Psycholinguistic Theory

- Human lexical memory for nouns as a hierarchy.
  - Can canary sing? - Pretty fast response.
  - Can canary fly? - Slower response.
  - Does canary have skin? - Slowest response.

Wordnet - a lexical reference system based on psycholinguistic theories of human lexical memory.
## Lexical Matrix

<table>
<thead>
<tr>
<th>Word Meanings</th>
<th>Word Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td>M1</td>
<td>E1,1</td>
</tr>
<tr>
<td></td>
<td>E1,2</td>
</tr>
<tr>
<td>M2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E2,2</td>
</tr>
<tr>
<td>M3</td>
<td></td>
</tr>
<tr>
<td></td>
<td>E3,3</td>
</tr>
<tr>
<td>Mm</td>
<td></td>
</tr>
</tbody>
</table>

### Synonymous Words

- F1
- F2
- F3
- Fn

### Polysemous Words

- Em,n
### Wordnet - Lexical Matrix (with examples)

<table>
<thead>
<tr>
<th>Word Meanings</th>
<th>Word Forms</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$F_1$</td>
</tr>
<tr>
<td>$M_1$</td>
<td>$(depend)$</td>
</tr>
<tr>
<td></td>
<td>$E_{1,1}$</td>
</tr>
<tr>
<td>$M_2$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>$M_3$</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
</tr>
<tr>
<td>$M_m$</td>
<td></td>
</tr>
</tbody>
</table>
Wordnet

- Wordnet is a network of words linked by lexical and semantic relations.
- The first wordnet in the world was for English developed at Princeton over 15 years.
- The Eurowordnet-linked structure of European language wordnets was built in 1998 over 3 years with funding from the EC as a mission mode project.
- Wordnets for Hindi and Marathi being built at IIT Bombay are amongst the first IL wordnets.
- All these are proposed to be linked into the IndoWordnet which eventually will be linked to the English and the Euro wordnets.
INDOWORDNET

- Hindi Wordnet
- Bengali Wordnet
- Dravidian Language Wordnet
- Sanskrit Wordnet
- Punjabi Wordnet
- Marathi Wordnet
- Konkani Wordnet
- English Wordnet
- North East Language Wordnet

- Urdu Wordnet
Syntagmatic vs. Paradigmatic relations?
Psycholinguistics is the basis of the design.
When we hear a word, many words come to our mind by association.
For English, about half of the associated words are syntagmatically related and half are paradigmatically related.
For cat
- animal, mammal - paradigmatic
- mew, purr, furry - syntagmatic
Stated Fundamental Application of Wordnet: Sense Disambiguation

Determination of the correct sense of the word
The crane ate the fish vs.
The crane was used to lift the load
bird vs. machine
The problem of Sense tagging

- Given a corpora **To Assign correct sense to the words.**
- This is sense tagging. Needs **Word Sense Disambiguation (WSD)**
- Highly important for Question Answering, Machine Translation, Text Mining tasks.
Classification of Words

Word
- Content Word
  - Verb
  - Noun
  - Adjective
  - Adverb
- Function Word
  - Preposition
  - Conjunction
  - Pronoun
  - Interjection
Example of sense marking: its need

(According to a new research, those people who have a busy social life, have larger space in a part of their brain).

(As per a recent research, those individuals who have a busy social life, have larger space in a part of their brain).
Ambiguity of लोगों (People)

- लोग, जन, लोक, जनमानस, पब्लिक - एक से अधिक व्यक्ति "लोगों के हित में काम करना चाहिए"
  - (English synset) multitude, masses, mass, hoi_polloi, people, the_great_unwashed - the common people generally "separate the warriors from the mass" "power to the people"

- दुनिया, दुनियाँ, संसार, विश्व, जगत, जहाँ, जहान, ज़माना, जमाना, लोक, दुनियावाले, दुनियाँवाले, लोग - संसार में रहने वाले लोग "महात्मा गांधी का सम्मान पूरी दुनिया करती है / में इस दुनिया की परवाह नहीं करता / आज की दुनिया पैसे के पीछे भाग रही है"
  - (English synset) populace, public, world - people in general considered as a whole "he is a hero in the eyes of the public"
Basic Principle

- Words in natural languages are polysemous.
- However, when synonymous words are put together, a unique meaning often emerges.
- Use is made of Relational Semantics.
- Componential Semantics where each word is a **bundle of semantic features** (as in the Schankian Conceptual Dependency system or Lexical Componential Semantics) is to be examined as a viable alternative.
Componential Semantics

- Consider cat and tiger. Decide on componential attributes.

