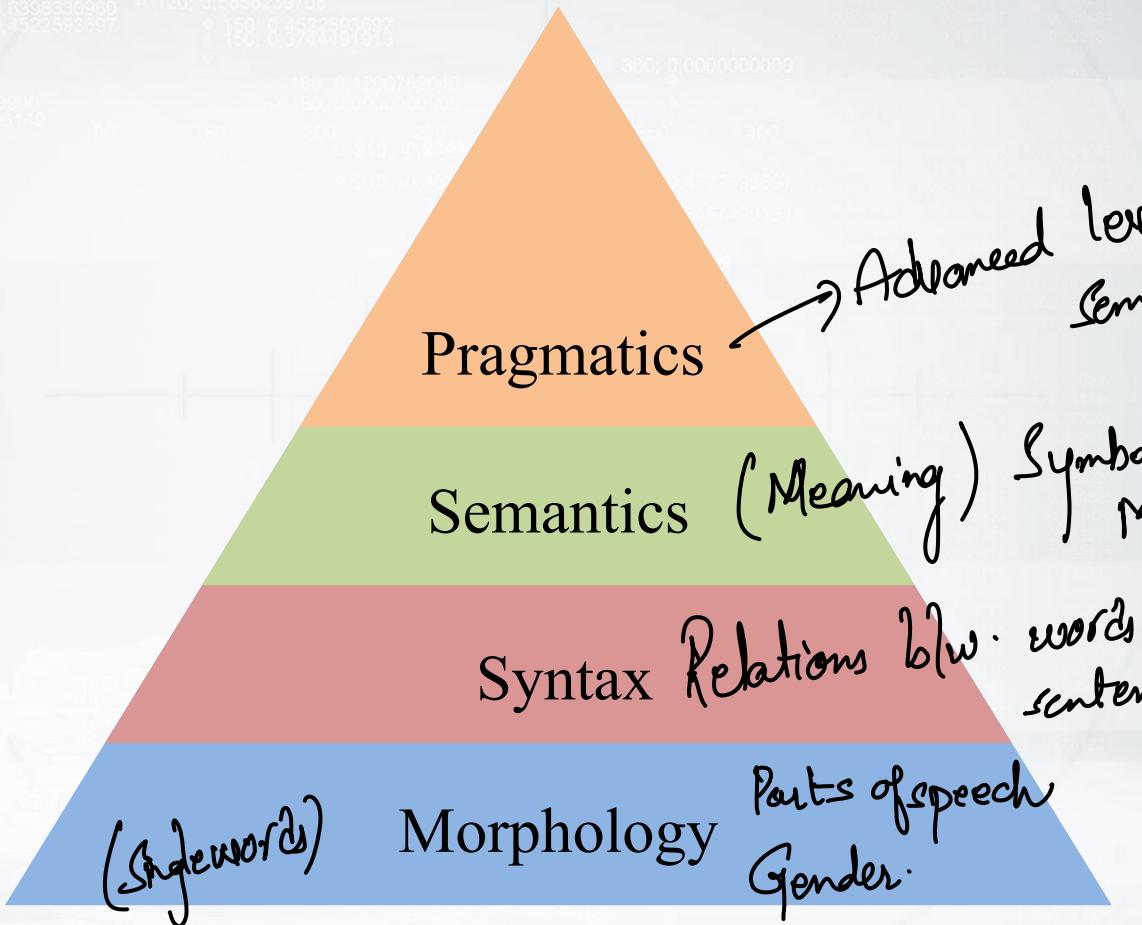


NLP is not only about Mathematics but also linguistics.

Linguistic knowledge in NLP

NLP Pyramid



Natural Language Processing Pyramid

Libraries and tools

- **NLTK** (Morphological & Syntax)

- Small but useful datasets with markup
- Preprocessing tools: tokenization, normalization...
- Pre-trained models for POS-tagging, parsing...

