CIS-700

Spring 2020

Reasoning for Natural Language Understanding

Dan Roth

Computer and Information Science

University of Pennsylvania

This class

- Understand early and current work on Reasoning
  - (Learn to) read critically, present, and discuss papers
- Understand some of the difficulties in NLU from the perspective of reasoning
  - Conceptual and technical
- Try some new ideas

- How:
  - Presenting/discussing papers
    - Probably: 2 presentations each; 4 discussants
  - Writing a few critical reviews
  - "Small" individual project (reproducing);
  - Large project (pairs)
  - Tentative details are on the web site.

- Machine Learning
  - 519/419
  - 520
  - Other?
- NLP
  - Yoav Goldberg's book
  - Jurafsky and Martin
  - Jacob Eisenstein

- Attendance is mandatory
- Participation is mandatory

- **Time of class?**

What's Important in order to make progress in NLU

- How to make progress towards natural language understanding
  - Learning and Reasoning; knowledge
- Dispel with [some] of the currently hot trends
  - If we want to reach the moon…
- What is Reasoning?
- Reasoning in Natural Language Understanding
- Today: Examples & Discussion

Natural Language Understanding

No connective

- At least 14 people have been killed in southern Sri Lanka, police said. The telecoms minister was among about 35  in the blast site at the town of Akuressa, 160km (100 miles) south of the capital, Colombo. Government officials were attending a function at a mosque to celebrate an Islamic holiday at the time. The minister said later that the suicide attack was carried out by ….
  - ➔ 49 people were hit by a suicide bomber in Akuressa.

This is an Inference Problem

Reasoning is about providing Reasons

- Requires dealing with a large number of phenomena
  - lexical, quantitative, co-reference, semantic types, discourse convention, knowledge…
  - But, this discussion indicates that we “think” when we understand a story.
- **Understanding is telling ourselves a story about the story**

John McCarthy on Natural Language Understanding

AN EXAMPLE FOR NATURAL LANGUAGE UNDERSTANDING AND THE AI PROBLEMS IT RAISES

John McCarthy
Computer Science Department
Stanford University
Stanford, CA 94305
jmc@cs.stanford.edu
http://www-formal.stanford.edu/jmc/

1976

A New York Times Story

"A 61-year old furniture salesman was pushed down the shaft of a freight elevator yesterday in his downtown Brooklyn store by two robbers while a third attempted to crush him with the elevator car because they were dissatisfied with the $1,200 they had forced him to give them.

The buffer springs at the bottom of the shaft prevented the car from crushing the salesman, John J. Hug, after he was pushed from the first floor to the basement. The car stopped about 12 inches above him as he flattened himself at the bottom of the pit.

A New York Times Story (Cont.)

Mr. Hug was pinned in the shaft for about half an hour until his cries attracted the attention of a porter. The store at 340 Livingston Street is part of the Seaman's Quality Furniture chain.

Mr. Hug was removed by members of the Police Emergency Squad and taken to Long Island College Hospital. He was badly shaken, but after being treated for scrapes of his left arm and for a spinal injury was released and went home. He lives at 62-01 69th Lane, Maspeth, Queens.

He has worked for seven years at the store, on the corner of Nevins Street, and this was the fourth time he had been held up in the store. The last time was about one year ago, when his right arm was slashed by a knife-wielding robber."

New York Times Story: Questions

- An intelligent person or program should be able to answer the following questions based on the information in the story:
  - The article proceeds with 22 questions:

1. Who was in the store when the events began?
   - Probably Mr. Hug alone, although the robbers might have been waiting for him, but if so, this would have been stated.

2. What did the porter say to the robbers?
   - Nothing, because the robbers left before he came.

20. Why did Mr. Hug yell from the bottom of the elevator shaft?
   - So as to attract the attention of someone who would rescue him.

- "The above list of questions is rather random. I doubt it covers all facets of understanding the story."

McCarthy's Challenges

The QA module is not being trained
Once the program knows English, and has the relevant background knowledge, it should answer the questions

- A formalism capable of expressing the assertion of the sentences free from dependence on the grammar of the English language. ("Artificial Natural Language", ANL)
  - Semantic Parser
- An "understander" that constructs the "facts" from the text.
  - Information Extraction: Entities, Relations, Temporal, Quantities,…
- Expression of the "general information" about the world that could allow getting the answers to the questions from the "facts" and the "general information"
  - Background Knowledge
- A "problem solver" that could answer the above questions on the basis of the "facts".
  - Question Answering Engine

Lessons

- What can we learn from this example?
  - Difficulties of NLU
  - Importance of reasoning
    - Part of Reasoning here seems to be “providing the reasons”, not only the “answers”
  - Decoupling learning from reasoning
    - McCarthy thinks that there is a need for some level of abstraction – an abstract representation of the text and the relevant knowledge so that a generic module can work on it and “do the reasoning”.
- Is this important/Essential?

