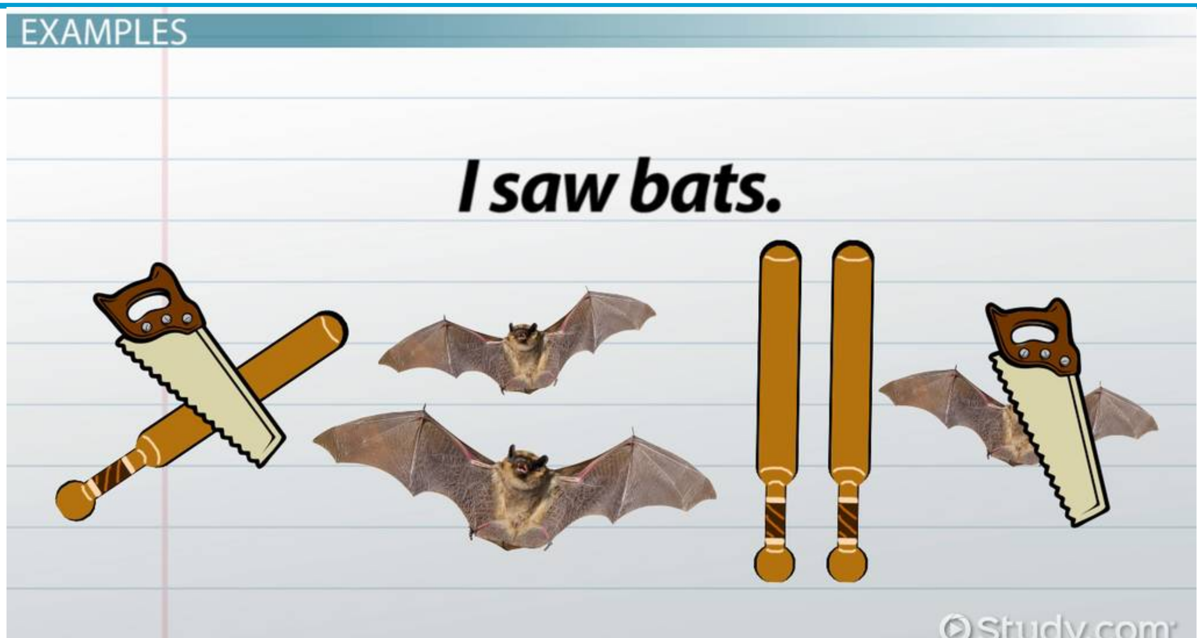
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Jurafsky, D. and Martin, J. H. (2009): Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics and Speech Recognition. Second Edition. Pearson: New Jersey: Chapter 20
- Agirre, E., Edmonds, P. (2006): Word Sense Disambiguation: Algorithms and Applications (Text, Speech and Language Technology). Springer, Heidelberg

knowledge-based, supervised, word sense induction, substitution

WORD SENSE AND MEANING

MOTIVATING EXAMPLE

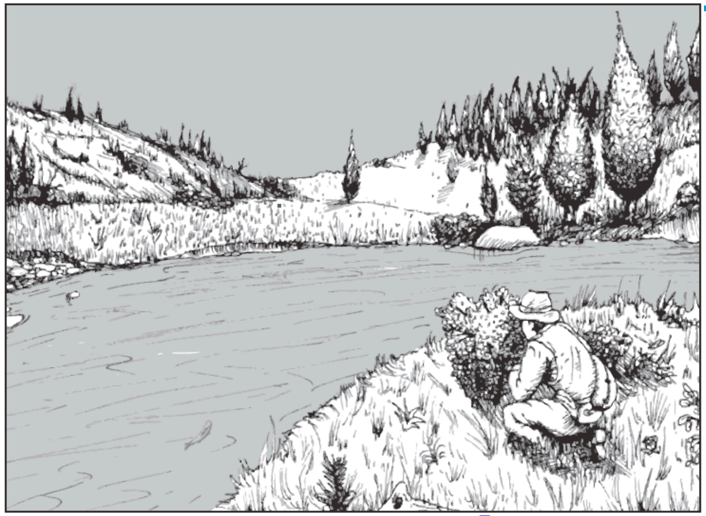


http://study.com/academy/lesson/lexical-ambiguity-definition-examples.html

- ambiguous words have multiple senses
- Word Sense Disambiguation assigns a single (out of multiple) sense in a given context
- Semantic disambiguities: assume that syntactic disambiguation has already been performed

WHY LANGUAGE IS HARD ..

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG



polysemous

synonymous

Concept Layer

Lexical Layer

He sat on the river bank and counted his dough.

She went to the bank and took out some money.

