

Natural Language Processing

Lecture 9: Lexical Semantics -
Word Senses and Lexical Relations

11/23/2020

COMS W4705
Yassine Benajiba

Natural Language Semantics

- Semantics is concerned with the meaning of language.
- Lexical Semantics: What is the meaning of individual words?
- Computational Semantics: How do we compute language meaning from word meaning? (next week)
- How do we represent meaning?

Two Approaches to Language Meaning

- *Usage-based semantics:*
 - Distributional / Vector-space semantics. Word embeddings.
 - Core concept: Semantic similarity.
 - History in connectionist approaches (neural networks).
- Formal Semantics:
 - Use formal Meaning Representations for words /sentences / discourse.
 - Based on lexical resources.
 - History in cognitive science (mind as symbol manipulation device).

Why Natural Language Semantics?

- Meaning representations bridge between linguistic input and extra-linguistic knowledge.
 - Support inference and reasoning.
 - Examples:
 - Answering questions on an exam.
 - Deciding what to order at a restaurant by reading the menu.
 - Learning how to use a device by reading the manual.
 - Finding a joke funny.
 - Following a recipe.

Syntax and Semantics

- Can we do syntax without semantics?
 - *"He read the book with the blue cover"*
 - *"He saw the hill with the telescope"*
 - *"Last night I shot an elephant in my pajamas"*
- These examples require world knowledge (books have covers, telescopes are used for seeing, elephants are large, ...)

Semantics and Machine Translation

L'Avocat général_{fr} \longrightarrow the general avocado_{en}
??

Word Senses

- Different word senses of the same word can denote different (more or less related) concepts.

bank (of a river) vs. *bank* (financial institution) vs. *bank* (storage facility)

mouse (animal) vs. *mouse* (computer accessory)

bright [light] vs. *bright* [idea] vs. *bright* [student] vs. *bright* [future]

- "Lexeme" - a pairing of a particular word form with its sense.

Homonymy

- Homonymy is a relation between concepts/senses:
 - Multiple **unrelated** concepts correspond to the same word form.
 - Examples:
 - *bank*
 - *check*
 - *kind*
 - *bass*

Polysemy

- Multiple semantically related concepts correspond to the same word form.
- Examples:
 - *wood* (material that trees are made of) vs. *wood* (a forested area)
 - *bank* (financial institution) vs. *bank* (building)

Metonymy

- A subtype of polysemy.
- Systematic and productive.
- One aspect of a concept is used to refer to other aspects of a concept (or the concept itself).
- BUILDING <-> ORGANIZATION (*bank, school,...*)
- ANIMAL <-> MEAT (the *chicken* was overcooked,
the *chicken* eats a worm)

Zeugma

- when a single word is used with two other parts of a sentence but must be understood differently (word sense) in relation to each.
- *Does United serve breakfast and JFK?*
- *He lost his gloves and his temper.*

Semantic Relations

- Synonym / Antonym
- Hypernym / Hyponym (IS-A)
- Meronym / Holonym (part-of relationship)

Synonyms

- Two lexemes that share a sense.
 - couch/sofa
 - vomit/throw up
 - car/automobile
 - hazelnut/filbert
 - water/H₂O

Note that even though the sense is the same, there may be differences in politeness, slang, register, genre, etc.

Lexical Substitution

- Two lexemes are synonyms if they can be substituted for each other in a sentence, such that the sentence retains its meaning (truth conditions).
- Note that synonymy is not a relationship between words, but between lexemes.

large
How big is that plane?

large (?)
She was like a big sister to him.