Cognitive Psychology of Reasoning

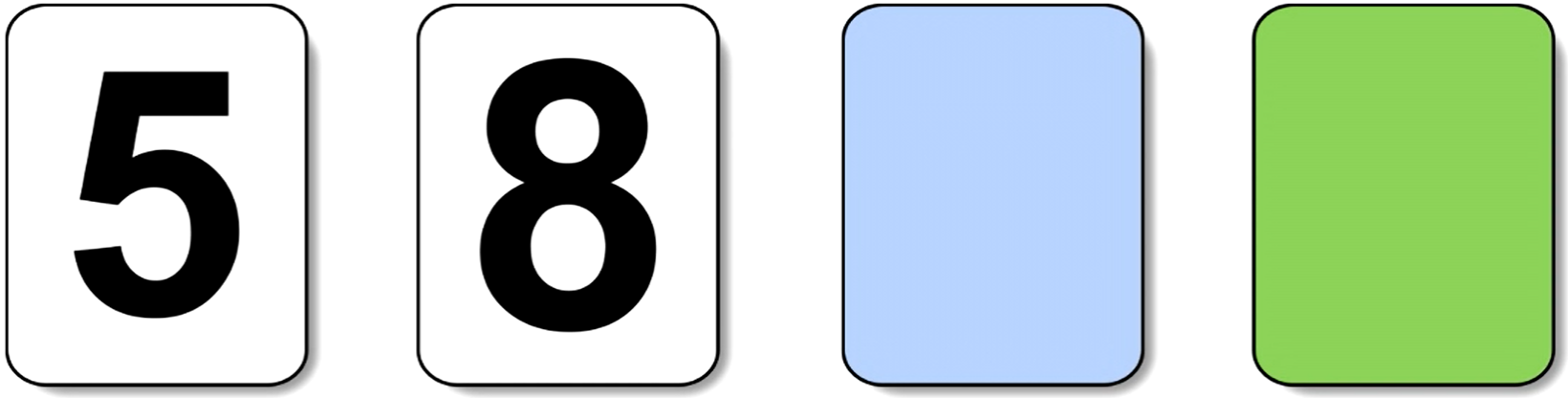
- McCarthy's argument starts with “most people should…”

- There have been many studies by cognitive psychologists that try to identify the types of reasoning people can do, and how.
- Earlier theories:
  - “human reasoning depended on formal rules of inference akin to those of a logical calculus”

- Philip N. Johnson-Laird: Mental models and human reasoning
  - http://www.modeltheory.org/papers/2010mms&human-reasoning.pdf

Can **you** Reason? (30 seconds; write it down; don't communicate)

- You are shown 4 cards; each one has a number on one face and a color on the other.
- You are asked to determine the truth value of the following statement:
  - "If a card shows an even number on one face then its opposite face is blue."

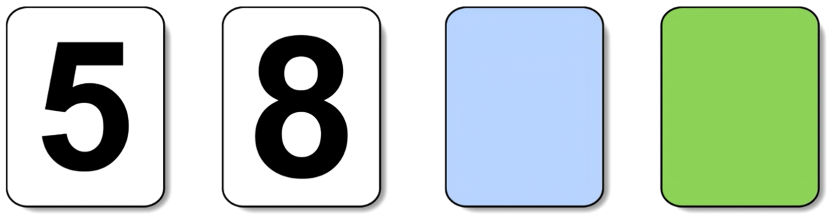


  - Your goal: to turn over the minimal # or cards to allow you to determine if this statement is true.
  - **Which card(s) should you turned over?**
  - (30 seconds; write it down)

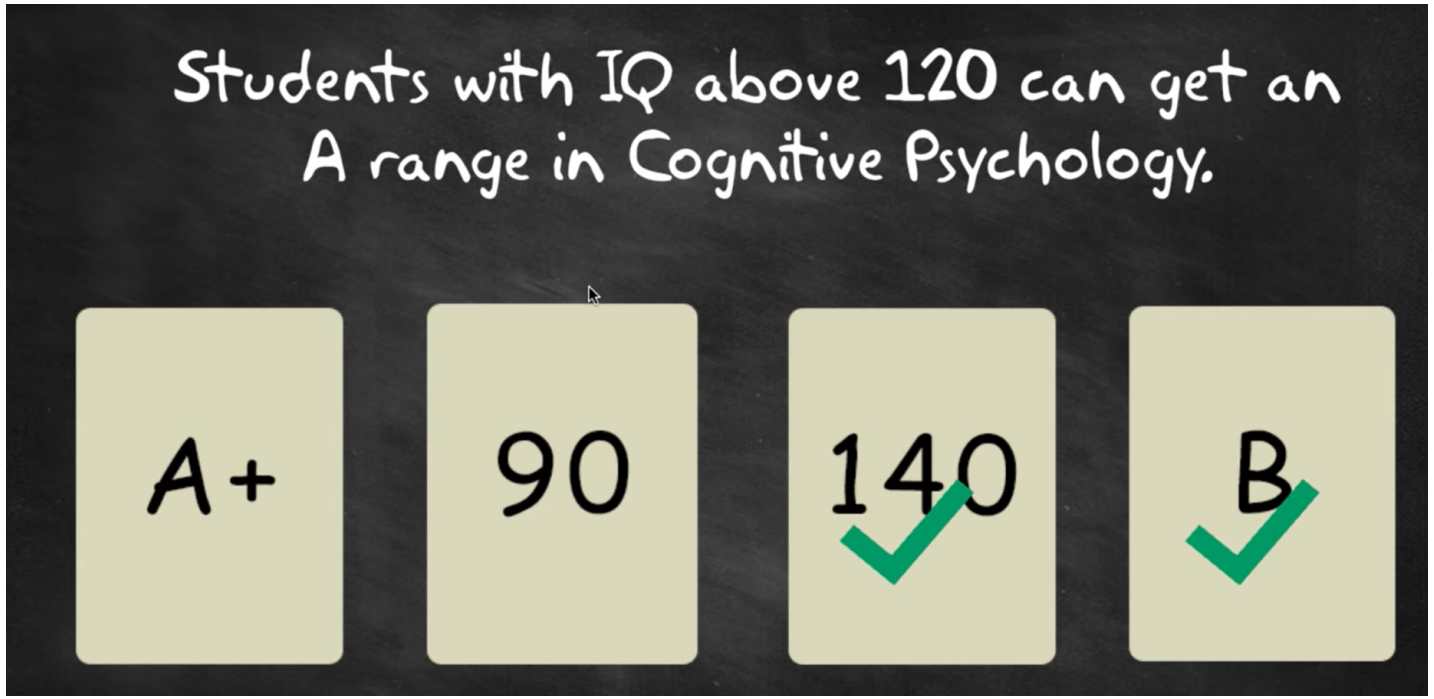
Can you Reason?

