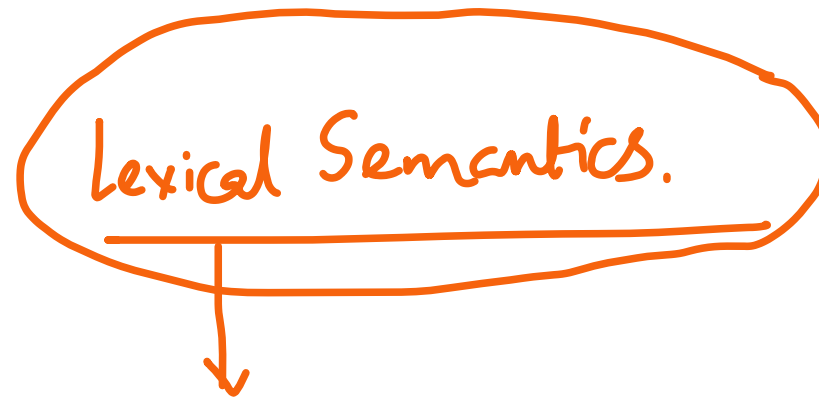
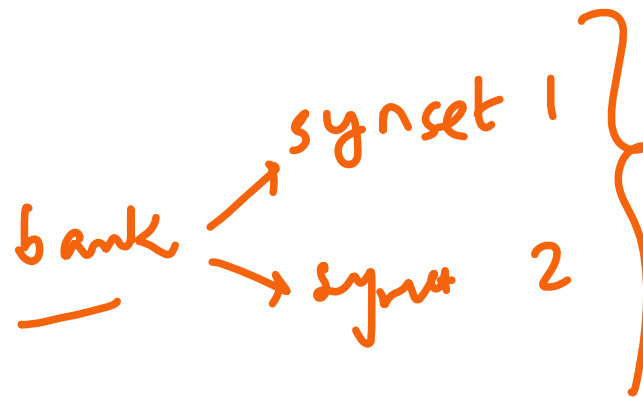


Lexical Semantics.



Word Sense Disambiguation (WSD)



References

- ACM Survey paper by Roberto Navigli
 - ACM Computing Surveys, Vol. 41, No. 2, Article 10, Publication date: February 2009.
- Statistical Language Learning by Eugene Charniak }

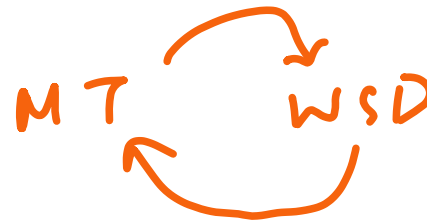
WSD

1. Challenges:

1. Different formalizations
 1. Representation of senses
 2. Granularity of senses
 3. Domain-oriented vs Unrestricted
 4. Set of words to disambiguate
2. Knowledge Acquisition Bottleneck

2. Applications:

1. Web
2. Machine Translation]



Refer: Section 1.1

Coarse and fine-grained sense distinction

- Coarse (homonyms):
 - The **bank** cashed my check.
 - The sun climbed slowly above the far **bank**, melting the mist, driving away all haze.
- Fine:
 - She chopped the vegetables with a chef's **knife**.
 - A man was beaten and cut with a **knife**.

Sense Enumeration versus Generation

- Enumerative Approach: more common
- Generative Approach:
 - Four qualia roles are provided, namely:
 - formal (a superordinate of knife)
 - constitutive (parts of a knife)
 - telic (the purpose of a knife)
 - agentive(who uses a knife).
 - The instantiation of a combination of roles allows for the creation

Refer: Section 2.1

Knowledge Sources for WSD

- Structured Resources : Thesauri, Machine Readable Dictionaries, Ontologies
- Unstructured Corpora :
 - Corpora
 - Raw
 - Sense-Annotated. Example: Semcor
 - Collocation resources
 - Other resources

Refer: Section 2.2

SemCor

As of Sunday_n¹ night_n¹ there was_v⁴ no word_n² of a resolution_n¹ being offered_v² there_r¹ to rescind_v¹ the action_n¹. Pelham pointed out_v¹ that Georgia_n¹ voters_n¹ last_r¹ November_n¹ rejected_v² a constitutional_a¹ amendment_n¹ to allow_v² legislators_n¹ to vote_n¹ on pay_n¹ raises_n¹ for future_a¹ Legislature_n¹ sessions_n².

