Topics in Natural Language Processing

Shay Cohen

Institute for Language, Cognition and Computation

University of Edinburgh

Lecture 1

Topics in Natural Language Processing

This course:

- An advanced course about natural language processing and machine learning
- Composed of two parts:
  - Lectures
  - Presentations by students

What's Important?

To have a career in NLP or ML (Andrew Ng):

- Read papers
- Try to replicate results
- Do dirty work

This course is in a seminar format: it tries to make existing papers with state-of-the-art methodology more accessible to all students in class

Use this opportunity to improve your public speaking skills, your immersion in reading research papers and learning how to ask the right questions during lectures/presentations

Topics in NLP

Prerequisites:

- Some familiarity with machine learning and probability
- If something is unclear, ask!

Things to Do:

- Student presentations (20%)
- Brief paper responses (15%)
- Assignment (10%)
- Essay (55%)

Office hours: By appointment

Essay

- Find a specific problem in NLP
- ... Preferably one with “structure”
- Cover the core approaches to solve it
- ... Preferably heavily oriented towards statistical learning
- Cover the way evaluation is done
- Current obstacles to solve the problem and future directions

I will send the criteria I will use to grade the essays

NLP in the Old Days

1950s-1980s: rule based systems

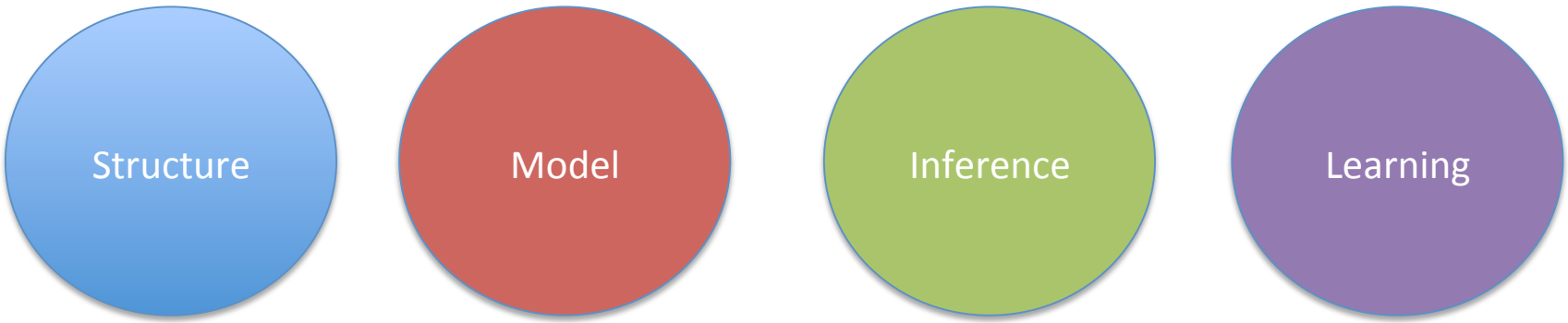
IBM'S WATSON (right) AND FRIENDS:† For a mathematical wizard . . .

NLP Now

late 1980s until now: statistical learning

Solving an NLP Problem

When modelling a new problem in NLP, need to address four issues:



Roadmap

- Introduction and Basic Refresher (today)
- Statistical paradigms in NLP
- Bayesian learning
- Grammar models – beyond context free
- Inference
- Linear models and learning
- Neural networks and deep learning
- Learning from incomplete data
- Evaluation and experimental design