- Most people that are not trained in logic get it wrong.
- But, when given equivalent tests that are grounded in some familiar situations, they get it right.
  - “If a student scores less than 650 in a quantitative GRE test they cannot get into a top CS program”
    - Cards with scores and programs the student is enrolled in.
  - “If a student is younger than 21 they cannot buy beer in the US”
    - Cards with age and buy/cannot buy statements

If a card shows an even number on one face then its opposite face is blue



Reasoning in Explicit Situations



- These experiments were part of the development of a theory of Mental models:
- https://www.researchgate.net/publication/228408902_Mental_models_a_gentle_guide_for_outsiders
- These tests have probabilistic analogies, by Amos Tversky and Daniel Kahneman
  - Check: **The Undoing Project: A Friendship That Changed Our Minds,** by Michael Lewis**.**

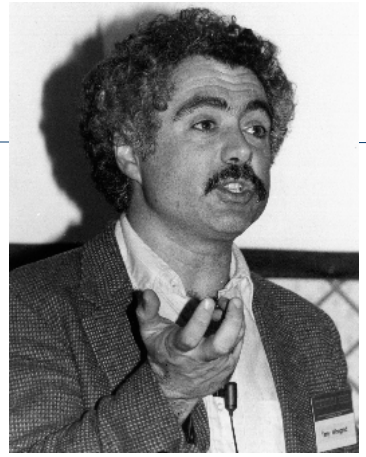
Should We Care?

- Mr. Micawber's famous advice (in Dickens's, **David Copperfield**):
  - Annual income twenty pounds, annual expenditure nineteen pounds nineteen and six, result happiness. Annual income twenty pounds, annual expenditure twenty pounds ought and six, result misery
- What is its logical form?
  - Do we need to work it out?
  - Do we need to do it for all everyday assertions?
- Clearly, converting into a "logical form" requires more than just syntax.
  - It depends on knowledge.
- Here:
  - nineteen pounds nineteen (shillings) and six (pence) is less than twenty pounds.
  - happiness and misery are inconsistent properties.
  - Possibly other contextual cues.

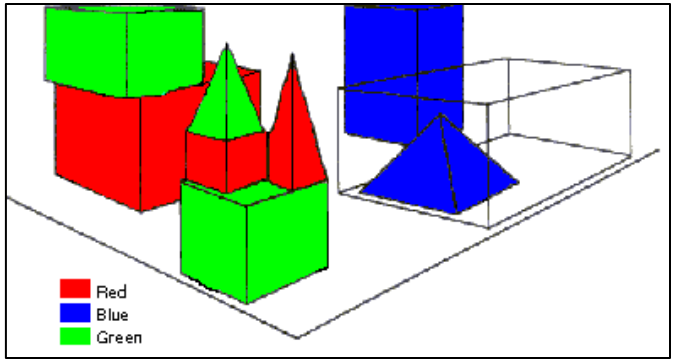
Should We Care (2)?

- Do we need a “formal” abstractions to support reasoning?
- Developmental psychologists think that there is merit to this.
  - That “reasoning is commonsense”
  - A toddler knows that a cloth dropped into a bucket of water comes out wet **(deduction)**, and that
  - A dry cloth that is put into a bucket and comes out wet probably means that the bucket contains water **(abduction)**.
- But, what does that mean to the computational theories we need to develop to pursue AI, and NLU?
- Perhaps a more basic question is that of “what representations could support this level of behavior/reasoning”?