Antonyms

- Senses are opposites with respect to one specific feature of their meaning.
- Otherwise, they are very similar!
 - *dark / light* (level of luminosity)
 - *short / long* (length)
 - *hot / cold* (temperature)
 - *rise / fall* (direction)
 - *front / back* (relative position)

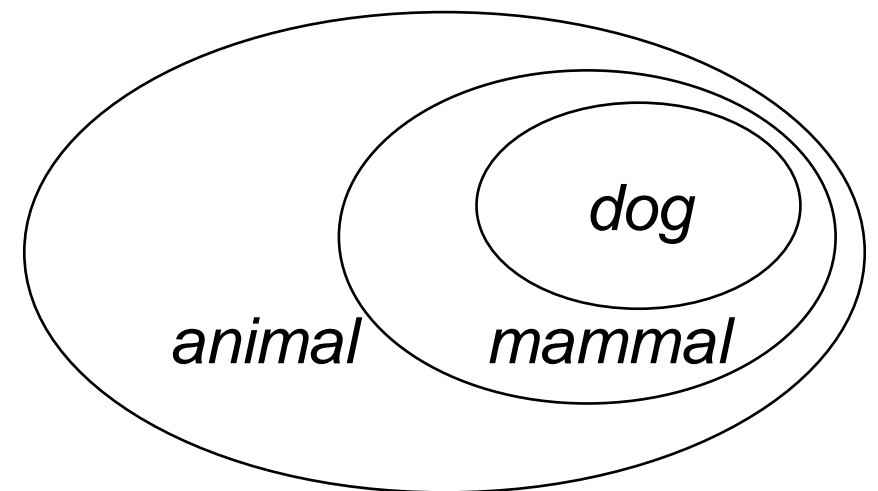
Antonyms typically describe opposite ends of a scale, or opposite direction/position with respect to some landmark.

Hyponymy

- One sense is a hyponym (or subordinate) of another sense if the first sense is more specific, denoting a subclass of the other. (IS-A relationship).
 - *dog* is a hyponym of *mammal*.
 - *mammal* is a hyponym of *animal*.
 - *desk* is a hyponym of *furniture*.
 - *sprint* is a hyponym of *run*.
- The inverse relation is called **hypernymy**, so *furniture* is a hypernym (or superordinate) of *desk*.

Hyponymy

- Can think of Hyponymy as a set relationship.
- *for all x, if x is a mammal, then x is an animal.*
- Related to **entailment**:



The woman ate spaghetti with cheese.



The person consumed pasta with dairy product.

Meronymy

- Part-whole relationship.
- A meronym is a part of another concept.
 - *leg* is a meronym of *chair*.
 - *wheel* is a meronym of *car*.
 - *cellulose* is a meronym of *paper*. (substance meronymy)
- The inverse relation is **holonymy**.
Car is a holonym of *wheel*.

WordNet

- WordNet is a lexical database containing English word senses and their relations.
- Represents word sense as **synsets**, sets of lemmas that have synonymous lexemes in one context.

{composition, paper, report, theme}

{newspaper, paper}

{paper}

- Version 3.1 Contains synonyms, antonyms, hypernyms, (some) meronyms, and frequency information for about 117.000 nouns, 11.500 verbs, 22.000 adjectives, and 4.500 adverbs.

WordNet Senses

- each sense comes with a *gloss* (dictionary definition).

```
$ wn mouse -over
```

Overview of noun mouse

The noun mouse has 4 senses (first 1 from tagged texts)

1. (14) mouse -- (any of numerous small rodents typically resembling diminutive rats having pointed snouts and small ears on elongated bodies with slender usually hairless tails)
2. shiner, black eye, mouse -- (a swollen bruise caused by a blow to the eye)
3. mouse -- (person who is quiet or timid)
4. mouse, computer mouse -- (a hand-operated electronic device that controls the coordinates of a cursor on your computer screen as you move it around on a pad; on the bottom of the device is a ball that rolls on the surface of the pad; "a mouse takes much more room than a trackball")

