Outline:

- Lexical Semantics: introduction
- Word senses
- Relations between senses
- WordNet
  - WordNet extensions: MultiWordNet, WordNet Domains
- Semantic roles and selectional restrictions
  - PropBank
  - FrameNet
What do we mean by “word”?

- **Lexeme**: a pairing of an orthographic or phonological form with its meaning
  - Orthographic form: the way the lexeme looks on the page
  - Phonological form: the way the lexeme sounds
- **Lexicon**: a finite list of lexemes
- **Lemma**: the grammatical form used to represent a lexeme

**WORDFORMS**

\[
\begin{align*}
\{ \text{dog, dogs} & \rightarrow \text{dog} \\
\{ \text{run, runs, ran, running} & \rightarrow \text{run} 
\end{align*}
\]
What do we mean by “word”? 

- **Lexeme**: a pairing of an orthographic or phonological form with its meaning
  - **Orthographic form**: the way the lexeme looks on the page
  - **Phonological form**: the way the lexeme sounds
- **Lexicon**: a finite list of lexemes
- **Lemma**: the grammatical form used to represent a lexeme

**WORDFORMS**

\[
\text{dog, dogs} \quad \rightarrow \quad \text{dog}
\]

\[
\text{run, runs, ran, running} \quad \rightarrow \quad \text{run}
\]

**LEMMAS**

- Lemmatization may depend on the contest:
  
  I *found* ten dollars in the street yesterday.
  
  The community used the money to *found* a church.

  LEMMA: *find* (“to locate”)
  
  LEMMA: *found* (“to build”)

- Lemmas are part of speech specific

28.10.2010 – Lexical Semantics
Word senses

- The meaning of a lemma can vary enormously given the context

  *John cashed a check at the *bank*.*
  
  *Mary sat on the *bank* of the river and watched the currents.*

- A *sense* is a discrete representation of one aspect of the meaning of a word.

  *bank1:* financial institution
  
  *bank2:* sloping mound
Homonymy

John cashed a check at the **bank**.

Mary sat on the **bank** of the river and watched the currents.

**bank1:** financial institution

**bank2:** sloping mound

The senses of the word do not have any particular relation among them.
Polysemy

*John cashed a check at the **bank**.*

*Mary sat on the **bank** of the river and watched the currents.*

*Donors give blood for storage at a **bank** for emergency transfusions.*

**bank**: financial institution

**bank**: sloping mound

**bank**: repository of biological entities

The senses of the word are related semantically.
Metonymy

John cashed a check at the bank.
Mary sat on the bank of the river and watched the currents.
Donors give blood for storage at a bank for emergency transfusions.
The bank is on the corner of Regent street and Queen avenue.

**bank1:** financial institution
**bank2:** sloping mound
**bank3:** repository of biological entities
**bank4:** the building belonging to a financial institution

A subtype of polysemy: METONYMY

The use of one aspect of a concept or entity to refer to other aspects of the entity, or to the entity itself, e.g.
BUILDING ↔ ORGANIZATION
How many senses a word has?

Which of those flights *serve* breakfast?

Does Midwest Express *serve* Philadelphia?

Does Midwest Express *serve* breakfast and Philadelphia?

- **ZEUGMA:** practical technique for determining if two senses are distinct.
Homonymy

- **Homophones** (same pronunciation, different spelling)
  
  *would/wood*

- **Homographs** (same spelling, different pronunciation)

  - **bass1**: *the lowest part of the musical range*
  - **bass2**: *spiny-finned fish*
Two senses of two different lemmas are **synonyms** when their meaning is identical or nearly identical (i.e. if they are substitutable one for the other in any sentence without changing the truth conditions of the sentence).

- **couch/sofa**
- **car/automobile**
- **water/H₂O**
- **big/large**

John forgot to bring **H₂O** supplies when hiking in Palm Springs.

Sara’s **large** sister is a teacher.
Two senses are **antonyms** if they define a binary opposition, or are at opposite ends of some scale.

- long/short
- fast/slow
- rise/fall
- up/down

**REVERSIVE:** describe some sort of change or movement in opposite directions.
Relations between senses: hyponymy

One sense is a **hyponym** of another sense if the first sense is more specific, denoting a subclass of the other.