SHRDLU (Terry Winograd, 1968)



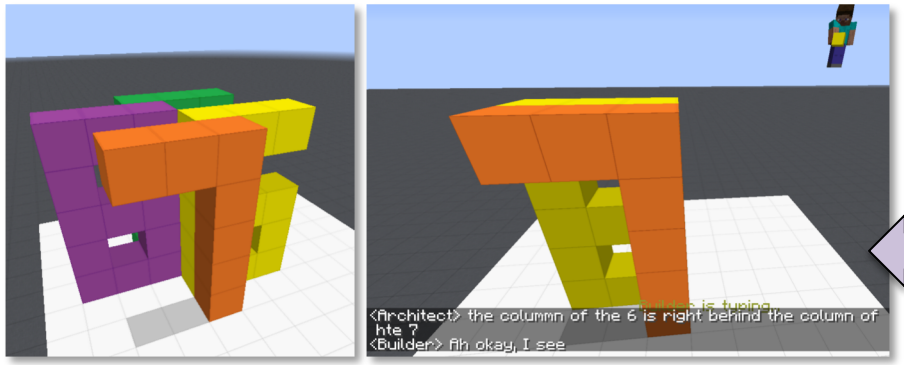
- **Person:** Pick up a big red block.
- **Computer:** OK.
- **Person:** Grasp the pyramid.
- **Computer:** I DON'T UNDERSTAND WHICH PYRAMID YOU MEAN.
- **Person:** Find a taller block than the one you are holding and put it into the box.
- **Computer:** BY "IT", I ASSUME YOU MEAN…..
- **Computer:** OK.
- **Person:** What does the box contain?
- **Computer:** THE BLUE PYRAMID AND THE BLUE BLOCK.
- **Person:** What is the pyramid supported by?
- **Computer:** **WHAT DO YOU MEAN BY “SUPPORTED”?
- …….
- **Person:** How many blocks are not in the box?
- **Computer:** FOUR OF THEM.



Understanding, executing and interacting requires language understanding & reasoning to recognize goals and advance the task

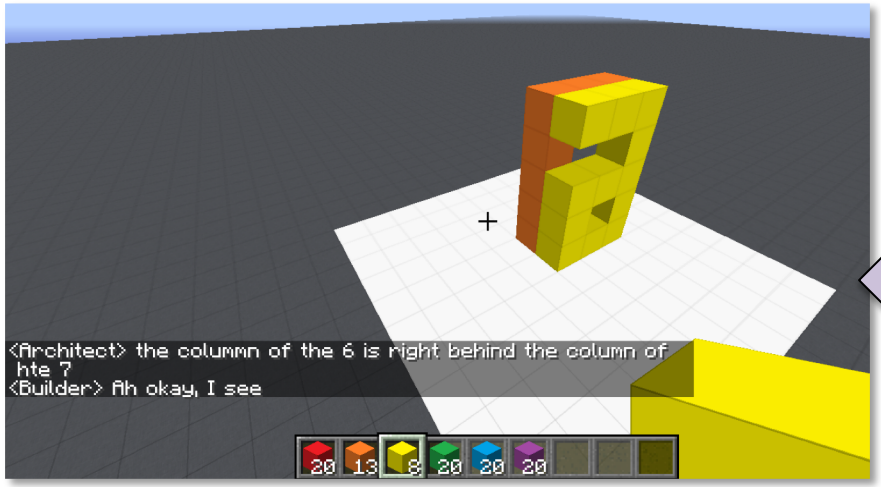
Communication (CwC Project, 2017-2019)

ARCHITECT



Target Structure Build Region

BUILDER



CHAT INTERFACE

Architect: in about the middle build a column five tall

(Builder puts down five orange blocks)

Architect: then two more to the left of the top to make a 7

(Builder puts down two orange blocks)

Architect: now a yellow 6

Architect: the long edge of the 6 aligns with the stem of the 7 and faces right

Builder: Where does the 6 start?

Architect: behind the 7 from your perspective

Builder: Is it directly adjacent?

Architect: yes directly behind it. touches it

(Builder puts down twelve yellow blocks, in the shape of a 6)

Architect: too much overlap unfortunately

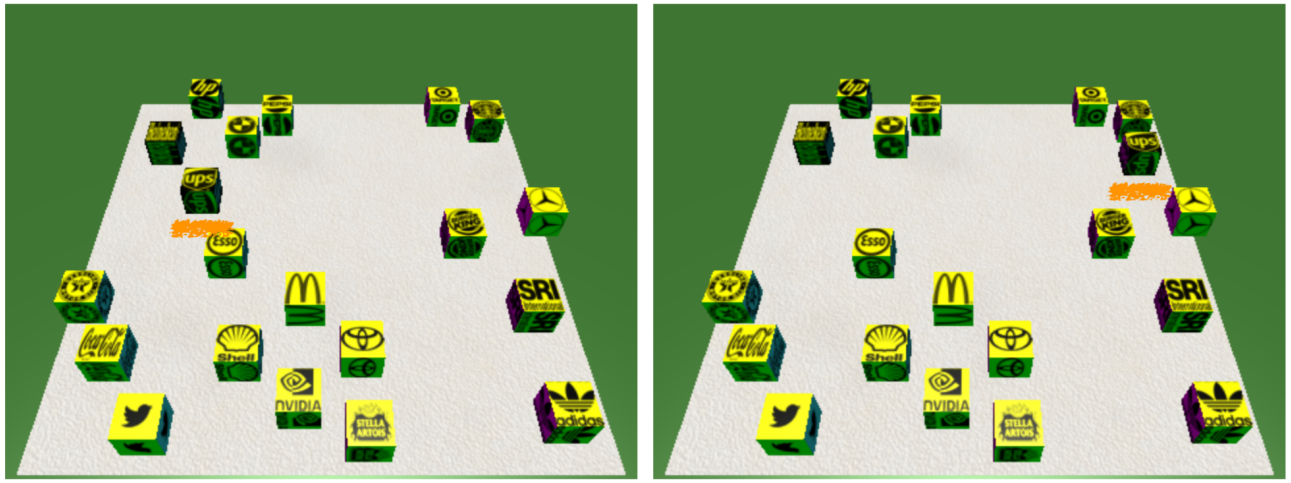
Architect: **the columnn of the 6 is right behind the column of hte 7**