<table>
<thead>
<tr>
<th></th>
<th>Furry</th>
<th>Carnivorous</th>
<th>Heavy</th>
<th>Domesticable</th>
</tr>
</thead>
</table>

- For cat (Y, Y, N, Y)
- For tiger (Y, Y, Y, N)

Complete and correct Attributes are difficult to design.
Semantic relations in wordnet

1. Synonymy
2. Hypernymy / Hyponymy
3. Antonymy
4. Meronymy / Holonymy
5. Gradation
6. Entailment
7. Troponymy

1, 3 and 5 are lexical (word to word), rest are semantic (synset to synset).
Synset: the foundation
(house)

1. **house** -- (a dwelling that serves as living quarters for one or more families; "he has a house on Cape Cod"; "she felt she had to get out of the house")
2. **house** -- (an official assembly having legislative powers; "the legislature has two houses")
3. **house** -- (a building in which something is sheltered or located; "they had a large carriage house")
4. **family, household, house, home, menage** -- (a social unit living together; "he moved his family to Virginia"; "It was a good Christian household"; "I waited until the whole house was asleep"; "the teacher asked how many people made up his home")
5. **theater, theatre, house** -- (a building where theatrical performances or motion-picture shows can be presented; "the house was full")
6. **firm, house, business firm** -- (members of a business organization that owns or operates one or more establishments; "he worked for a brokerage house")
7. **house** -- (aristocratic family line; "the House of York")
8. **house** -- (the members of a religious community living together)
9. **house** -- (the audience gathered together in a theatre or cinema; "the house applauded"; "he counted the house")
10. **house** -- (play in which children take the roles of father or mother or children and pretend to interact like adults; "the children were playing house")
11. **sign of the zodiac, star sign, sign, mansion, house, planetary house** -- ((astrology) one of 12 equal areas into which the zodiac is divided)
12. **house** -- (the management of a gambling house or casino; "the house gets a percentage of every bet")
Creation of Synsets

Three principles:

- Minimality
- Coverage
- Replacability
Synset creation (continued)

**Home**
John’s home was decorated with lights on the occasion of Christmas.
Having worked for many years abroad, John Returned home.

**House**
John’s house was decorated with lights on the occasion of Christmas.
Mercury is situated in the eighth house of John’s horoscope.
Synsets (continued)

{house} is ambiguous.
{house, home} has the sense of a social unit living together;
Is this the minimal unit?
{family, house, home} will make the unit completely unambiguous.

For coverage:
{family, household, house, home} ordered according to frequency.

Replacability of the most frequent words is a requirement.
Synset creation

From first principles

- Pick all the senses from good standard dictionaries.
- Obtain synonyms for each sense.
- Needs hard and long hours of work.
Synset creation (continued)

From the wordnet of another language in the **same family**

- Pick the synset and obtain the sense from the gloss.
- Get the words of the target language.
- Often same words can be used - especially for tatsama words.
- Translation, Insertion and deletion.

**Hindi Synset:** anubhavi, jankaar, manjaa hua (experienced person)

**Marathi Synset:** anubhavi, jankaar, tajna
Gloss and Example

Crucially needed for concept explication, wordnet building using another wordnet and wordnet linking.

**English Synset:** {earthquake, quake, temblor, seism} -- (shaking and vibration at the surface of the earth resulting from underground movement along a fault plane of from volcanic activity)

**Hindi Synset:** {भूकंप, भूचाल, भूडोल, जलजला, भूकंप, भू-कंप, भू-कंप, जलजला, भूमकंप, भूमिकंप - प्राकृतिक कारणों से पृथ्वी के भीतरी भाग में कुछ उथल-पुथल होने से ऊपरी भाग के सहस्त्र हिलने की क्रिया "२००१ में गुजरात में आये भूकंप में काफी लोग मारे गये थे"}

(shaking of the surface of earth; many were killed in the earthquake in Gujarat)

**Marathi Synset:** धरणीकंप,भूकंप - पृथ्वीच्या पोटात द्रव्यक्षोभ होऊन पृष्ठभाग हालण्याची क्रिया "२००१ साली गुजरातमध्ये झालेल्या धरणीकंपात अनेक लोक मृत्युमुखी पडले"
Semantic Relations