<table>
<thead>
<tr>
<th>hypernym</th>
<th>vehicle</th>
<th>fruit</th>
<th>furniture</th>
<th>mammal</th>
</tr>
</thead>
<tbody>
<tr>
<td>hyponym</td>
<td>car</td>
<td>mango</td>
<td>chair</td>
<td>dog</td>
</tr>
</tbody>
</table>

ontologies

taxonomies
Relations between senses: meronymy

Part-whole relation

A leg is a meronym of chair.
A wheel is a meronym of car.
Car is holonym of wheel.

Semantic field: a set of lexemes which cover a certain conceptual domain and which bear certain specifiable relations to one another (Lehrer, 1985)

reservation, flight, travel, buy, price, cost, fare
WordNet: an online lexical database

- WordNet is an online lexical reference system whose design is inspired by current psycholinguistic theories of human lexical memory.
- Developed at Princeton University by George Miller’s team.
- It is a public domain resource:
  
  http://wordnet.princeton.edu/

- Three separate databases:
  - Nouns
  - Verbs
  - Adjective and adverbs
  - No closed class words
Lemma Entry


Word to search for: bank

Display Options: [Select option to change] Change

Key: "S." = Show Synset (semantic) relations, "W." = Show Word (lexical) relations

Noun

- bank (clayey land (especially the slope beside a body of water): "they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the current")
- (a) depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities) "he cashed a check at the bank"; "the bank holds the mortgage on my home"
- (a) bank (a long ridge or pile): "a huge bank of earth"
- (a) bank (an arrangement of similar objects in a row or in tiers) "he operated a bank of switches"
- (a) bank (a supply or stock held in reserve for future use (especially in emergencies))
- (a) bank (the funds held by a gambling house or the dealer in some gambling games) "he tried to break the bank at Monte Carlo"
- (a) bank, cant, cambare (a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force)
- (a) savings bank, coin bank, money box, bank (a container (usually with a slot in the top) for keeping money at home) "the coin bank was empty"
- (a) bank, bank building (a building in which the business of banking transacted) "the bank is on the corner of Nassau and Witherspoon"
- (a) bank (a flight maneuver; aircraft tips laterally about its longitudinal axis (especially in turning)): "the plane went into a steep bank"

Verb

- (a) bank (tip laterally): "the pilot had to bank the aircraft"
- (a) bank (enclose with a bank): "bank roads"
- (a) bank (do business with a bank or keep an account at a bank): "Where do you bank in this town?"
- (a) bank (act as the banker in a game or in gambling)
- (a) bank (be in the banking business)
- (a) deposit, bank (put into a bank account) "She deposits her paycheck every month"
- (a) bank (cover with ashes so as to control the rate of burning): "bank a fire"
- (a) trust, swear, rely, bank (have confidence or faith in): "We can trust in God"; "Rely on your friends"; "bank on your good education"; "I swear by my grandmother's recipes"
Noun

- S1 (n) bank (spreading land (especially the slope beside a body of water): "they pulled the canoe up on the bank"); "he sat on the bank of the river and watched the currents"
- S2 (n) depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities); "he cashed a check at a bank"; "that bank holds the mortgage on my home"
- S3 (n) bank (a long ridge or pile: "a huge bank of earth"
- S4 (n) bank (an arrangement of similar objects in a row or in tiers); "he operated a bank of switches"
- S5 (n) bank (a supply or stock held in reserve for future use (especially in emergencies))
- S6 (n) bank (the funds held by a gambling house or the dealer in some gambling games); "he tried to break the bank at Monte Carlo"
- S7 (n) bank, cant, cambre (a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force)
- S8 (n) savings bank, coin bank, money box, bank (a container (usually with a slot in the top) for keeping money at home); "the coin bank was empty"
- S9 (n) bank, bank building (a building in which the business of banking transacted); "the bank is on the corner of Nassau and Witherspoon"
- S10 (n) bank (a flight maneuver, aircraft tips laterally about its longitudinal axis (especially in turning)); "the plane went into a steep bank"

Verb

- S1 (v) bank (tip laterally); "the pilot had to bank the aircraft"
- S2 (v) bank (enclose with a bank); "bank roads"
- S3 (v) bank (do business with a bank or keep an account at a bank); "Where do you bank in this town?"
- S4 (v) bank (act as the banker in a game or in gambling)
- S5 (v) bank (be in the banking business)
- S6 (v) deposit, bank (put into a bank account); "She deposits her paycheck every month"
- S7 (v) bank (cover with ashes so as to control the rate of burning); "bank a fire"
- S8 (v) trust, swear, rely, bank (have confidence or faith in); "We can trust in God"; "Rely on your friends"; "bank on your good education"; "I swear by my grandmother’s recipes"
Noun