TYPES OF AMBIGUITY

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Homonymy: two or more meanings happen to be expressed with the same string
  - withdrawing money from the **bank**
  - embark on a boat from the river **bank**

- Polysemy: the same string has different, but related senses, stemming from the same origin
  - the **bank** was robbed by Billy the Kid
  - the **bank** was constructed by a famous architect

APPROACHES TO WSD

- Knowledge Based Approaches ('unsupervised')
  - Rely on knowledge resources like WordNet, Thesaurus etc.
  - May use grammar rules for disambiguation.
  - May use hand coded rules for disambiguation.
- Machine Learning Based Approaches ('supervised')
  - Rely on corpus evidence.
  - Train a model using tagged or untagged corpus.
  - Probabilistic/Statistical models.
- Hybrid Approaches
  - Use corpus evidence as well as semantic relations from WordNet.
- Unsupervised, knowledge-free Approaches
  - induce sense inventory
  - disambiguate in context

WSD USING SELECTIONAL PREFERENCES AND ARGUMENTS

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

Sense 1

This airline ***serves*** dinner in the evening flight.

- serve (Verb)
  - agent
  - object – edible

Sense 2

This airline ***serves*** the sector between Munich and Rome.

- serve (Verb)
  - agent
  - object – sector

Requires exhaustive enumeration of:

- Argument-structure of verbs.
- Selectional preferences of arguments.
- Description of properties of words such that meeting the selectional preference criteria can be decided.

  E.g. This flight serves the “region” between Paris and Warsaw.

  How do you decide if “region” is compatible with “sector”

OVERLAP-BASED APPROACHES

- Requires a Machine Readable Dictionary (MRD).
- Find the overlap between the features of different senses of an ambiguous word (sense bag) and the features of the words in its context.
- These features could be sense definitions, example sentences, etc.
- The sense which has the maximum overlap is selected as the contextually appropriate sense.

Concise Oxford English Dictionary

- **dance**
  - verb
    1. move rhythmically to music, typically following a set sequence of steps. ▶ perform (a particular dance or a role in a ballet).
    2. move in a quick and light or lively way. ▶ (of someone's eyes) sparkle with pleasure or excitement.
  - noun
    1. a series of steps and movements that match the speed and rhythm of a piece of music. ▶ an act or period of dancing.
    2. a social gathering at which people dance.
    3. (also **dance music**) music for dancing to, especially in a nightclub.

LESK (1986) ALGORITHM

- Identify senses of words in context using definition overlap
- Can do this either between gloss and context or between all sense combinations
- Main problem: zero overlap for most contexts

```
function SimplifiedLesk(word, sentence) {
 bestSense= mostFrequentSense(word);
 maxOverlap =0;
 context = allWords(sentence);

 foreach sense in allSenses(word) {
    signature=signature(sense);
    overlap = overlap(signature, context);
    if (overlap > maxOverlap) {
          maxOverlap = overlap;
          bestSense - sense
 }}
 return bestSense;
}
```

Dictionary functions

- `mostFrequentSense`: returns most frequent / first sense identifier from dictionary
- `allSenses`: returns all sense identifiers for a word from dictionary
- `signature`: returns set of words from sense definition in dictionary

Lesk, M. (1986). Automatic sense disambiguation using machine readable dictionaries: how to tell a pine cone from an ice cream cone. In SIGDOC '86: Proceedings of the 5th annual international conference on Systems documentation, pages 24-26, New York, NY, USA. ACM.

SIMPLIFIED LESK ALGORITM

- Lesk algorithm relies on definitions of context words to disambiguate the senses of a target word
- Simplified Lesk:
  – Measure the overlap between the (sentence) context of the target, and the definition of its senses
  – If no overlap, use most frequent sense (MFS)

■ noun
1 ▶ a series of steps and movements that match the speed and rhythm of a piece of music. ▶ an act or period of dancing.
2 ▶ a social gathering at which people dance.
3 ▶ (also **dance music**) music for dancing to, especially in a nightclub.

1, 2 or 3 ?

"With the music, the **dance** started with slow movements."

WORDNET – AN ONLINE LEXICAL DATABASE

HTTP://WORDNET.PRINCETON.EDU/ - CURRENT VERSION: 3.1

- High coverage lexical-semantic network built by psychologists

POS	Monosemous Words and Senses	Polysemous Words	Polysemous Senses
Noun	101863	15935	44449
Verb	6277	5252	18770
Adjective	16503	4976	14399
Adverb	3748	733	1832
Totals	**128391**	**26896**	**79450**

- Relations:
  - ISA-relation (hyponom - hypernym, taxonomic backbone)
  - Part-of (meronym - holonym)
  - Type-instance (e.g. Obama is an instance of President)
  - Opposite-of (antonym), mostly for adjectives
  - Derivative (pertainym), e.g. crime – criminal
  - some semantic roles between verbs and nouns, e.g. AGENT, INSTRUMENT …

SYNSETS FOR "MAGAZINE#N"