WordNet Hypernyms

```
$ wn mouse -hyphen
```

```
Synonyms/Hypernyms (Ordered by Estimated Frequency) of noun mouse
```

```
4 senses of mouse
```

```
Sense 1
```

```
mouse
```

```
=> rodent, gnawer
```

```
=> placental, placental mammal, eutherian, eutherian mammal
```

```
=> mammal, mammalian
```

```
=> vertebrate, craniate
```

```
=> chordate
```

```
=> animal, animate being, beast, brute, creature, fauna
```

```
=> organism, being
```

```
=> living thing, animate thing
```

```
=> whole, unit
```

```
=> object, physical object
```

```
=> physical entity
```

```
=> entity
```

WordNet Meronyms

```
$ wn mouse -meron
```

```
Meronyms of noun mouse
```

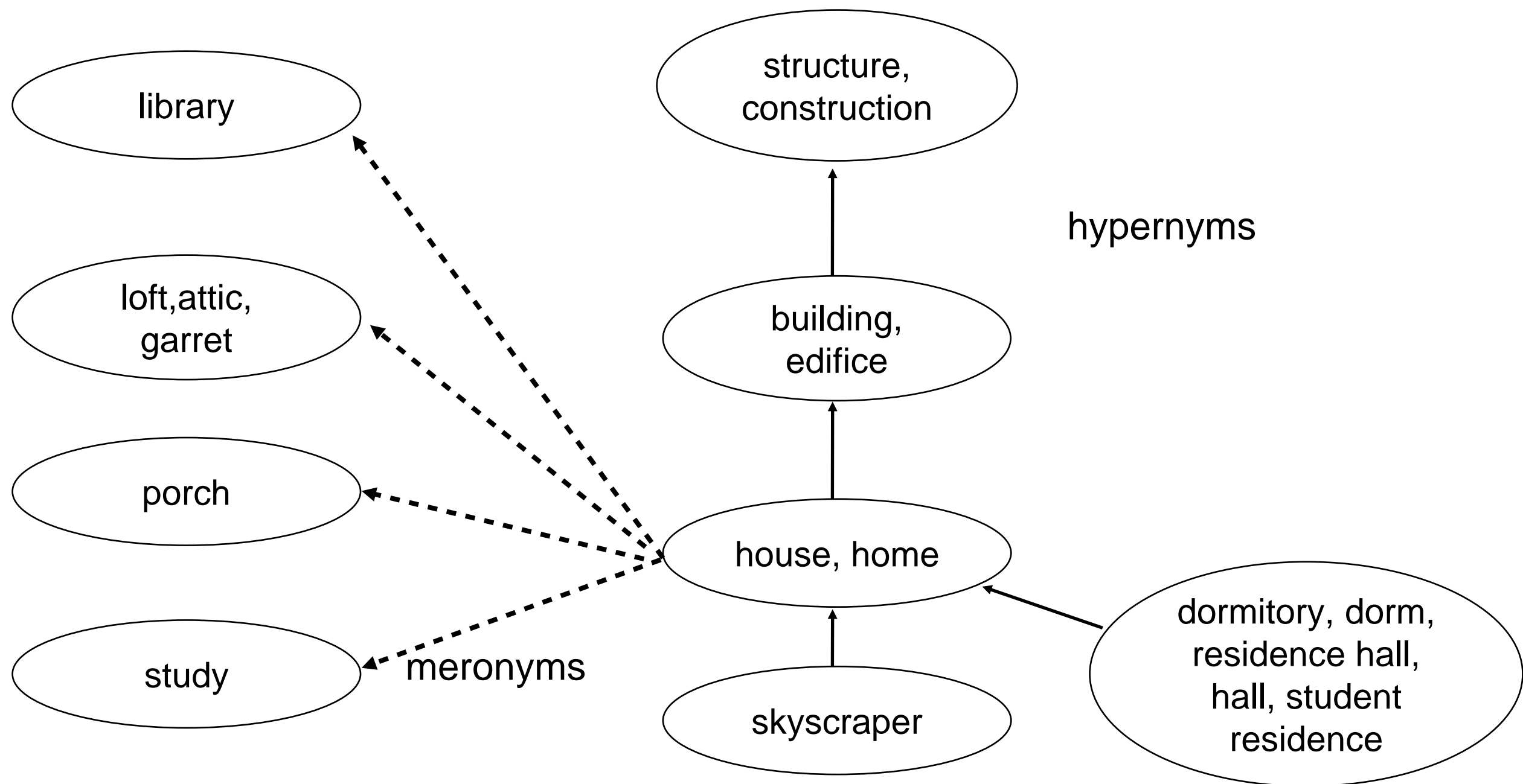
```
1 of 4 senses of mouse
```

```
Sense 4
```

```
mouse, computer mouse
```

```
    HAS PART: mouse button
```