- bank (sloping land (especially the slope beside a body of water), "they pulled the canoe up on the bank"); "he sat on the bank of the river and watched the current"
- bank (depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities), "he cashed a check at the bank"); "that bank holds the mortgage on my home"
- bank (a long ridge or pile), "a huge bank of earth"
- bank (an arrangement of similar objects in a row or in tiers), "he operated a bank of switches"
- bank (a supply or stock held in reserve for future use (especially in emergencies))
- bank (the funds held by a gambling house or the dealer in some gambling games), "he tried to break the bank at Monte Carlo"
- bank, ramp, ramp (a slope in the turn of a road or track, the outside is higher than the inside in order to reduce the effects of centrifugal force)
- savings bank, coin bank, money box, bank (a container (usually with a slot in the top) for keeping money at home), "the coin bank was empty"
- bank, bank building (a building in which the business of banking transacted), "the bank is on the corner of Nassau and Witherspoon"
- bank (a flight maneuver, aircraft tips laterally about its longitudinal axis (especially in turning)), "the plane went into a steep bank"

Verb

- bank (tip laterally), "the pilot had to bank the aircraft"
- bank (enclose with a bank), "bank roads"
- bank (do business with a bank or keep an account at a bank), "Where do you bank in this town?"
- bank (act as the banker in a game or in gambling)
- bank (be in the banking business)
- deposit, bank (put into a bank account), "she deposits her paycheck every month"
- bank (cover with ashes so as to control the rate of burning), "bank a fire"
- trust, swear, reh, bank (have confidence or faith in), "We can trust in God"); "Rely on your friends"; "bank on your good education"; "I swear by my grandmother's recipes"
The synsets

- The set of near synonym for a WordNet sense is called synset (for synonym set).
- Concepts are represented as a list of the word senses that can be used to express the concept.

- S: (n) depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities)
# Noun relations in WordNet

<table>
<thead>
<tr>
<th>RELATION</th>
<th>Also called</th>
<th>DEFINITION</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hypernym</td>
<td>Superordinate</td>
<td>From concepts to superordinate</td>
<td>breakfast → meal</td>
</tr>
<tr>
<td>Hyponym</td>
<td>Subordinate</td>
<td>From concepts to subtypes</td>
<td>meal → lunch</td>
</tr>
<tr>
<td>Member meronym</td>
<td>Has-Member</td>
<td>From groups to their members</td>
<td>faculty → professor</td>
</tr>
<tr>
<td>Has-Instance</td>
<td></td>
<td>From concepts to instances of the concept</td>
<td>composer → Bach</td>
</tr>
<tr>
<td>Instance</td>
<td></td>
<td>From instances to their concepts</td>
<td>Austen → author</td>
</tr>
<tr>
<td>Member Holonym</td>
<td>Member-Of</td>
<td>From members to their groups</td>
<td>copilot → crew</td>
</tr>
<tr>
<td>Part Meronym</td>
<td>Has-Part</td>
<td>From whole to parts</td>
<td>table → leg</td>
</tr>
<tr>
<td>Part Holonym</td>
<td>Part-Of</td>
<td>From parts to whole</td>
<td>course → meal</td>
</tr>
<tr>
<td>Antonym</td>
<td></td>
<td>Opposites</td>
<td>leader → follower</td>
</tr>
</tbody>
</table>