WordNet senses

$Senses_{WN}(car_n) = \{ \text{car}_n^1, \text{auto}_n^1, \text{automobile}_n^1, \text{machine}_n^4, \text{motorcar}_n^1 \},$
 $\{ \text{car}_n^2, \text{rail car}_n^1, \text{rail way car}_n^1, \text{rail road car}_n^1 \},$
 $\{ \text{cable car}_n^1, \text{car}_n^3 \},$
 $\{ \text{car}_n^4, \text{gondola}_n^3 \},$
 $\{ \text{car}_n^5, \text{elevator car}_n^1 \} \}.$

Refer: Section 2.2.2

WordNet senses

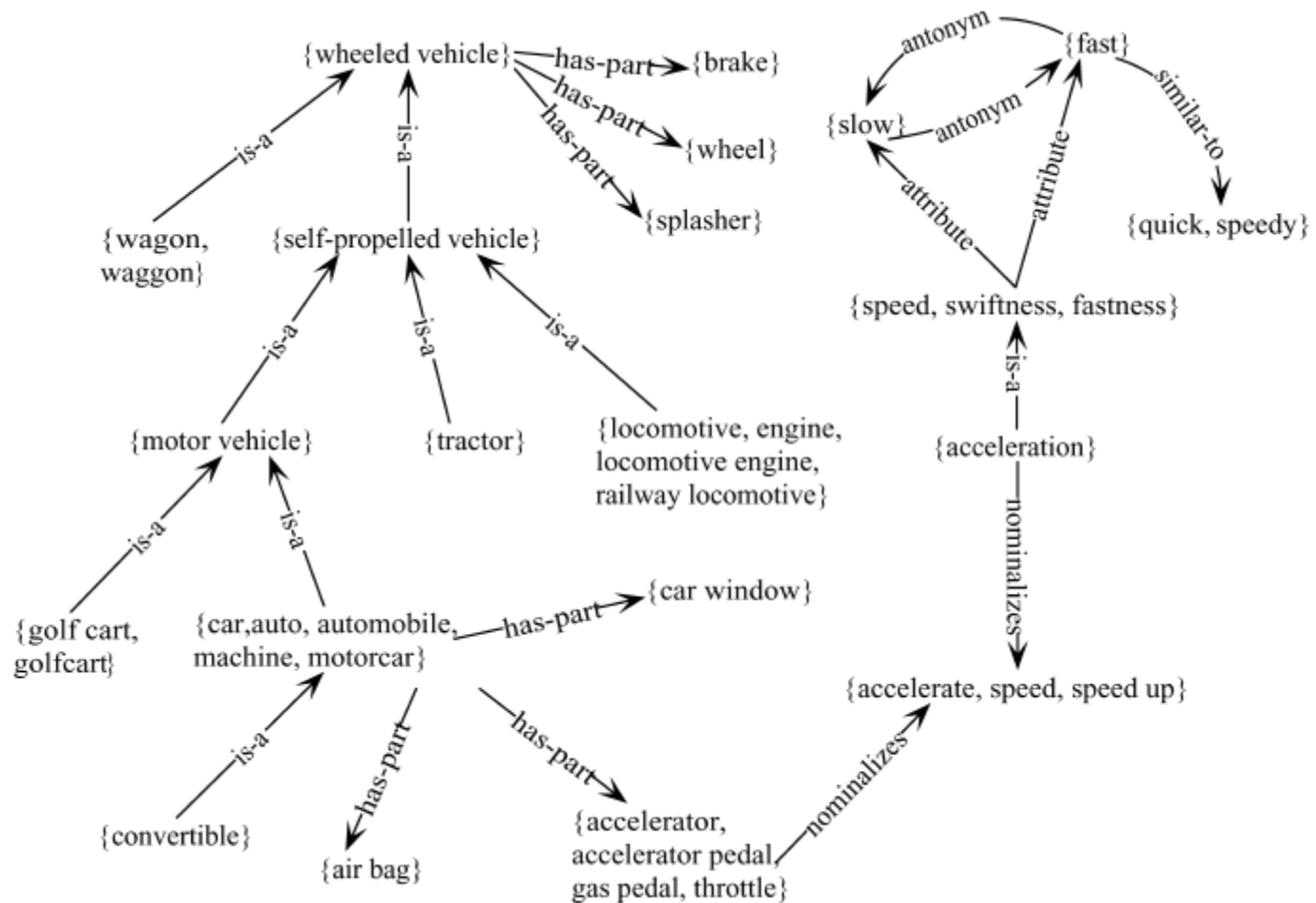


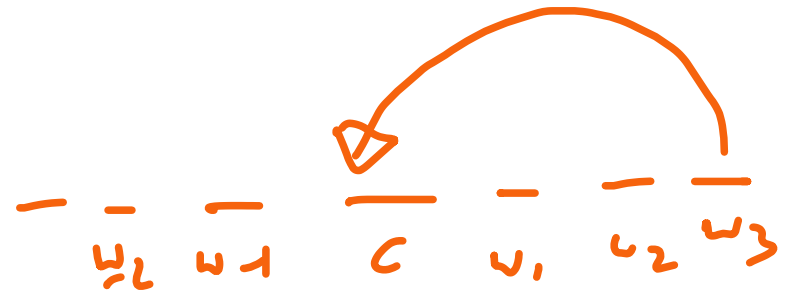
Fig. 3 in paper

WSD as classification

1. Significant difference wrt other classification tasks:
WSD involves n different classification tasks where n is the size of the lexicon