Semantic Relations as a Graph



Word Sense Disambiguation (WSD)

- Given a word token in context, identify its correct word sense (from a list of possible word-senses).
- Why? Machine translation. Question answering. Information Retrieval, Speech Synthesis...
- What set of senses:
 - For MT: Possible translations of a word (e.g. English to Spanish)
 - For Speech Synthesis: Homographs (*bass*, *bread*,...)
 - Senses from a dictionary like WordNet

Two types of WSD tasks

- Lexical Sample task
 - Small pre-selected set of target words.
 - Inventory of senses for each word.

*Largemouth **bass** are willing to bite a variety of baits.*

- All-words task
 - Every word in an entire text.
 - A lexicon with senses for each word.
 - Similar to POS tagging (but specific set of tags for each word).

WSD Methods

- Supervised Machine Learning
 - Train a classifier for each word.
 - Requires hand-labeled training data (sense annotations).
- Dictionary Methods
 - No training data. Instead use a sense dictionary (like WordNet).
 - Or exploit sense relations (WordNet graph).
- Semi-supervised learning
 - Use a small hand-annotated data set and generalize ("bootstrapping").

Supervised WSD

- Given:
 - A sense inventory (for example, WordNet senses).
 - An annotated training corpus.
 - A set of features extracted from the training corpus.
 - A classifier.
- How would you solve this?

Annotated Training Corpus

- Lexical sample task:
 - SENSEVAL 1 to 3 (1998,2001). Small number of target words (~50), 12.000 to 20.000 instances.
SemEval 2007.
 - "line, hard, serve" dataset. 4.000 instances each for the words line, hard, and serve.
- All-words task:
 - SENSEVAL 3, SemEval 2007.
 - SemCor ("Semantic concordance"). 234,000 words from Brown Corpus, manually tagged with WordNet senses.

Feature Definition - Intuition

"If one examines the words in a book, one at a time as through an opaque mask with a hole in it one word wide, then it is obviously impossible to determine, one at a time, the meaning of the words...

But if one lengthens the slit in the opaque mask, until one can see not only the central word in question but also say N words on either side, then if N is large enough one can unambiguously decide the meaning of the central word...

The practical question is : ``What minimum value of N will, at least in a tolerable fraction of cases, lead to the correct choice of meaning for the central word?"

Feature Definition

- Can you identify the word sense?

(: bass)

Feature Definition

- Approaches for representing context:
 - Collocational features:
 - Take the position of each context word into account.
 - Bag-of-word features:
 - Simply keep the set of all context words.
- Using word embeddings (or phrase embeddings)