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

WordNet Search - 3.0 - WordNet home page - Glossary - Help

Word to search for: magazine [Search WordNet]

Display Options: (Select option to change) [Change]

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

Noun

Synset

sample use

- (13)S: (n) **magazine#1**, mag#1 (a periodic publication containing pictures and stories and articles of interest to those who purchase it or subscribe to it) *"it takes several years before a magazine starts to break even or make money"*
- (2)S: (n) **magazine#2** (product consisting of a paperback periodic publication as a physical object) *"tripped over a pile of magazines"*
- (1)S: (n) **magazine#3**, magazine publisher#1 (a business firm that publishes magazines) *"he wo*

  gloss

  for a magazine"
- S: (n) **magazine#4**, cartridge#2 (a light-tight supply chamber holding the film and supplying it for exposure as required)
- S: (n) **magazine#5**, powder store#1, powder magazine#1 (a storehouse (as a compartment on a warship) where weapons and ammunition are stored)
- S: (n) cartridge holder#1, cartridge clip#1, clip#1, **magazine#6** (a metal frame or container holding cartridges; can be inserted into an automatic gun)

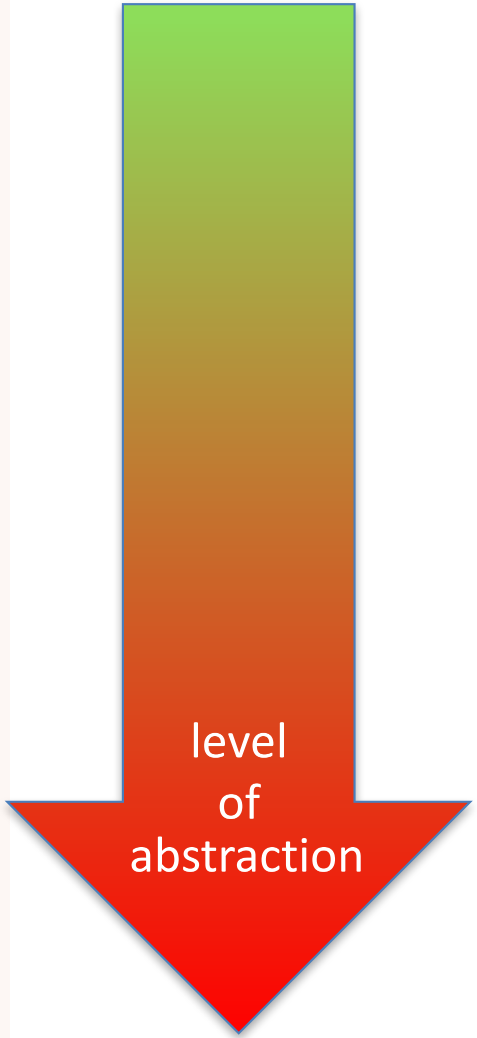
Lexical members

SemCor count

WORDNET HYPERNYM CHAIN

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- (2)S: (n) **magazine#2** (product consisting of a paperback periodic publication as a physical object) *"tripped over a pile of magazines"*
  - *direct hypernym* | *inherited hypernym* | *sister term*
    - S: (n) product#2, production#3 (an artifact that has been created by someone or some process) *"they improve their product every year"; "they export most of their agricultural production"*
      - S: (n) creation#2 (an artifact that has been brought into existence by someone)
        - S: (n) artifact#1, artefact#1 (a man-made object taken as a whole)
          - S: (n) whole#2, unit#6 (an assemblage of parts that is regarded as a single entity) *"how big is that part compared to the whole?"; "the team is a unit"*
            - S: (n) object#1, physical object#1 (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#1 (an entity that has physical existence)
                - S: (n) entity#1 (that which is perceived or known or inferred to have its own distinct existence (living or nonliving))