- **Stanford parser** (Syntax analysis)

- **spaCy**: python and cython library for NLP
- **Gensim**: python library for text analysis, e.g. for word embeddings and topic modeling
- **MALLET**: Java-based library, e.g. for classification, sequence tagging, and topic modeling

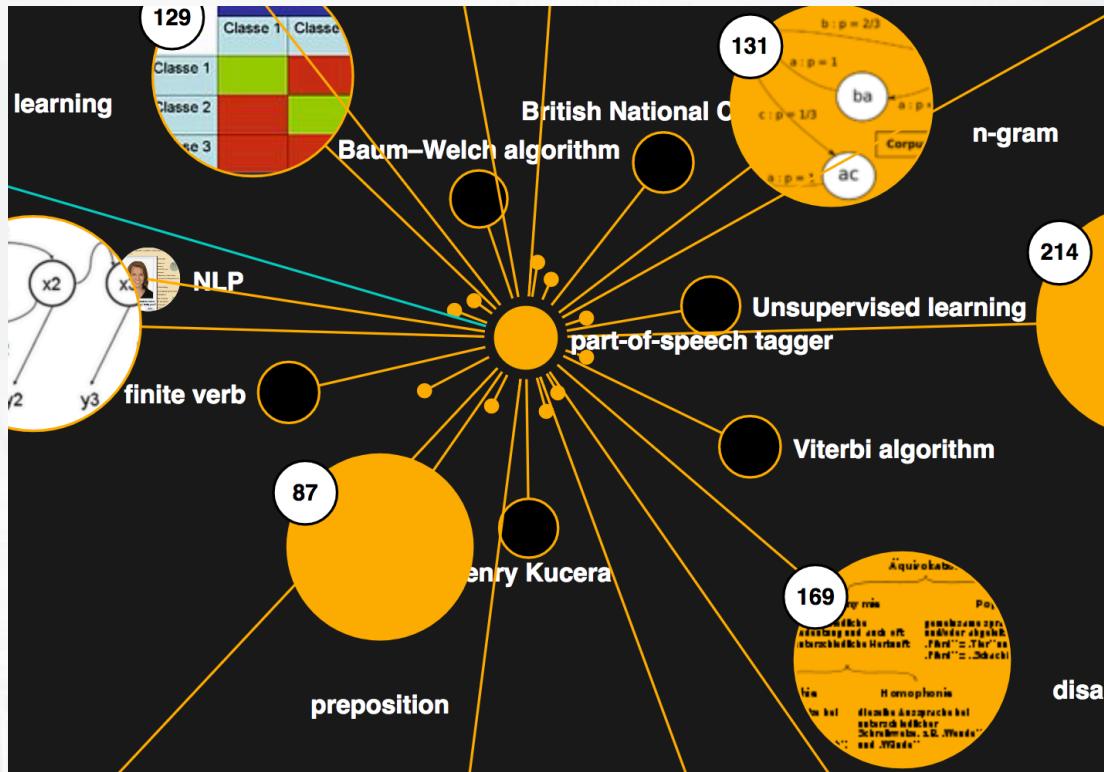
high level abstractions
of topic model ...

Linguistic knowledge

(Relations b/w words)

- Ideas and evaluation
- External resources: WordNet, BabelNet, etc.

→ multilingual
(Relations)
(Nearest
Neighbors)



