Introduction to NLP

Ex. of application:
- ↳ Question Answering
- ↳ Information Extraction
- ↳ Sentiment Analysis
- ↳ Machine Translation

Spam detection   Mostly solved
Part-of-Speech (POS) Tagging
Named entity recognition (NER)

Question Answering (QA)
Paraphrase   Still
Summarization   really
                    hard
Dialog

Sentiment Analysis   making
Coreference resolution   good
                            progress
Word sense disambiguation
Parsing
Machine Translation (MT)
Information Retrieval (IE)

neologisms   segmentation   idioms
                issues

Why is NLP difficult? → non-standard (language)

↳ Ambiguity "crash blossoms"

─ Violinist linked to JAL Crash Blossoms ← multiple
                                            interpretation
─ Red Tape Holds up new bridge   words sense
                                            1. delay
                                            2. to support

Ambiguity is pervasive

phrase structure

Fed raises interest rates

Fed raises interest rates
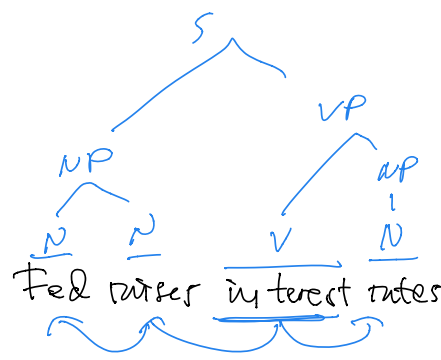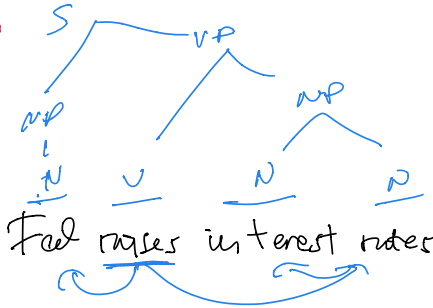
So, how do tackle and make progress
  on this problem? → need tools → knowledge about language
                                  → knowledge about the world
        use
  probabilistic model ←————————— → A way to combine knowledge
  built from data   generally   resources