[Minecraft example from Julia Hockenmaier]

Communication (CwC Project, 2017-2019)

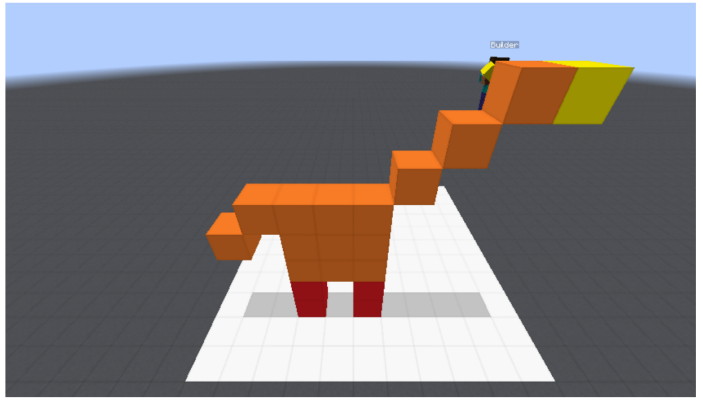
- Consider the following explanation:
  - **[Example from Yonatan Bisk]**
  - Imagine that this is a chess board
  - Place the UPS in H2, and McDonald in G6.

- Can you follow these instructions?
  - What's needed for us to be able to **write programs** that can do it?



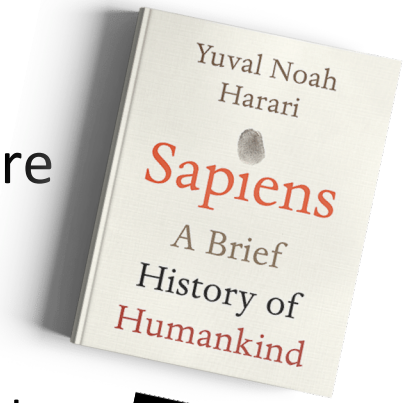
[I need to] move UPS from the left side of the board to just below Starbucks, leaving a small gap.

THE TARGET CONFIGURATION

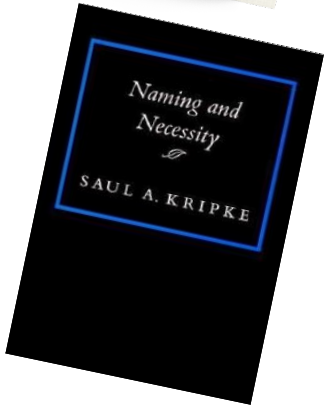


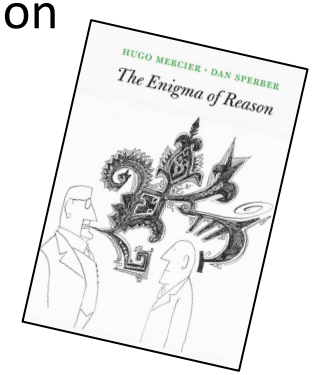
Symbolic Reasoning?

- Some people think that **symbols** is an evil invention of old AI people.
- It's not.
- Language is a symbolic system
  - Table is a collection of things in world.
- Even though we communicate via speech, gestures, writing, which are continuous, symbols are the invariants of this communication.



  - Harari: Language – the ability to assign symbols to "things" and "reason" about them is key to human cognitive revolution
  - Kripke: "Naming" things is key to communication and to cognition
  - The Enigma of Reason: "Reasoning is about giving reasons"