2. Two categories of WSD :

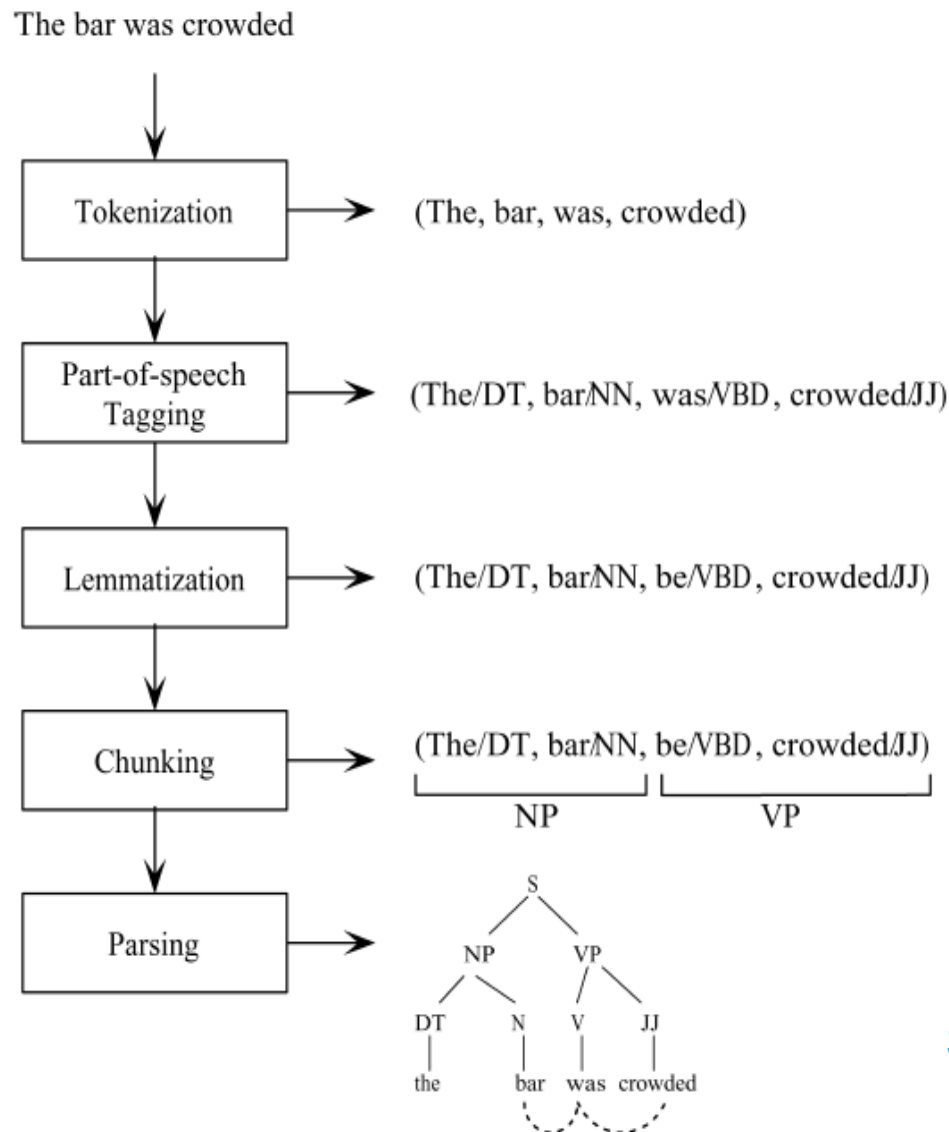
1. Lexical Sample WSD
2. All-words WSD



Features for WSD

1. Part-of-speech ✓
2. Context words —
3. Stemming/morpho processing on context words
4. Partial parsing : thematic or grammatical roles
Alternately : Chunking + parsing

Steps in Feature Extraction



