

# Lec8: Lexicon-II

HUL 242

## Minor Exam

- Exam is on 11/2 4:00 PM - 5:00 PM (Location LH 504)
- Time: 1 hr
- Make sure you can login to Moodle (no overquota)
- Be there 15 minutes prior to start time (i.e. 3:45 PM)
- Read the instructions written in the quiz!
- You will still be asked to write explanations- But keep them short!

## Minor Exam

- NOT an open book exam
- Items in the notes not covered in class will not be tested
- Concepts covered in tutorial problems are relevant!

## Recap: Lexical Network

- Lexicon: collection of vocabulary items, dictionary
- Structured lexicon: WordNet
- Hierarchical organization of concepts
- Using insights from psychology, spreading activation

## Challenges for WordNet

- How to map of word to meaning, when multiple meanings exist?
- Matrix of such mappings word  $\rightarrow$  meaning
- $F_1$  and  $F_2$  are synonyms i.e. share the same meaning
- Two entries in the same column ?

| Word Meanings | Word Forms |           |           |         |           |
|---------------|------------|-----------|-----------|---------|-----------|
|               | $F_1$      | $F_2$     | $F_3$     | $\dots$ | $F_n$     |
| $M_1$         | $E_{1,1}$  | $E_{1,2}$ |           |         |           |
| $M_2$         |            | $E_{2,2}$ |           |         |           |
| $M_3$         |            |           | $E_{3,3}$ |         |           |
| $\vdots$      |            |           |           | $\dots$ |           |
| $M_m$         |            |           |           |         | $E_{m,n}$ |

# Synonymy

- Are words that are polysemous likely to have more synonyms as well ? Examples?

## Synonymy

- Are words that are polysemous likely to have more synonyms as well ? Examples?
- At first, WordNet was only concerned with network relations between concepts (column-wise)

# Synonymy

- Are words that are polysemous likely to have more synonyms as well ? Examples?
- At first, WordNet was only concerned with network relations between concepts (column-wise)
- But ignoring relations between the words themselves not practical either



## Synsets

- Words are arranged in synonym sets **synsets**; a set of words interchangeable in some context

# Synsets

- Words are arranged in synonym sets **synsets**; a set of words interchangeable in some context
- Members of each synset → lemmas

## Synsets

- Words are arranged in synonym sets **synsets**; a set of words interchangeable in some context
- Members of each synset → lemmas
- Word meaning represented by listing all the words used to express it (row-wise) {board, plank}. But also {board, company}

## Synsets

- Words are arranged in synonym sets **synsets**; a set of words interchangeable in some context
- Members of each synset → lemmas
- Word meaning represented by listing all the words used to express it (row-wise) {board, plank}. But also {board, company}
- Which is the more frequent sense ?

## Example

- In WordNet, the most frequently occurring synset will show up first

### Noun

- (28)[S:](#) (n) **board** (a committee having supervisory powers) *"the board has seven members"*
- (18)[S:](#) (n) **board, plank** (a stout length of sawn timber; made in a wide variety of sizes and used for many purposes)
- (4)[S:](#) (n) **board** (a flat piece of material designed for a special purpose) *"he nailed boards across the windows"*

Is it too fine grained?

## Word sense disambiguation

- Selecting the correct sense of a word
- Take as input a word (+context) + WordNet senses
- Disambiguate *bank*

### Example

The **bank** can guarantee deposits will eventually cover future tuition costs because it invests in adjustable-rate mortgage securities.

- S: (n) 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 currents"*
- 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"*

## Overlap of words

- A simple approach to the problem:
- Remove function words/ stop words
- Choose the sense whose dictionary gloss or definition shares the most words with the target words context
- The sense *bank* (financial institution) has some overlapping words with the target sentence
- Problems: glosses, definitions too short
- Collect more senses from hand-labelled corpus data

## Synsets: basis of lexical relations

- WordNet synsets are used to organize the semantic network
- Synsets are nodes in the lexical relations in WordNet

Table : Noun relations

| Relation   | Aka           | Example             |
|------------|---------------|---------------------|
| Hypernym ✓ | Superordinate | breakfast → meal    |
| Hyponym ✓  | Subordinate   | meal → lunch        |
| Meronym    | Has-member    | faculty → professor |
| Holonym    | Member-of     | copilot → crew      |



## Example

- **S: (n) 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 currents"*
  - direct hyponym / full hyponym
    - **S: (n) riverbank, riverside** (the bank of a river)
    - **S: (n) waterside** (land bordering a body of water)
  - direct hypernym / **inherited hypernym** / 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 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))

## Adjectives

- Difficult to organize into a hierarchy
- Attributes are usually bipolar (with opposing values)
- Basic semantic relation is antonymy
- Applies mainly to descriptive adjectives *good, big, interesting*
- What are exceptions to this rule?

## Adjectives

- Difficult to organize into a hierarchy
- Attributes are usually bipolar (with opposing values)
- Basic semantic relation is antonymy
- Applies mainly to descriptive adjectives *good, big, interesting*
- What are exceptions to this rule?
- Colours *green* or relational adjectives *musical, presidential* (often similar to nouns)

# Organization

- Division of lexicon into the following categories
- Nouns → topical hierarchy (82115 synsets)
- Verbs → entailment relations (13767 synsets)
- Adjectives, Adverbs → no hierarchical arrangement (21777 synsets)

# Organization

- Division of lexicon into the following categories
- Nouns → topical hierarchy (82115 synsets)
- Verbs → entailment relations (13767 synsets)
- Adjectives, Adverbs → no hierarchical arrangement (21777 synsets)

# Organization

- Division of lexicon into the following categories
- Nouns → topical hierarchy (82115 synsets)
- Verbs → entailment relations (13767 synsets)
- Adjectives, Adverbs → no hierarchical arrangement (21777 synsets)

# Organization

- Division of lexicon into the following categories
- Nouns → topical hierarchy (82115 synsets)
- Verbs → entailment relations (13767 synsets)
- Adjectives, Adverbs → no hierarchical arrangement (21777 synsets)