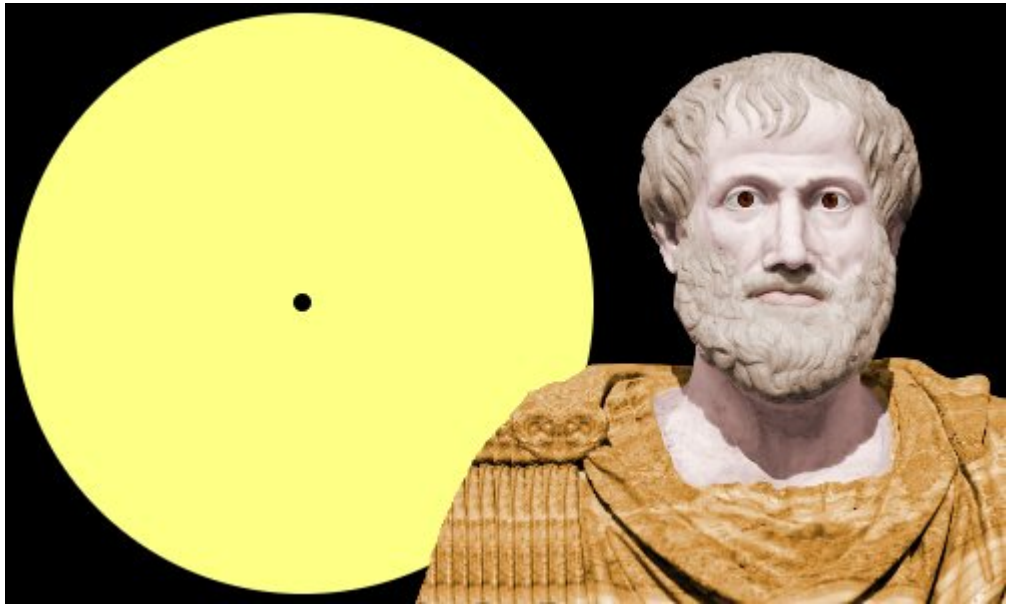




Aristotle's Logic https://www.famousscientists.org/aristotle/

- Aristotle founded the study of formal logic, systematizing logical arguments – he is famous for the syllogism, a method by which known information can be used to prove a point.

- Here is a famous example, from Aristotle himself, of a simple syllogism:
  - All men are mortal.
  - Socrates is a man.
  - Therefore, Socrates is mortal.

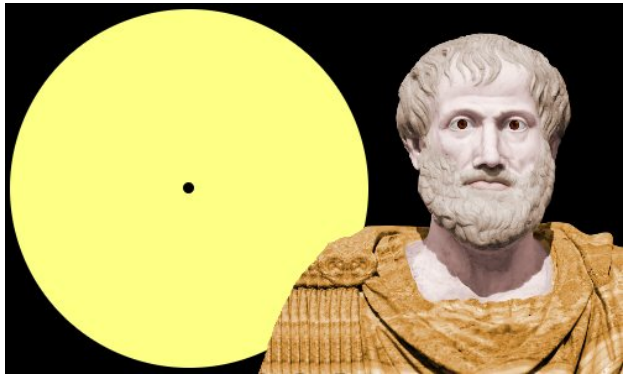


So, did Aristotle have a laptop?

[Example due to Les Valiant]