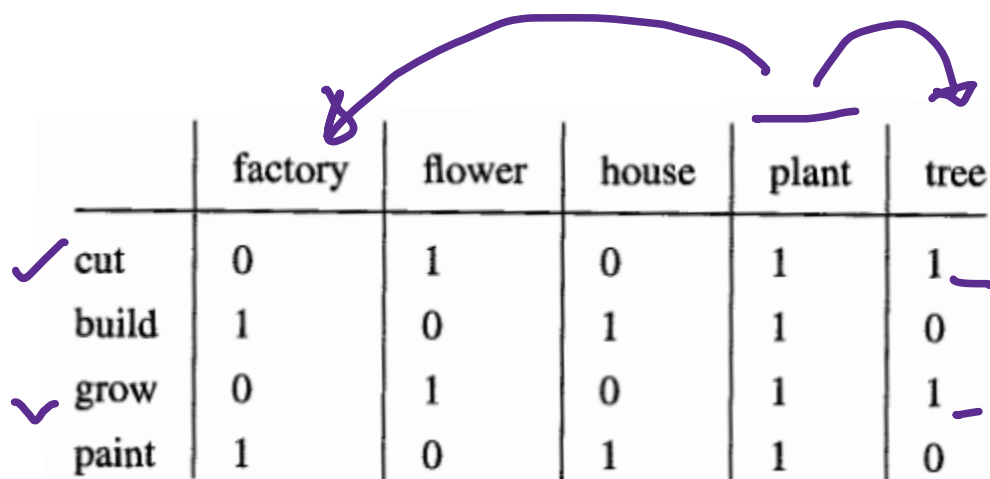
Section 2.3

Features for WSD

1. Local features
2. Topical features
3. Syntactic features
4. Semantic features

Features for WSD

- Selectional Restrictions



	factory	flower	house	plant	tree
✓ cut	0	1	0	1	1
build	1	0	1	1	0
✓ grow	0	1	0	1	1
paint	1	0	1	1	0