28.10.2010 – Lexical Semantics (20)
Noun relations in WordNet

S (a) bank (sloping land (especially the slope beside a body of water)) "they pulled the canoe up on the bank"; "he sat on the bank of the river and watched the current"
- direct hyponym / full hyponym
- direct hyponym / inherited hyponym / sister term
  - S (n) slope, incline, side (an elevated geological formation) "he climbed the steep slope"; "the house was built on the side of a mountain"
    - S (n) geological formation, formation (geology) the geological features of the earth
    - S (n) object, physical object (a tangible and visible entity; an entity that can cast a shadow) "it was full of rackets, balls and other objects"
      - S (n) physical entry (an entity that has physical existence)
      - S (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- derivationally related form
  S (n) depository financial institution, bank, banking concern, banking company (a financial institution that accepts deposits and channels the money into lending activities) "he cashed a check at the bank"; "that bank holds the mortgage on my home"
- direct hyponym / full hyponym
- member holonym
- direct hyponym / inherited hyponym / sister term
  - S (n) financial institution, financial organization, financial organisation (an institution (public or private) that collects funds (from the public or other institutions) and invests them in financial assets)
    - S (n) institution, establishment (an organization founded and united for a specific purpose)
    - S (n) organization, organisation (a group of people who work together)
      - S (n) social group (people sharing some social relation)
    - S (n) group, grouping (any number of entities (members) considered as a unit)
      - S (n) abstraction, abstract entity (a general concept formed by extracting common features from specific examples)
      - S (n) entity (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))
- derivationally related form
Verb relations in WordNet

<table>
<thead>
<tr>
<th>RELATION</th>
<th>DEFINITION</th>
<th>EXAMPLE</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Hypernym</strong></td>
<td>From events to superordinate events</td>
<td>fly9 → travels5</td>
</tr>
<tr>
<td><strong>Troponym</strong></td>
<td>From a verb (event) to a specific manner elaboration of that verb</td>
<td>walk1 → stroll1</td>
</tr>
<tr>
<td><strong>Entails</strong></td>
<td>From verbs (events) to the verbs (events) that entail</td>
<td>snore1 → sleep1</td>
</tr>
<tr>
<td><strong>Antonym</strong></td>
<td>Opposite</td>
<td>increase1 ↔ decrease1</td>
</tr>
</tbody>
</table>
Bipolar adjective structure

- swift
- prompt
- alacritous
- quick
- rapid
- fast

- dilatory
- sluggish
- leisurely
- tardy
- laggard
- slow

similar to
antonymy
## WordNet: some figures

<table>
<thead>
<tr>
<th>POS</th>
<th>lemma</th>
<th>synsets</th>
<th>word senses</th>
</tr>
</thead>
<tbody>
<tr>
<td>Noun</td>
<td>117798</td>
<td>82115</td>
<td>146312</td>
</tr>
<tr>
<td>Verb</td>
<td>11529</td>
<td>13767</td>
<td>25047</td>
</tr>
<tr>
<td>Adjective</td>
<td>21479</td>
<td>18156</td>
<td>30002</td>
</tr>
<tr>
<td>Adverb</td>
<td>4481</td>
<td>3621</td>
<td>5580</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>155287</td>
<td>117659</td>
<td>206941</td>
</tr>
</tbody>
</table>
Other theories of lexical meaning

- Lexical decomposition
- Meaning postulates
- Prototypes
- Semantic network
Lexical decomposition

- Word meaning = composition of meaning primitives
  - Ex. to buy (Jackendoff 1983)

\[
\text{[EXCH}[\text{GO} ([\text{MONEY}, \text{FROM}[k] \text{TO}[i])] ]
\]

28.10.2010 – Lexical Semantics
Meaning postulates

Mental representations of the meaning relations between words (*Fodor 1970*)

- **Ex. to buy**
  - `buy(x,y,z) → get (x,y,z)`
  - `buy(x,y,z) → pay(x,y,z)`
  - `buy(x,y,z) → choose(x,y)`
  - `buy(x,y) → sell(z,y,x)`

- **Ex. bachelor**
  - `bachelor(x) → man(x) ^ ¬ married(x)`

28.10.2010 – Lexical Semantics
Prototypes

- Word meaning = information holding of the most typical examples of a concept but not necessarily of all instances (Rosch 1975)
- *Ex. tiger*

<table>
<thead>
<tr>
<th>Property</th>
<th>Possible value</th>
<th>Default value</th>
</tr>
</thead>
<tbody>
<tr>
<td>is-a</td>
<td>feline</td>
<td></td>
</tr>
<tr>
<td>weight</td>
<td>&lt;180 kg</td>
<td>120 kg</td>
</tr>
<tr>
<td>height</td>
<td>&lt; 106 cm</td>
<td>80 cm</td>
</tr>
<tr>
<td>length</td>
<td>&lt; 250 cm</td>
<td>180 cm</td>
</tr>
<tr>
<td>colour</td>
<td>yellow with black stripes, white, black, yellow</td>
<td>yellow with black stripes</td>
</tr>
<tr>
<td>habitat</td>
<td>jungle, river banks</td>
<td>jungle</td>
</tr>
</tbody>
</table>
Semantic networks

- Word meaning = set of relations with other meanings of the lexicon (Quillian 1968)
  - Ex. to buy

![Diagram showing semantic network with words connected by relations such as antonymy and troponymy]

- buy is connected to:
  - sell
  - get
  - pay
  - take over
  - pick up
  - choose
  - entails doing
  - entails doing
What is not present in Wordnet?