GLOSS OVERLAPS ≈ RELATEDNESS

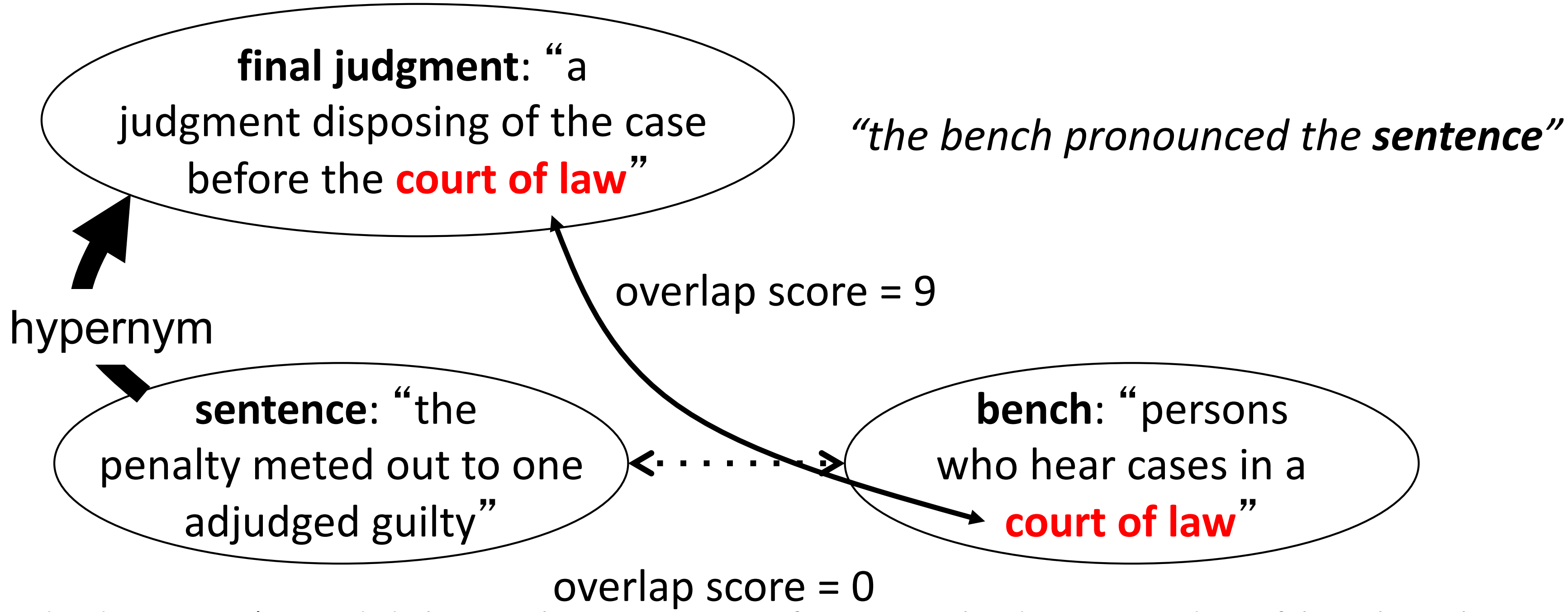
Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Lesk's (1986) idea: Related word senses are (often) defined *using the same words.* E.g:
  - bank(1): "a financial institution"
  - bank(2): "sloping <u>land</u> beside a <u>body</u> of <u>water</u>"
  - lake: "a <u>body</u> of <u>water</u> surrounded by <u>land</u>"
- Gloss overlaps = # content words common to two glosses ≈ relatedness
  - relatedness (bank(2), lake) = 3
  - relatedness (bank(1), lake) = 0

Problem: "lexical gap": Same or similar meaning can be expressed with a large variety of words from the vocabulary. For most pairs of glosses, overlap = 0 .

EXTENDED LESK (BANERJEE AND PEDERSEN, 2002)

- Utilize link structure of WordNet to pull in related glosses for overlap computation
- Addresses the overlap sparseness issue
- do this for one ambiguous word at-a-time
- Reweighting: For n-gram overlaps, add a score of n^2



(Banerjee and Pedersen, 2002). Extended Gloss Overlaps as a Measure of Semantic Relatedness. Proceedings of the Eighteenth International Joint Conference on Artificial Intelligence, pp. 805-810, August 9-15, 2003, Acapulco, Mexico.

ONE-SENSE-PER-DISCOURSE HYPOTHESIS

Universität Hamburg
DER FORSCHUNG | DER LEHRE | DER BILDUNG

- Idea: When a word is used several times in a document, it probably has the same meaning in all occurrences.
- This means that we can gather evidence from all contexts per document, which reduces sparseness
- holds for homonymous, not polysemous nouns, does not hold for verbs and adjectives as much
- Measuring the validity of the hypothesis (Small study on 12 nouns)
- Applicability: How often do we in fact observe an ambiguous word more than once in a document?
- Accuracy: If we observe an ambiguous word more than once per document, how often do these occurrences have the same meaning?

The one-sense-per-discourse hypothesis:

Word	Senses	Accuracy	Applicblty
plant	living/factory	99.8 %	72.8 %
tank	vehicle/contnr	99.6 %	50.5 %
poach	steal/boil	100.0 %	44.4 %
palm	tree/hand	99.8 %	38.5 %
axes	grid/tools	100.0 %	35.5 %
sake	benefit/drink	100.0 %	33.7 %
bass	fish/music	100.0 %	58.8 %
space	volume/outer	99.2 %	67.7 %
motion	legal/physical	99.9 %	49.8 %
crane	bird/machine	100.0 %	49.1 %
Average		99.8 %	50.1 %

Gale William, Kenneth Church, and David Yarowsky, "One Sense Per Discourse", in Proceedings of the ARPA Workshop on Speech and Natural Language Processing, pp. 233– 237, 1992.