<http://babelnet.org/synset?word=NLP&lang=EN&details=1&orig=NLP>

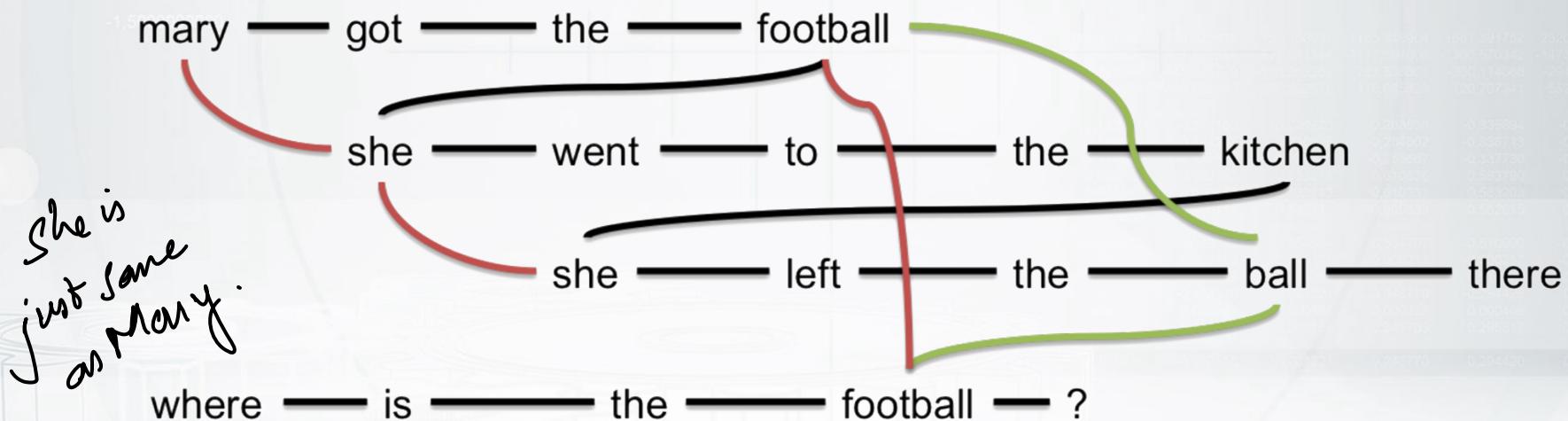
Linguistic knowledge + Deep Learning

Task: Question Answering / Reasoning

Linguistic links: co-reference (red), hypernyms (green)

Method: DAG-LSTM

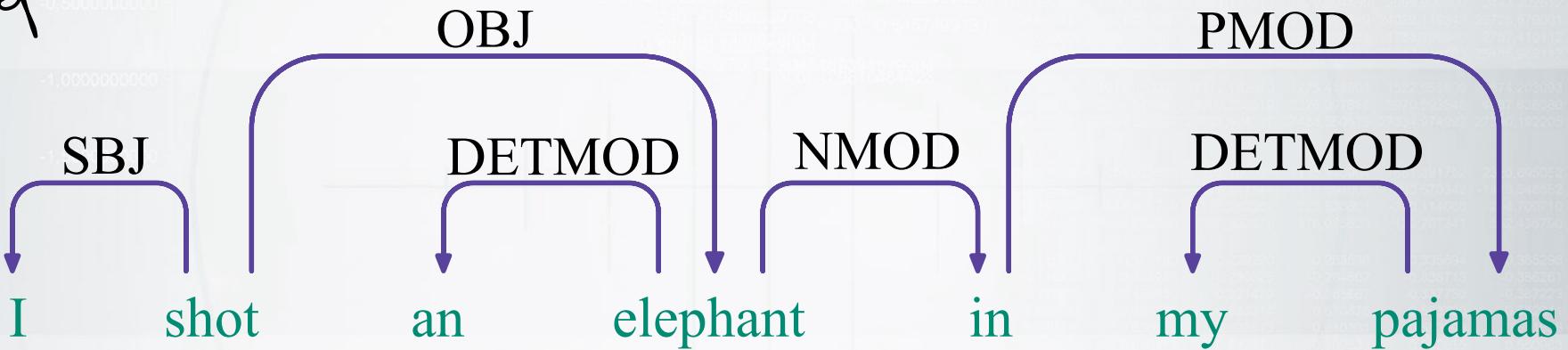
(Dynamic Asyclic Graph LSTM)



Dhingra et. al. Linguistic Knowledge as Memory for Recurrent Neural Networks, 2017.