Ack: Charniak's book

Features for WSD

- Higher Order Associations

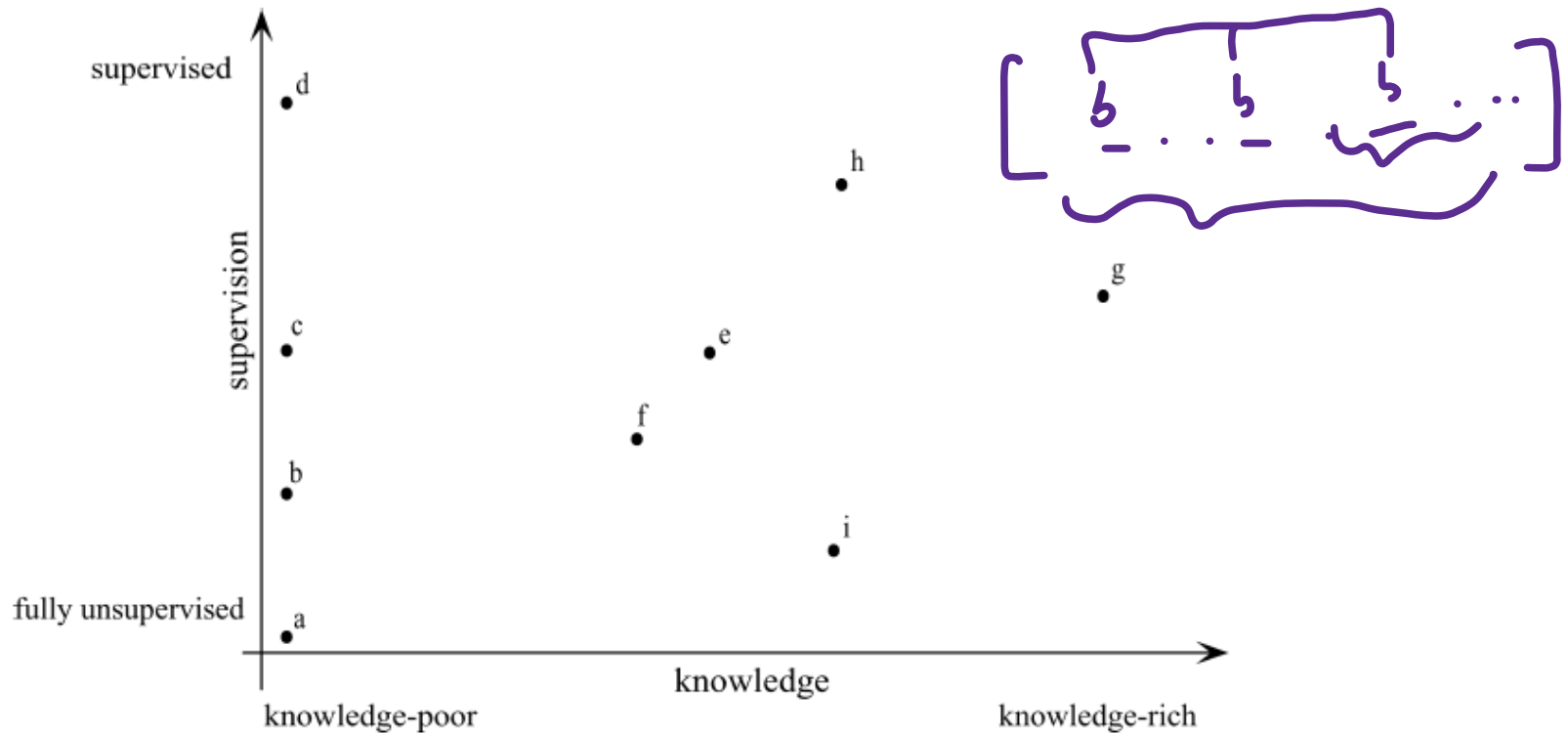
that program . . .” Hence the line to be adopted at the UN against Soviet verbal assaults: the United States would admit to having military programs in *space*, but stress their importance to peace and contrast American candor to Soviet secrecy. Even liberal Senator Albert Gore (D., Tenn.) concurred: “no workable dividing line ([33], p. 348)

of the world, then the United States must show that its affluence reached the poor and colored at home. Rostow also had strong views on *space* policy. As a member of the Greenewalt Committee, he considered technological competition to be critical and had “a bias toward hope rather than skepticism.” But ([33], p. 217)

Ack: Charniak’s book

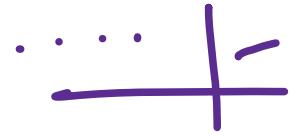
WSD approaches



- Three axes:
1. Supervised/Unsupervised
 2. Knowledge-Rich/Knowledge-Poor
 3. Type-based/Token-based

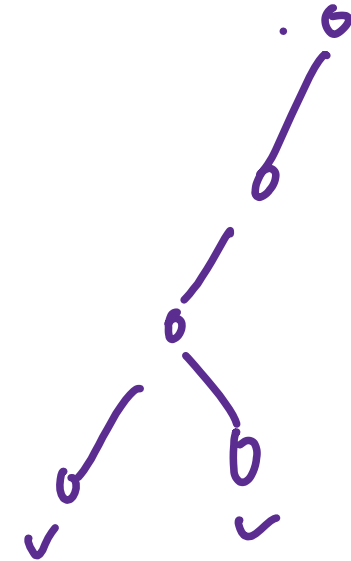


Refer: Section 2.4

Supervised WSD



- Decision Lists
- Decision Trees (rules) 
- Naïve Bayes
- Neural Networks
- Support Vector Machines
- Instance Based Learning 



Refer: Section 3

To be continued