- Hypernymy and Hyponymy
  - Relation between word senses (synsets)
  - X is a hyponym of Y if X is a kind of Y
  - Hyponymy is transitive and asymmetrical
  - Hypernymy is inverse of Hyponymy
  (lion->animal->animate entity->entity)
Semantic Relations (continued)

- **Meronymy and Holonymy**
  - Part-whole relation, branch is a part of tree
  - X is a meronymy of Y if X is a part of Y
  - Holonymy is the inverse relation of Meronymy

{kitchen} .......................... {house}
Lexical Relation

- **Antonymy**
  - Oppositeness in meaning
  - Relation between word forms
  - Often determined by phonetics, word length etc. ({rise, ascend} vs. {fall, descend})
Dwelling, abode

- kitchen
- bedroom
- house, home
- guestroom
- veranda
- backyard
- study
- guestroom
- hermitage
- cottage

A place that serves as the living quarters of one or more families

WordNet Sub-Graph
Troponym and Entailment

- Entailment
  \{snoring \rightarrow sleeping\}

- Troponym
  \{limp, strut \rightarrow walk\}
  \{whisper \rightarrow talk\}
Entailment

Snoring entails sleeping.
Buying entails paying.

- **Proper Temporal Inclusion.**
  Inclusion can be in any way.
Sleeping temporally includes snoring.
Buying temporally includes paying.

- **Co-extensiveness.** (Troponymy)
Limping is a manner of walking.
Opposition among verbs.

- \{\text{Rise, ascend}\} \{\text{fall, descend}\}
  - Tie-untie (do-undo)
  - Walk-run (slow, fast)
  - Teach-learn (same activity different perspective)
  - Rise-fall (motion upward or downward)

- Opposition and Entailment.
  - Hit or miss (entail aim). Backward presupposition.
  - Succeed or fail (entail try.)
The causal relationship.

Show- see.
Give- have.

Causation and Entailment.
  Giving entails having.
  Feeding entails eating.
Entailment

+Temporal Inclusion

-Troponymy (Proper Inclusion)
  - Troponymy (Co-extensiveness)
    - limp-walk
    - lisp-talk
  - snore-sleep
  - buy-pay

-Temporal Inclusion

Backward Presupposition
  - succeed-try
  - untie-tie

Cause
  - raise-rise
give-have
## Kinds of Antonymy

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Size</strong></td>
<td><strong>Small - Big</strong></td>
</tr>
<tr>
<td><strong>Quality</strong></td>
<td><strong>Good – Bad</strong></td>
</tr>
<tr>
<td><strong>State</strong></td>
<td><strong>Warm – Cool</strong></td>
</tr>
<tr>
<td><strong>Personality</strong></td>
<td><strong>Dr. Jekyl- Mr. Hyde</strong></td>
</tr>
<tr>
<td><strong>Direction</strong></td>
<td><strong>East- West</strong></td>
</tr>
<tr>
<td><strong>Action</strong></td>
<td><strong>Buy – Sell</strong></td>
</tr>
<tr>
<td><strong>Amount</strong></td>
<td><strong>Little – A lot</strong></td>
</tr>
<tr>
<td><strong>Place</strong></td>
<td><strong>Far – Near</strong></td>
</tr>
<tr>
<td><strong>Time</strong></td>
<td><strong>Day - Night</strong></td>
</tr>
<tr>
<td><strong>Gender</strong></td>
<td><strong>Boy - Girl</strong></td>
</tr>
</tbody>
</table>
## Kinds of Meronymy

<table>
<thead>
<tr>
<th>Component-object</th>
<th>Head - Body</th>
</tr>
</thead>
<tbody>
<tr>
<td>Staff-object</td>
<td>Wood - Table</td>
</tr>
<tr>
<td>Member-collection</td>
<td>Tree - Forest</td>
</tr>
<tr>
<td>Feature-Activity</td>
<td>Speech - Conference</td>
</tr>
<tr>
<td>Place-Area</td>
<td>Palo Alto - California</td>
</tr>
<tr>
<td>Phase-State</td>
<td>Youth - Life</td>
</tr>
<tr>
<td>Resource-process</td>
<td>Pen - Writing</td>
</tr>
<tr>
<td>Actor-Act</td>
<td>Physician - Treatment</td>
</tr>
</tbody>
</table>
# Gradation