Learning

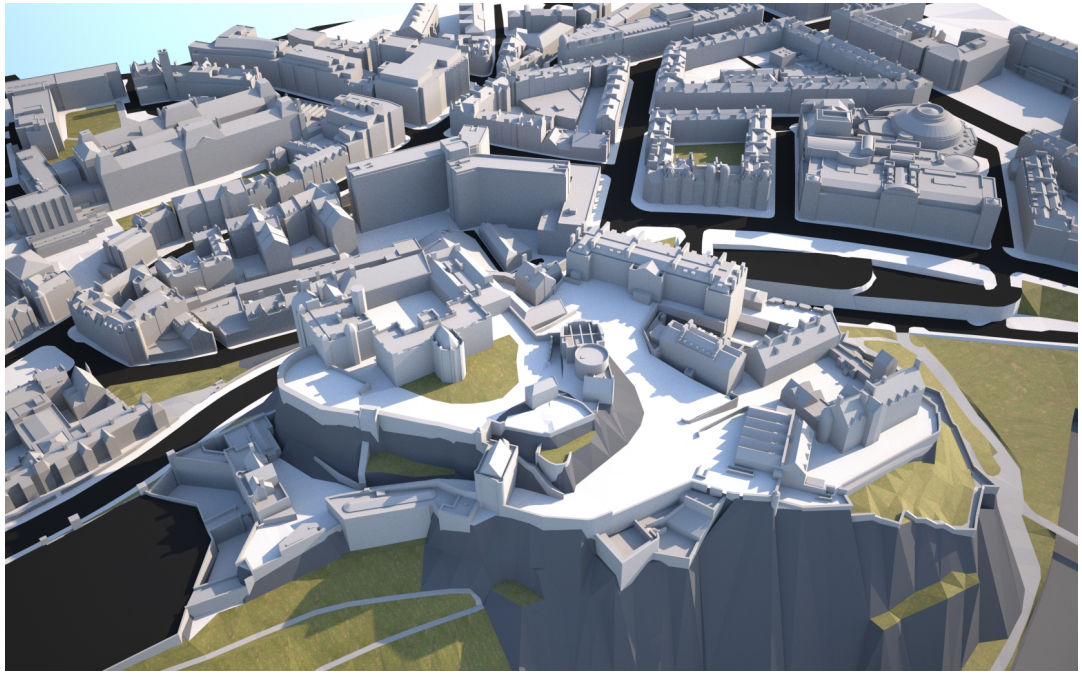
Learning is:

- Experience translated into expertise/knowledge
- Memorisation with generalisation

Machine learning and NLP:

- Experience = Training data
- Knowledge = Decoder or Prediction Model
- Used to either **mimic** humans or **transcend** their abilities

What is a Model?



From Merriam-Webster:

- a usually small copy of something
- a set of ideas and numbers that describe the past, present, or future state of something (such as an economy or a business)

When is a model a good model?

When is a Model a Good Model?

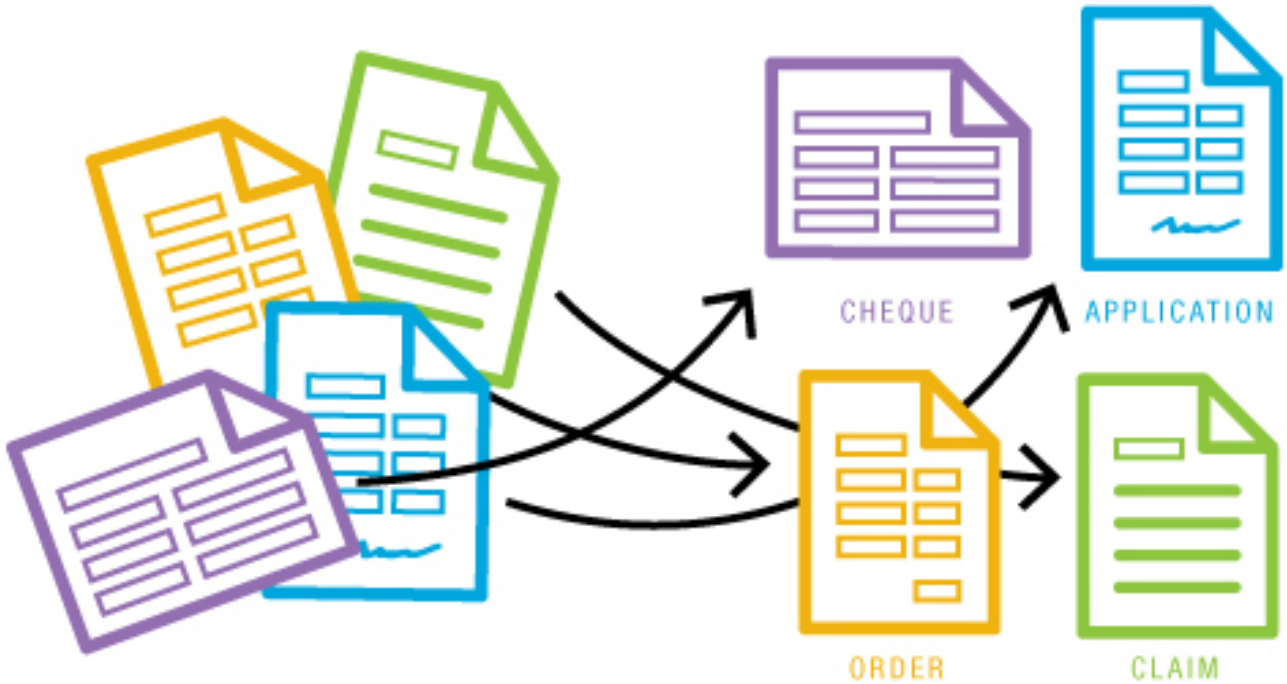
- Its predictions are correct
- Interpretable (Lipton, 2017):
  - So that we can trust it
  - So that it reflects causal relationships
  - So that we can transfer the knowledge to other domains
  - So that we can make fair and ethical decisions