Bag-of-Words features

*He learned how to play **guitar and bass** watching Youtube videos.*



fish	0
⋮	⋮
guitar	1
⋮	⋮
watching	1
⋮	⋮
Youtube	1

Collocational Features

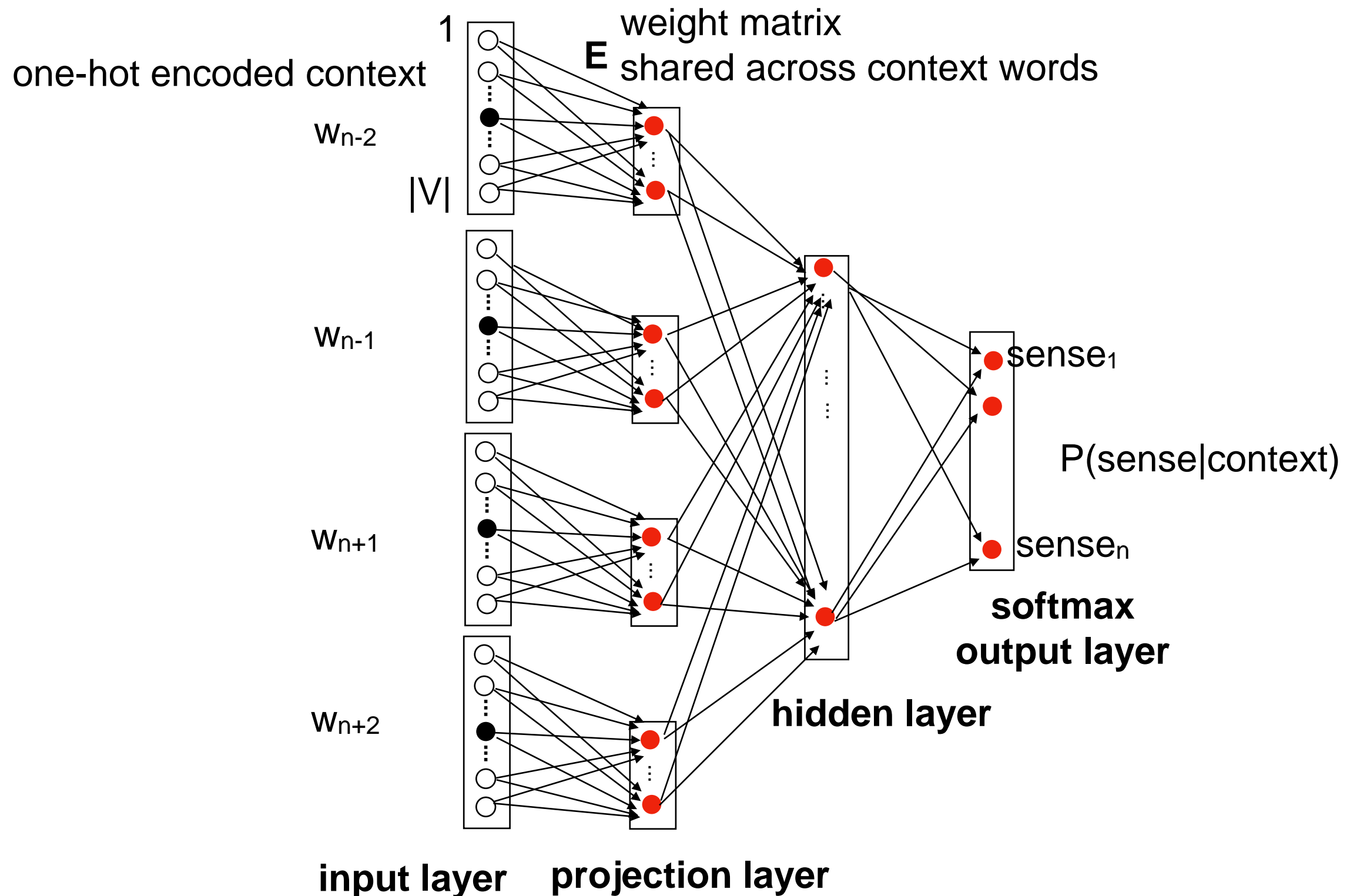
*He learned how to play **guitar and bass** watching Youtube videos.*

- Assume feature window: +/- 2 words.
- Position specific information about the words in the window

[guitar, NN, and, CC, watching, VBG, Youtube, NNP]

[word_{n-2}, POS_{n-2}, word_{n-1}, POS_{n-1}, word_{n+1}, POS_{n+1}]

Feed-forward NN for WSD



Dictionary-Based Methods

- Supervised WSD requires a lot of annotated training data.
- Instead, we can use a separate dictionary, such as WordNet:
 - to obtain candidate senses.
 - to obtain information that allows us to identify which of the candidate senses is correct.