<table>
<thead>
<tr>
<th>State</th>
<th>Childhood, Youth, Old age</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>Hot, Warm, Cold</td>
</tr>
<tr>
<td>Action</td>
<td>Sleep, Doze, Wake</td>
</tr>
</tbody>
</table>
Metonymy

- Associated with Metaphors which are epitomes of semantics
- Oxford Advanced Learners Dictionary definition: “The use of a word or phrase to mean something different from the literal meaning”
- Does it mean Careless Usage?!
Insight from Sanskritic Tradition

- Power of a word
  - Abhidha, Lakshana, Vyanjana

- Meaning of **Hall:**
  - The hall is packed (avidha)
  - The hall burst into laughing (lakshana)
  - The Hall is full (unsaid: and so we cannot enter) (vyanjana)
Metaphors in Indian Tradition

- upamana and upameya
  - Former: object being compared
  - Latter: object being compared with
  - Puru was like a lion in the battle with Alexander (Puru: upameya; Lion: upamana)
Upamana, rupak, atishayokti

- **upamana**: Explicit comparison
  - Puru was like a lion in the battle with Alexander

- **rupak**: Implicit comparison
  - Puru was a lion in the battle with Alexander

- **Atishayokti** (exaggeration): upamana and upameya dropped
  - Puru’s army fled. But the lion fought on.
Modern study (1956 onwards, Richards et al.)

- Three constituents of metaphor
  - Vehicle (items used metaphorically)
  - Tenor (the metaphorical meaning of the former)
  - Ground (the basis for metaphorical extension)
- “The foot of the mountain”
  - Vehicle: :foot”
  - Tenor: “lower portion”
  - Ground: “spatial parallel between the relationship between the foot to the human body and the lower portion of the mountain with the rest of the mountain”
Interaction of semantic fields (Haas)

- Core vs. peripheral semantic fields
- Interaction of two words in metonymic relation brings in new semantic fields with selective inclusion of features
- Leg of a table
  - Does not stretch or move
  - Does stand and support
Lakoff’s (1987) contribution

- Source Domain
- Target Domain
- Mapping Relations
Mapping Relations: ontological correspondences

- Anger is heat of fluid in container

<table>
<thead>
<tr>
<th>Heat</th>
<th>Anger</th>
</tr>
</thead>
<tbody>
<tr>
<td>(i) Container</td>
<td>Body</td>
</tr>
<tr>
<td>(ii) Agitation of fluid</td>
<td>Agitation of mind</td>
</tr>
<tr>
<td>(iii) Limit of resistance</td>
<td>Limit of ability to suppress</td>
</tr>
<tr>
<td>(iv) Explosion</td>
<td>Loss of control</td>
</tr>
</tbody>
</table>
Image Schemas

- Categories: Container Contained
- Quantity
  - More is up, less is down: Outputs rose dramatically; accidents rates were lower
  - Linear scales and paths: Ram is by far the best performer
- Time
  - Stationary event: we are coming to exam time
  - Stationary observer: weeks rush by
- Causation: desperation drove her to extreme steps
Patterns of Metonymy

- Container for contained
  - The kettle boiled (water)
- Possessor for possessed/attribute
  - Where are you parked? (car)
- Represented entity for representative
  - The government will announce new targets
- Whole for part
  - I am going to fill up the car with petrol
Patterns of Metonymy (contd)

- Part for whole
  - I noticed several new faces in the class

- Place for institution
  - Lalbaug witnessed the largest Ganapati

Question: Can you have part-part metonymy
Purpose of Metonymy

- More idiomatic/natural way of expression
  - More natural to say the kettle is boiling as opposed to the water in the kettle is boiling
- Economy
  - Room 23 is answering (but not *is asleep)
- Ease of access to referent
  - He is in the phone book (but not *on the back of my hand)
- Highlighting of associated relation
  - The car in the front decided to turn right (but not *to smoke a cigarette)
Feature sharing not necessary

- In a restaurant:
  - J alebii ko abhi dudh chaiye (no feature sharing)
  - The elephant now wants some coffee (feature sharing)
Proverbs

- Describes a specific event or state of affairs which is applicable metaphorically to a range of events or states of affairs provided they have the same or sufficiently similar image-schematic structure