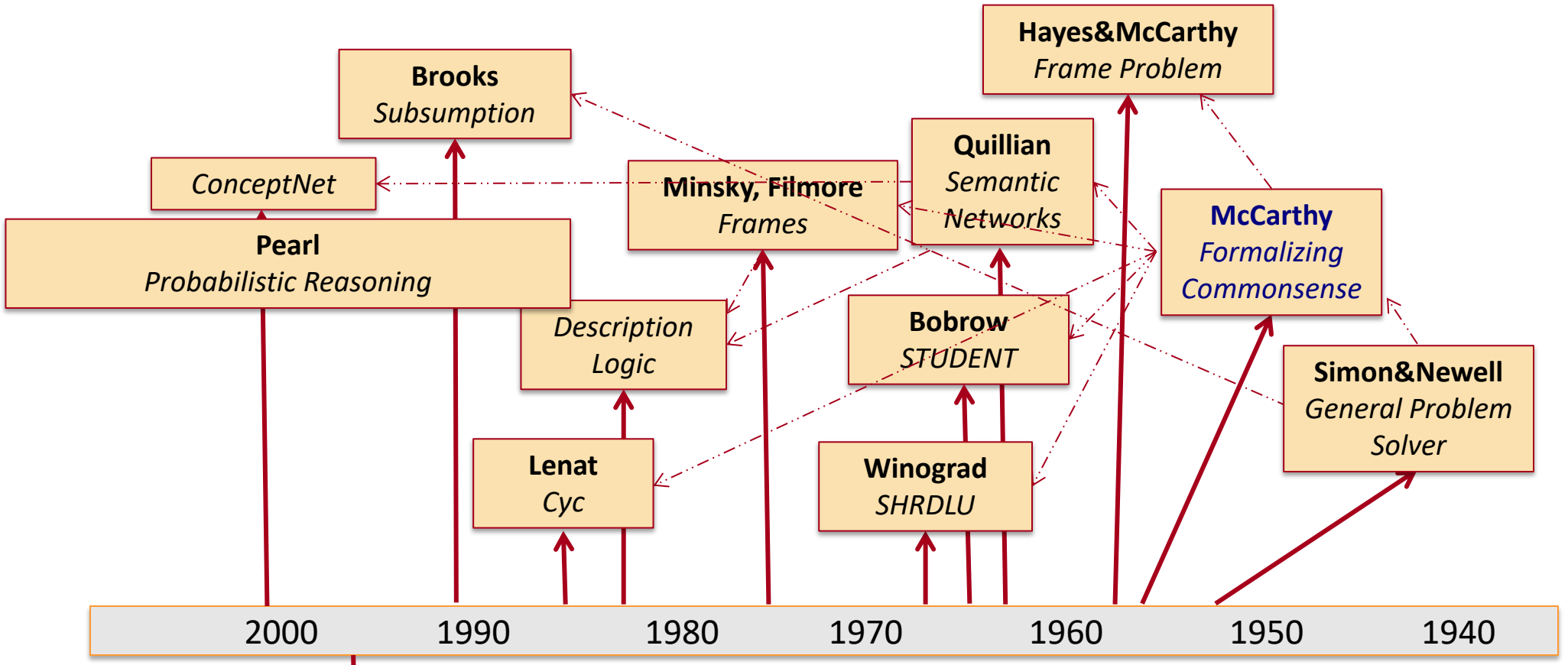
Key Challenges to NLU Researchers

- Today's NLP addresses well-defined tasks for which we have a lot of training data, but:
- Who asks the questions?
  - ~~ When did Aristotle die?
  - ~~ When was the laptop invented?
  - What happened first?



- The question isn't if we need reasoning, but rather:
  - How to do it?
    - E.g., how to determine "strategies" on the fly?
  - How to train for it?
    - How can we supervise for these "sparse" events? [A different talk: On Incidental Supervision]
    - Should we learn "everything" together? When/how should we decouple/decompose tasks?
- Of course, AI has thought about Reasoning a lot…
  - In most cases, in a way that decuples reasoning from learning
    - And often, commits to a representation that is independent of the task

A Biased View of Reasoning



Khardon & Roth
Learning to Reason

Common Sense Reasoning was formulated traditionally as a "reasoning" process, irrespective of learning and the resulting knowledge representation.

Learning to Reason I[Khardon & Roth JACM'96; 1994—2000]

- Proposed in reaction to the standard of complete decoupling of Learning and Reasoning
- A unifying computational theory of **Learning** and **Reasoning**.
  - Reasoning should be studied together with Learning and the representation it produces.
- Formally showing **the benefits in jointly studying Learning and Reasoning**
  - Some hard reasoning tasks become tractable if done on top of learning into an appropriate knowledge representation.
  - [Khardon & Roth JACM'96; 1994—2000]
- This is now called End-to-End….
- But, understanding **when** to decompose learning and **when** to decouple it from reasoning is also (still) very important.
  - At the heart of supporting transfer, compositionality, and reasoning.
- No better domain to think about this than Natural Language

A Few More Examples

Two Events

People were angry

Police used tear gas

Time

People **were angry** at something (which ended in violent conflicts with the police)…The police finally **used tear gas** (to restore order).

Two Events

Police used tear gas

People were angry

Time

Police **used tear gas**…People **were angry** at the police.

Two Events

Police used tear gas

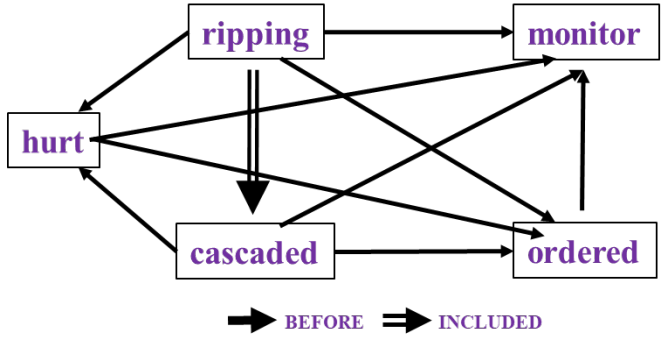
People were angry

Time

In natural language, we rarely see explicit **timestamps**, so we have to figure out the temporal order **from cues in the text**.

Time and Events [Ning et al. *SEM'2018; ACL'18, EMNLP'18, EMNLP'19]

- In Los Angeles that lesson was brought home today when tons of earth **cascaded** down a hillside, **ripping** two houses from their foundations. No one was **hurt**, but firefighters **ordered** the evacuation of nearby homes and said they'll **monitor** the shifting ground until March 23rd.



- Very difficult task— hinders exhaustive annotation ($O(N^2)$ edges)
- But, it's rather easy to get partial annotation – some relations.
- And, we have **strong expectations** from the output
  - Transitivity
  - Some events tend to precede others, or follow others

How does this knowledge impacts how we support these types of decisions?