What is a Statistical Model?

Predict the future. Probabilistically.

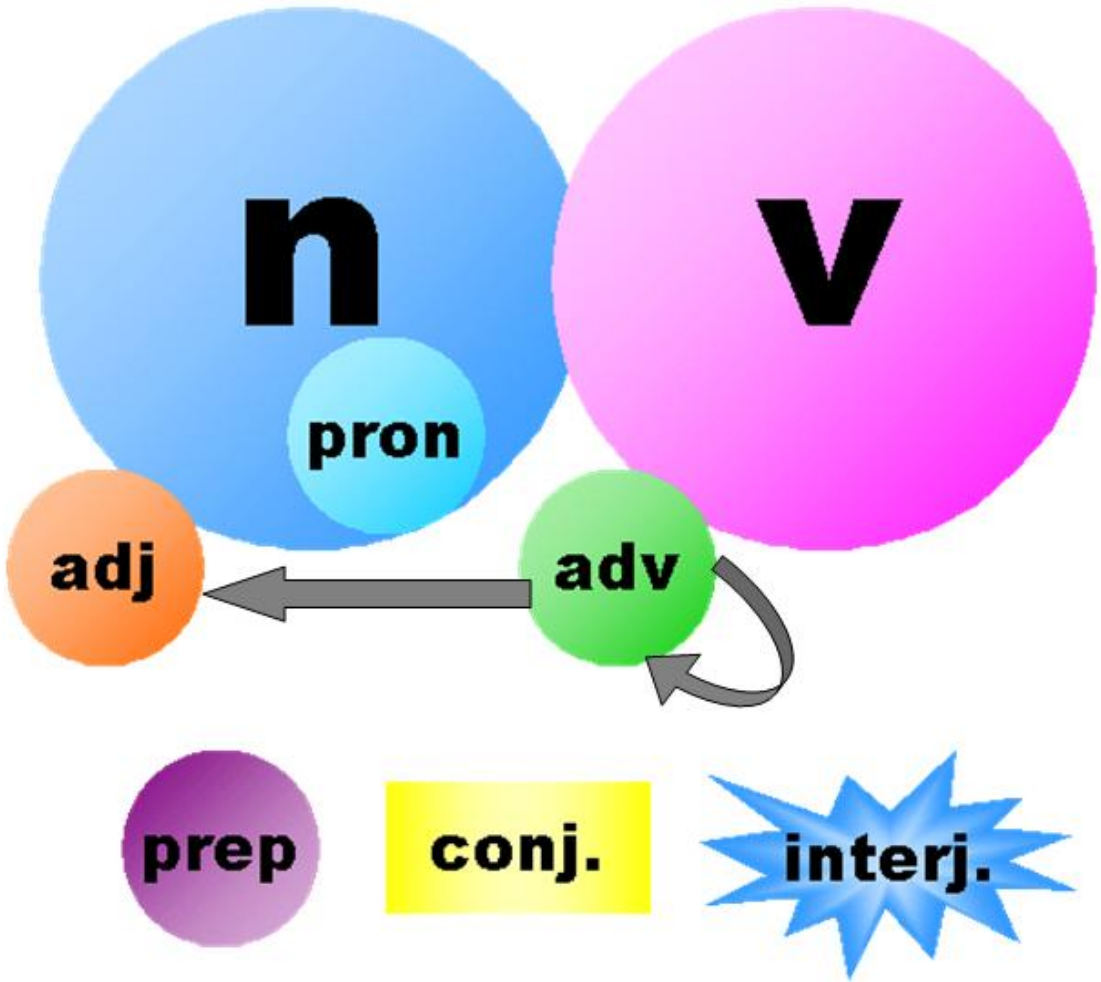
NLP Problem Example: Document Classification

sentiment analysis, document topic, ...