- very little syntax
  - Syntax would be very useful for verb subcategorizations
- almost no links among the principal part-of-speechs
- sometimes the semantic categorization is not accurate (or too accurate)
The SemCor corpus is a subset of the English Brown corpus containing almost 700,000 words:
- all the words are tagged by PoS
- more than 200,000 content words are also lemmatized and sense-tagged according to WordNet 1.6.

In detail:
- 352 texts
- all the open class words (N,V,adj,adv) annotated with PoS, lemma and sense: 186 texts
- only verbs are annotated with lemma and sense: 166 texts

The "all-words" component:
- 359,732 tokens
  - 192,639 are semantically annotated

The "only-verbs" component
- 16,814 tokens
  - 41,497 verb occurrences are semantically annotated
Extensions of Wordnet

- Many research groups are extending WordNet in various aspects
- Multilinguality
  - MultiWordNet
  - EuroWordNet
  Based on the assumption that the meaning networks already defined for the original English version may, for the most part, be reused for other languages.
- WordNet-Domains
- WordNet-Affect
- ...
MultiWordNet

- A multilingual lexical database in which the Italian WordNet is strictly aligned with Princeton WordNet 1.6.
- Italian synsets are created in correspondence with the Princeton WordNet synsets, whenever possible, and semantic relations are imported from the corresponding English synsets.
- Allows for the access to the Spanish, Portuguese, Hebrew, Romanian and Latin WordNets.
The word "bank" has 17 senses:

<table>
<thead>
<tr>
<th>Noun</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. depository_financial_institution, bank, banking_concern, banking_company</td>
<td>(Banking) a financial institution that accepts deposits and channels the money into lending activities; &quot;he cashed a check at the bank&quot;; &quot;that bank holds the mortgage on my home&quot;</td>
<td></td>
</tr>
<tr>
<td>2. bank</td>
<td>(Geography, Geology) sloping land (especially the slope beside a body of water); &quot;they pulled the canoe up on the bank&quot;; &quot;he sat on the bank of the river and watched the current&quot;</td>
<td></td>
</tr>
<tr>
<td>3. bank</td>
<td>(Money) a supply of stock held in reserve especially for future use (especially in emergencies)</td>
<td></td>
</tr>
<tr>
<td>4. bank, bank_building</td>
<td>(Banking) a building in which commercial banking is transacted; &quot;the bank is on the corner of Nassau and Witherspoon&quot;</td>
<td></td>
</tr>
<tr>
<td>5. bank</td>
<td>(Furnish) an arrangement of similar objects in a row or manner; &quot;he operated a bank of switches&quot;</td>
<td></td>
</tr>
<tr>
<td>6. savings_bank, coin_bank, money_box, bank</td>
<td>(Money) a container (usually with a slot in the top) for keeping money at home; &quot;the coin bank was empty&quot;</td>
<td></td>
</tr>
<tr>
<td>7. bank</td>
<td>(Finance) a place where the funds held by a gambling house or the dealer in some gambling games; &quot;he tried to break the bank at Monte Carlo&quot;</td>
<td></td>
</tr>
<tr>
<td>8. bank</td>
<td>(Town, Planning, Transport) a slope in the turn of a road or track; the outside is higher than the inside in order to reduce the effects of centrifugal force</td>
<td></td>
</tr>
</tbody>
</table>

Synset: depository_financial_institution, bank, banking_concern, banking_company

Gloss: a financial institution that accepts deposits and channels the money into lending activities; "he cashed a check at the bank"; "that bank holds the mortgage on my home"
Use of MultiWordNet in NLP application

- **Information Retrieval**: synonymy relations are used for query expansion to improve the recall of IR; cross language correspondences between Italian and English synsets are used for Cross Language Information Retrieval.

- **Semantic tagging**: MultiWordNet constitutes a large coverage sense inventory which is the basis for semantic tagging, i.e. texts are tagged with synset identifiers.