Recap

Modified version of a question for AI2's DROP dataset

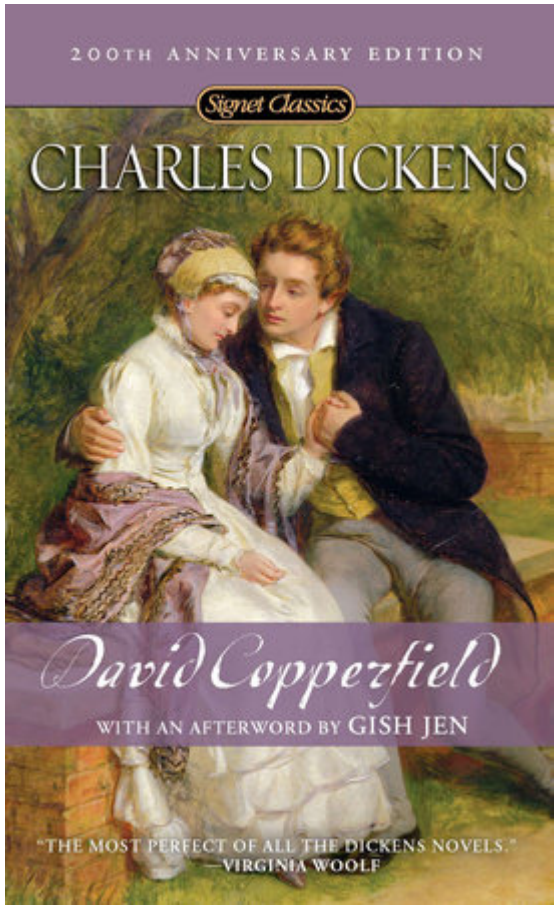
Coming off a road win over the Cowboys, the Redskins traveled to Lincoln Financial Field for a Week 5 NFC East duel with the Philadelphia Eagles. In the first quarter, the Redskins trailed early **as RB Brian Westbrook scored on a 9-yard TD run** and **the Eagles DeSean Jackson returned a punt 68 yards for a touchdown.** Washington still trailed at half time 14:9, **with field goals from Shaun Suisham**. In the third quarter, the Redskins took the lead on a trick play as WR Antwaan Randle El threw an 18-yard **TD pass to TE Chris Cooley**. In the fourth quarter, the Redskins increased their lead when **Clinton Portis scored on a 4-yard TD run**. The Eagles managed one more score in the final quarter for a **final score of 17:23**.

- What are the computational tasks that we should think about?

CliffsNotes

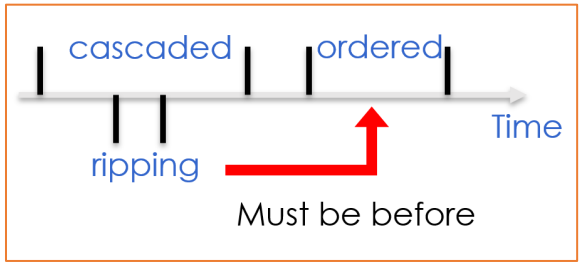
- Multiple natural language documents
  - Small units of text or large units of texts
  - Reading news about an event/situation **over time** and/or from **multiple sources**
  - **Reading a book**

  - The novel features the character David Copperfield, his journey of change and growth from infancy to maturity, as many people enter and leave his life and he passes through the stages of his development. (**Fiction, and you know it**)
  - London and England in the 19-th century; socio-economic state, child exploitation; schools, prisons, emigration to Australia (**true historical facts**)
- What are the computational tasks that we should think about?



What is Reasoning (today)?

- Assigning values to multiple interrelated variables
  - Given observations and knowledge
  - Most likely assignment: Abduction
  - Example: a graph



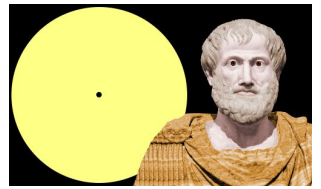
- Computing a function over multiple interrelated variables
  - Last; longest, most frequent, etc.

Who scored the last field goal for the Eagles? [given a recap of the game]

- Decisions over a dynamic set of interrelated variables
  - The Aristotle example
    - Need a strategy to decide what the variables are & a function computed over it
  - MC-Taco [EMNLP'19]

S1: *Growing up on a farm near St. Paul, L. Mark Bailey didn't dream of becoming a judge.*
Q1: *Is Mark still on the farm now?*

Will we make it to dinner before the movie?



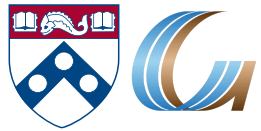
This Class

- I'd like to address all the questions brought up in the previous discussion, and more.
  - What is Reasoning?
  - What reasoning is needed to advance NLU?
  - What levels of abstractions are needed (if at all)?
  - If we need to abstract to "formal theories" – do we care about neural representations of these?
    - Actually, this is an old question…
  - What NLP tasks should we look at to advance our abilities in these directions?

- We will read a lot of papers
  - Feel free to propose additional papers that you think should be on our reading list.

This class

- We will open the reading list for bids in a few days
  - You will choose two papers to present and 4 to be a discussant on
- We will provide a small set of papers and ask that you choose one to reproduce
- Please feel free to suggest additional papers to the reading list.



- Understand early and current work on Reasoning
  - (Learn to) read critically, present, and discuss papers
- Understand some of the difficulties in NLU from the perspective of reasoning
  - Conceptual and technical
- Try some new ideas

- How:
  - Presenting/discussing papers
    - Probably: 2 presentations each; 4 discussants
  - Writing a few critical reviews
  - "Small" individual project (reproducing);
  - Large project (pairs)
  - Tentative details are on the web site.

- Machine Learning
  - 519/419
  - 520
  - Other?
- NLP
  - Yoav Goldberg's book
  - Jurafsky and Martin
  - Jacob Eisenstein

- Attendance is mandatory
- Participation is mandatory

- **Time of class?**

Questions?