1. (25) **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"

2. (20) **depository financial institution, bank, banking company**

a financial institution that accepts deposits and channels the money into

"he cashed a check at the bank"

"that bank holds the mortgage on my home"

Simplified Lesk Algorithm

Lesk 1986

- Use dictionary glosses for each sense.
- Choose the sense that has the highest word overlap between gloss and context (ignore function words).

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

1. (25) **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"

2. (20) **depository financial institution, bank, banking company**

a financial institution that accepts deposits and channels the money in

"he cashed a check at the bank"

"that bank holds the mortgage on my home"

Extensions to Lesk Algorithm

- Often the available definitions and examples do not provide enough information. Overlap is 0.
- Different approaches to extending definitions:
 - "Corpus-Lesk": Use a sense-tagged-example corpus, add context from example sentences.
 - Extended Gloss Overlap (Banerjee & Pedersen 2003):
Add glosses from related words (hypernyms, meronyms, ...)
- Use embedded representations for each word in the definition. Choose the sense with highest average vector similarity to the context.

WSD evaluation

- Ideally, we would like to use an **extrinsic** (task-based) evaluation.
 - Use WSD as a component of some task (for example, MT) and see if the results improve.
- For convenience we often use **intrinsic** evaluation.
 - Measure sense accuracy (% of correct sense tags).
- What are some good baselines for this task?
 - Most frequent sense (according to some tagged corpus).
 - Lesk often used as a baseline for more elaborate approaches.

WSD Performance

- Varies widely depending on how difficult the disambiguation task is.
- Accuracies of over 90% are commonly reported on some of the classical WSD tasks.
- Senseval brought careful evaluation of difficult WSD (many senses, complex relations).
- Senseval 1: more fine grained senses, wider range of types:
 - Overall: about 75% accuracy
 - Nouns: about 80% accuracy
 - Verbs: about 70% accuracy

Upper Bound

- How well do people do on this task?
- Inter-annotator agreement:
 - Compare multiple human annotations on the same data, given the same annotations guidelines.
- Human agreement on all-words task with WordNet senses:
 - 75%-80%

Semi-Supervised WSD

- What if we only have a few labeled training examples.
- Idea: Bootstrapping. Generalize from a small hand-labeled dataset.
Yarowsky, 1995
- "One sense per collocation" rule
 - A word reoccurring in collocation with the same word will almost surely have the same sense.
 - Example:
 - the word *play* occurs with the music sense of **bass**
 - the word *fish* occurs with the fish sense of **bass**

Extracting New Sentences

- Using a few keywords and the "One sense per collocation" rule we can label new sentences for the known senses.

We need more good teachers – right now, there are only a half a dozen who can play the free **bass** with ease.

An electric guitar and **bass** player stand off to one side, not really part of the scene, just as a sort of nod to gringo expectations perhaps

The researchers said the worms spend part of their life cycle in such **fish** as Pacific salmon and striped **bass** and Pacific rockfish or snapper.

And it all started when **fishermen** decided the striped **bass** in Lake Mead were too skinny.

A Lexical Substitution Task

- Instead of identifying word sense
 - find a substitute for the target word, such that the meaning is preserved.

*He was **bright** and independent and proud .*

*Snow covered areas appear **bright** blue in the image which was taken in early spring.*

*... an institution that nurtures the best and **brightest** young musicians.*

- How would you solve this task?

Acknowledgments

- Some slides by Kathy McKeown, Bob Coyne, Dan Jurafsky.