- **Disambiguation**: Semantic relationships are used to measure the semantic distance between words, which can be used to disambiguate the meaning of words in texts.

- **Ontologies**: MultiWordNet can be seen as an ontology to be used for a variety of knowledge-based NLP tasks.

- **Terminologies**: MultiWordNet constitutes a robust framework supporting the development of specific structured terminologies.
# WordNet Domains

- Wordnet 1.6 annotation with domain labels (e.g. **Medicine, Architecture, Sport**) to noun synsets (Magnini and Cavaglià, 2000)

## Objectives:
- provide a context for evaluating the quality of this resource
- verify the role domain labels may have in a word
- disambiguation task, in particular as far as sense clustering is concerned

- Exploit the role of domain labels in a multilingual task

```
<table>
<thead>
<tr>
<th>Sense Number</th>
<th>Synset (Gloss)</th>
<th>Domains</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>depository financial institution, bank, banking concern, banking company</td>
<td>Economy</td>
</tr>
<tr>
<td>2</td>
<td>bank (sloping land...)</td>
<td>Geography, Geology</td>
</tr>
<tr>
<td>3</td>
<td>bank (a supply or stock held in reserve...)</td>
<td>Economy</td>
</tr>
<tr>
<td>4</td>
<td>bank, bank building (a building...)</td>
<td>Architecture, Economy</td>
</tr>
<tr>
<td>5</td>
<td>bank (an arrangement of similar objects...)</td>
<td>Factotum</td>
</tr>
<tr>
<td>6</td>
<td>savings bank, coin bank, money box, bank (a container...)</td>
<td>Economy</td>
</tr>
<tr>
<td>7</td>
<td>bank (a long ridge or pile...)</td>
<td>Geography, Geology</td>
</tr>
<tr>
<td>8</td>
<td>bank (the funds held by a gambling house...)</td>
<td>Economy, Play</td>
</tr>
<tr>
<td>9</td>
<td>bank, cant, camber (a slope in the turn of a road...)</td>
<td>Architecture</td>
</tr>
<tr>
<td>10</td>
<td>bank (a flight maneuver...)</td>
<td>Transport</td>
</tr>
</tbody>
</table>
```
Domain Label organization

- 250 domain labels collected from dictionaries
- Four level hierarchy (Dewey Decimal Classification)
- 41 “basic” domains used for the experiments

**GOALS:**
- Integrate taxonomic and domain oriented information
- Reduce polysemy
Semantic roles and selectional restrictions

- An important aspect of lexical meaning has to do with the *semantics of events*
- Importance of predicate-argument structure for representing an event
- Semantic constraints on the arguments of event predicates:
  - Semantic roles (thematic roles)
  - Selectional restrictions
Thematic roles

\[ \exists e, x, y \text{Isa}(e, \text{Breaking}) \land \text{Breaker}(e, \text{Sasha}) \land \text{BrokenThing}(e, y) \land \text{Isa}(y, \text{Window}) \]

\[ \exists e, x, y \text{Isa}(e, \text{Opening}) \land \text{Opener}(e, \text{Pat}) \land \text{OpenedThing}(e, y) \land \text{Isa}(y, \text{Door}) \]

- **Deep roles**: Breaking events have Breakers, Opening events have Openers and so on.
- To understand these events, need to know more about their **semantics**
- **Thematic roles**: capture the semantic commonalities
  - **AGENTS**: represents an abstract ideas such as volitional causation
  - **THEME**: prototypically inanimate objects affected by the action
### Commonly used thematic roles