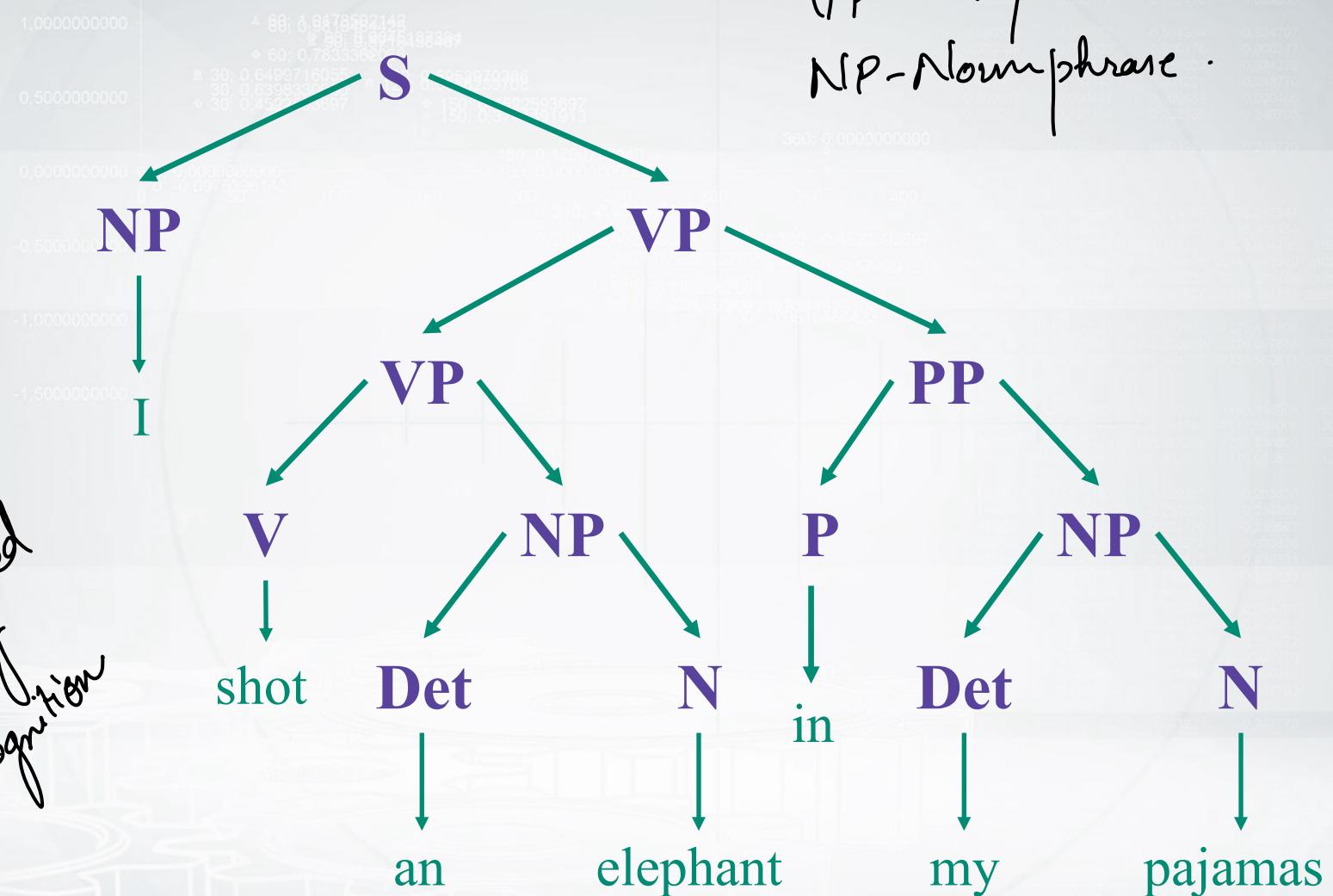
Syntax: dependency trees

dependency between the words.