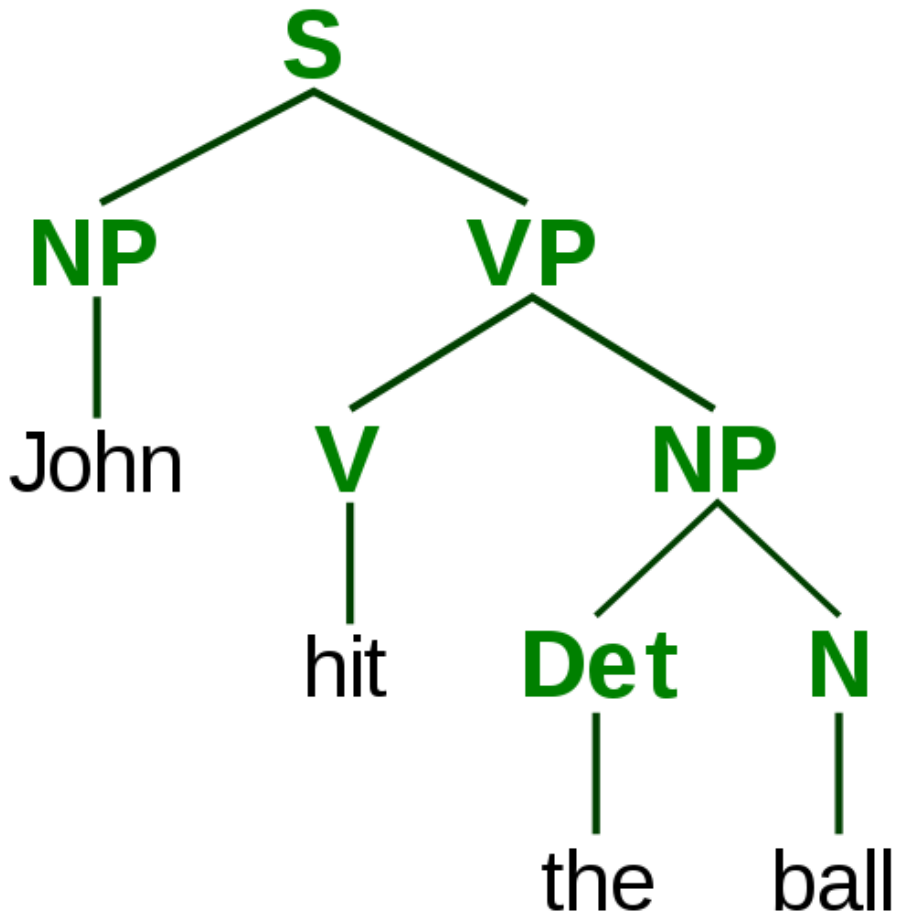
NLP Problem Example: POS Tagging

map words to their part-of-speech tags



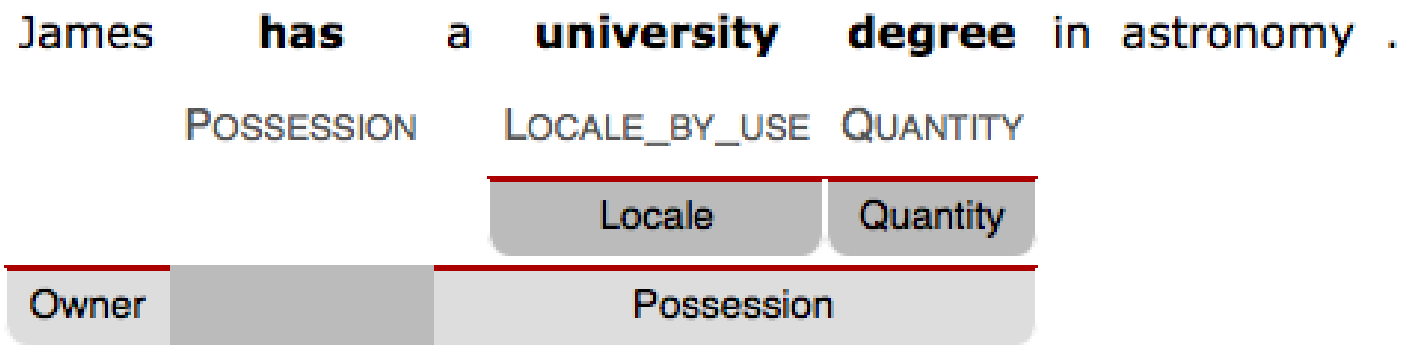
NLP Problem Example: Parsing

map sentences to their syntax



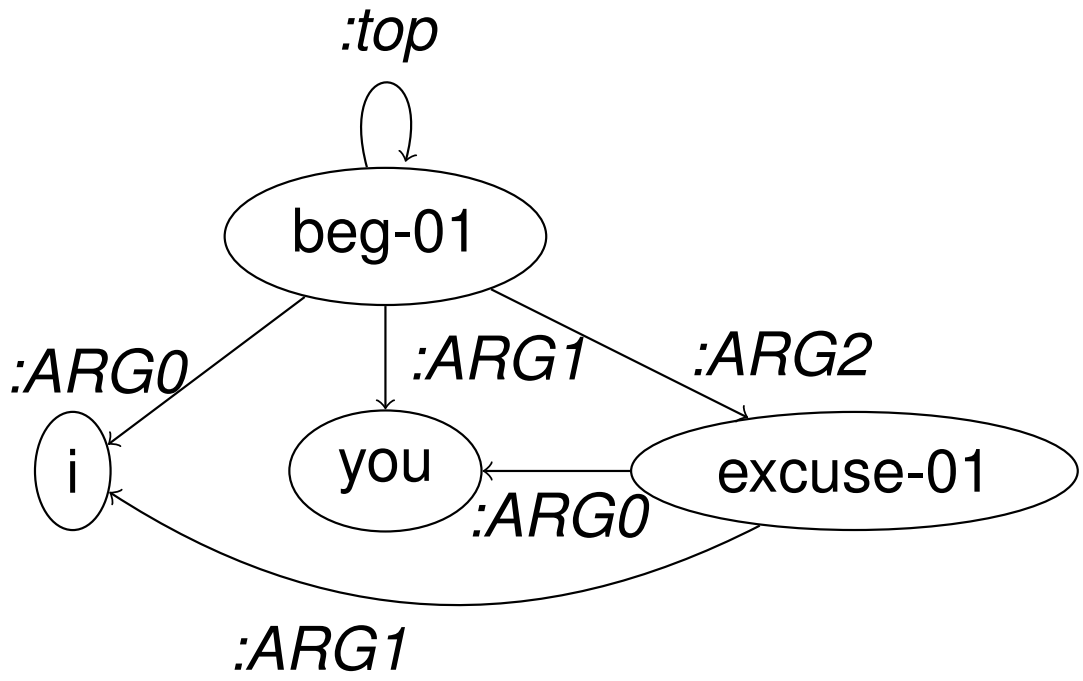
NLP Problem Example: FrameNet Parsing

find predicate-argument structure



NLP Problem Example: AMR Parsing

map sentences to a graph representation of their meaning



Back to Modelling

What if the space to model is complex? Modelling documents.

Modelling a Problem

- Define a sample space
- Define the structure of the sample space
- Decide on a parametrisation

Then one can proceed with data collection and learning

Modelling - Tradeoffs

- “Exact copy”, detailed
- Not too many parameters
- Efficient to work with

Our Endless Pursuit of Knowledge

Exact sciences:

- Deal in axioms and theorems. If you prove it, it is true under your axioms. Examples: Maths, CS theory

Empirical sciences:

- Deal in fact and theories. Knowledge is temporary. Examples: Physics, Biology, Linguistics

Engineering:

- Deals in artifacts. It is good if it works. Examples: Applied CS

Where is NLP situated?

Discussion

NLP: Science vs. engineering

Discussion

NLP: Science vs. engineering

- What makes NLP a science? Empirical or exact?
- What makes NLP an engineering field?

Next class

Paradigms in statistical learning

- The frequentist approach
- The Bayesian approach
- “Computer science approaches?”