<table>
<thead>
<tr>
<th>Thematic role</th>
<th>Definition</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>AGENT</strong></td>
<td>The volitional causer of an event</td>
<td><em>The waiter spilled the soup</em></td>
</tr>
<tr>
<td><strong>EXPERIENCER</strong></td>
<td>The experiencer of an event</td>
<td><em>John has a headache</em></td>
</tr>
<tr>
<td><strong>FORCE</strong></td>
<td>The non volitional causer of an event</td>
<td><em>The wind blows from the mall</em></td>
</tr>
<tr>
<td><strong>THEME</strong></td>
<td>The participant most directly affected by the event</td>
<td><em>Benjamin Franklin broke the ice</em></td>
</tr>
<tr>
<td><strong>RESULT</strong></td>
<td>The end product of an event</td>
<td><em>The French government has built a regulation-size baseball diamond</em></td>
</tr>
<tr>
<td><strong>CONTENT</strong></td>
<td>The proposition or content of a propositional event</td>
<td><em>Mona asked “you met Ann at the supermarket”?</em></td>
</tr>
<tr>
<td><strong>INSTRUMENT</strong></td>
<td>An instrument used in an event</td>
<td><em>He turned to poaching catfish, stunning them with a shocking device</em></td>
</tr>
<tr>
<td><strong>BENEFICIARY</strong></td>
<td>The beneficiary of an event</td>
<td><em>Ann makes hotel reservations for her boss</em></td>
</tr>
<tr>
<td><strong>SOURCE</strong></td>
<td>The origin of the object of a transfer event</td>
<td><em>I flew from Boston</em></td>
</tr>
<tr>
<td><strong>GOAL</strong></td>
<td>The destination of an object of a transfer event</td>
<td><em>I drove to Portland</em></td>
</tr>
</tbody>
</table>
Thematic roles

Thematic roles help us generalize over different surface realizations of predicate arguments

- John broke the window.
- John broke the window with a rock.
- The rock broke the door.
- The window broke.
- The window was broken by John.
Diathesis alternation

- **Thematic grid** (or case frame): the set of thematic role arguments taken by a verb

- Many verbs allow their thematic roles to be realized in various syntactic positions: *verb alternation* or *diathesis alternation*.

\[
\text{Doris gave the book to Cary.}
\]

AGENT     THEME     GOAL

\[
\text{Doris gave Cary the book.}
\]

AGENT     GOAL     THEME
Problems with thematic roles

- Very difficult to come up with a standard set of roles and to produce a formal definition of roles
- Alternative models of semantic roles: **generalized semantic roles** to abstract over specific thematic roles
  - Defined by a set of heuristic features (agent-like, patient-like meanings)
The proposition Bank (PropBank)

- Resource of sentences annotated with semantic roles
- Predicate-argument relations were added to the syntactic trees of the Penn Treebank.
- Each sense of each verb has a specific set of roles: \textit{Arg0}, \textit{Arg1}, \textit{Arg2} and so on.
  - \textit{Arg0}: PROTO-AGENT
  - \textit{Arg1}: PROTO-PATIENT
  - \textit{Arg2}: specific to verb sense

\textit{Increase.01} “go up incrementally”
- \textit{Arg0}: causer of increase
- \textit{Arg1}: thing increasing
- \textit{Arg2}: amount increased by
- \textit{Arg3}: start point
- \textit{Arg4}: end point

\begin{quote}
[\textit{Arg0 Big Fruit Co.}] increased [\textit{Arg1 the price of bananas}].
[\textit{Arg1 The price of bananas}] was increased again [\textit{Arg0 by Big Fruit Co.}]
\end{quote}
FrameNet (Fillmore et al. 2003)

- **What is FrameNet?**
  - A lexical resource for English
    - Based on *frame semantics*
    - And supported by *corpus evidence*

- **Aim:**
  - Documenting the range of *semantic and syntactic combinatory possibilities* of each word in each of its senses through annotation of *example sentences*
What is in a frame:

- **Semantic Frame**: conceptual structure that describes a specific type of *situation, object, event* and the *participants* involved in it.
  - Ex. REQUEST
  - Definition: *In this frame a Speaker asks an Addressee for something, or to carry out some action.*

- **Lexical Unit**: a word, a multiword or an idiomatic expression (typically a verb) that evokes a frame, also called *target*.
  - Ex. for REQUEST:
    - *ask, beg, command, demand, implore, order, petition, request, urge*

- **Frame Element**: roles, typically the syntactic dependents of the lexical unit.
  - Ex. for REQUEST:
    - Core: *Speaker, Addressee, Topic, Message, Medium*
    - Non-Core: *Beneficiary, Manner, Means, Time*

- **Relations** with other Frames: inheritance, part of
  - EX: REQUEST Uses: COMMUNICATION
Frame Example Annotation:

Annotate all chunks describing a participant in the frame

- *Frame Element Name* (manual): e.g. Speaker, Addressee
- *Phrase Type* (automatic): e.g. NP, PP, …
- *Grammatical Function* (automatic): e.g. External Argument (Subject), Object, …

[A supervisor] Speaker can REQUEST [a file]Message after reading the form.

In fact [I]Addressee was ASKED [to chair the meeting]Message.