Syntax: constituency trees

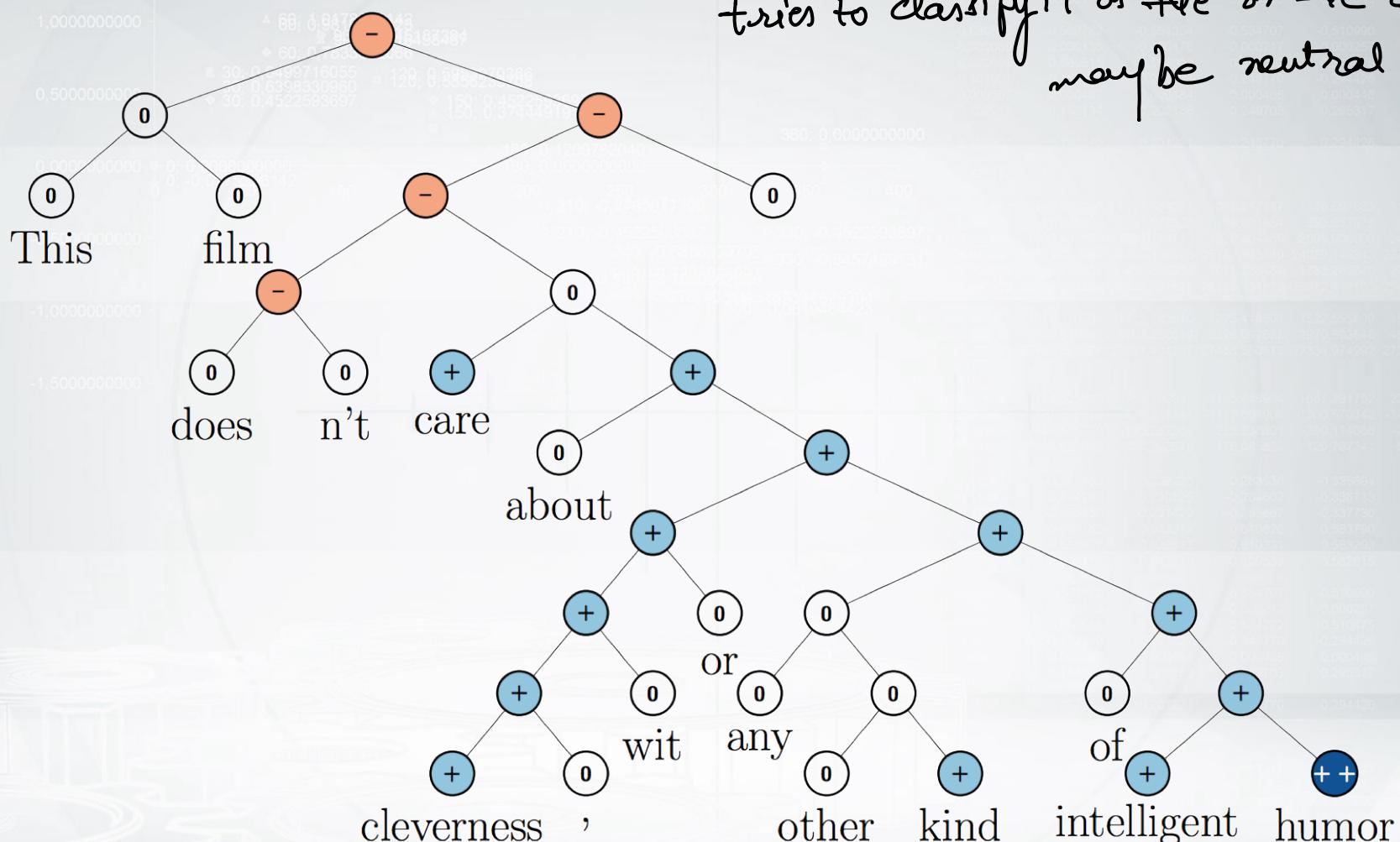
Named entity
recognition



VP - Verb phrase

NP - Noun phrase

Sentiment analysis → treats reviews as some pieces of text and tries to classify it as +ve or -ve or may be neutral -



Socher et. al. Recursive Deep Models for Semantic Compositionality Over a Sentiment Treebank, 2013.