Despite [their]Speaker PETITIONS [for mercy]Message, the three men were sentenced to death by firing squad.

...
The Berkeley FrameNet project:

- More than 10,000 LUs (more than 6,100 fully annotated). Frames come with definitions from the *Concise Oxford Dictionary* or written by FN staff member.
- More than 825 semantic frames.
- More than 135,000 annotated sentences. Roughly 20 example sentences per LU.
- The main FN corpus is the 100-million-word BNC + U.S. newswire texts provided by Linguistic Data Consortium.
Why is FrameNet useful?

- Provides a richer set of annotated semantic information than other existing frameworks (e.g. PropBank)
- Can be useful in different NLP tasks:
  - Textual Entailment:
    T: [The Everest summiter Hiddleston]\textit{PROTAGONIST} has passed away \textit{[in an avalanche of Mt. Tasman]}\textit{CAUSE} (frame: \textit{Death})
    H: [A person]\textit{PROTAGONIST} died \textit{[in an avalanche]}\textit{CAUSE} (frame: \textit{Death})
  - Question Answering:
    [Who]\textit{COGNIZER} discovered \textit{[prions]}\textit{NEW.IDEA}? (frame: \textit{Achieving\_first})
    [1997]\textit{PLACE: [Stanley B. Prusiner]}\textit{COGNIZER}, [United States]\textit{TIME, discovery} \textit{[of prions]}\textit{NEW.IDEA} (frame: \textit{Achieving\_first})
Selectional Restrictions

- Semantic roles gave us a way to express some of the semantics of an argument in its relation to the predicate

*Selectional restriction:* semantic type constraints that a verb imposes on the kind of concepts that are allowed to fill its arguments roles

- Selectional restrictions are associated with senses, not entire lexemes

They *serve* **green-lipped mussels** from New Zealand  
Which airlines *serve* **Denver**?  

**THEME:** foodstaff  
**THEME:** location
Representing selectional restrictions

$$\exists e, x, y \ \text{Eating}(e)^{\wedge} \ \text{Agent}(e,x)^{\wedge} \ \text{Theme}(e,y)$$

$$\exists e, x, y \ \text{Eating}(e)^{\wedge} \ \text{Agent}(e,x)^{\wedge} \ \text{Theme}(e,y)^{\wedge} \ \text{Isa}(y, \text{EdibleThing}) \ \text{SELECTIONAL RESTRICTION}$$

To represent a sentence like: I ate an apple.

$$\exists e, x, y \ \text{Eating}(e)^{\wedge} \ \text{Agent}(e,x)^{\wedge} \ \text{Theme}(e,y)^{\wedge} \ \text{Isa}(y, \text{EdibleThing})^{\wedge} \ \text{Isa}(y, \text{apple})$$

- Two practical problems:
  - the use of First Order Predicate Calculus to perform the simple task of enforcing selectional restriction is overkill
  - a large logical knowledge-base of facts about the concepts that make up the selectional restrictions is presupposed
Selectinal restrictions in terms of WordNet synsets

Each predicate specifies a WordNet synset as selectinal restriction of its arguments.

A meaning representation is well-formed if the role filler word is a hyponym of this synset.

"I ate an apple."

THEME role of the verb to eat: synset \{food, nutrient\}

- S: (n) apple (fruit with red or yellow or green skin and sweet to tart crisp whitish flesh)
  - direct hyponym / full hyponym
  - direct hyponym / inherited hyponym / sister term
    - S: (n) edible fruit (edible reproductive body of a seed plant especially one having sweet flesh)
      - S: (n) produce, green goods, green groceries, garden truck (fresh fruits and vegetables grown for the market)
    - S: (n) food, solid food (any solid substance (as opposed to liquid) that is used as a source of nourishment) "food and drink"
    - S: (n) solid (matter that is solid at room temperature and pressure)
      - 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 inferred to have its own distinct existence (living or nonliving))
Next Lab:

- Exercises of text annotation using WordNet synsets
- Semantic relatedness using lexical resources (WordNet similarity measures)
- Reading group: discussion of the paper
  - Ponzetto, S. P. Strube, M. *Knowledge Derived From Wikipedia For Computing Semantic Relatedness*
    

*Please, read it by next Wednesday!*
BIBLIOGRAPHY:


CREDITS: thanks to Carlo Strapparava and Sara Tonelli (FBK-Irst)