A maybe not yet unified theory of inference for text understanding

Christopher Manning

Stanford University

@chrmanning / @stanfordnlp

Workshop on Uphill Battles in Language Processing

An earlier era of work had lofty goals, but modest realities

Today, we have *much* better realities, but usually fail to enunciate lofty goals

We should start enunciating lofty goals, or others will do it for us

People out to "solve" language

"The role of FAIR is to advance the science and the technology of AI and do experiments that demonstrate that technology for new applications like computer vision, dialogue systems, virtual assistants, speech recognition, natural language understanding, translation, things like that."

Yann LeCun

http://www.businessinsider.com/interview-yann-lecun-facebook-ai-deepmind-2016-10

Norvig (1986) Ph.D.

Peter Norvig's thesis – almost 30th anniversary

A Unified Theory of Inference for Text Understanding

By

Peter Norvig

B.S. (Brown University) 1978

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

COMPUTER SCIENCE

in the

GRADUATE DIVISION

OF THE

UNIVERSITY OF CALIFORNIA, BERKELEY

Robert Wilensky
Lofti Zadeh
Chuck Fillmore

Approved: 11/24/86
Chairman   Date
L. A. Zadeh   11/25/86
Charles J. Fillmore   11/25/86

The language analyzed

In a poor fishing village built on an island not far from the coast of China, a young boy named Chang Lee lived with his widowed mother. Every day, little Chang bravely set off with his net, hoping to catch a few fish from the sea, which they could sell and have a little money to buy bread.

(a) There is a sea, which surrounds the island, is used by the villagers for fishing, and forms part of the coast of China

(b) Chang intends to trap fish in his net, which is a fishing net

(c) The word *which* refers to *the fish*

(d) The word *they* refers to Chang and his mother

The unified theory of inference

"As we have just seen, a suitable knowledge base is a prerequisite for making proper inferences" (p. 4). It's built to enable inferences

System had 6 general forms of inference; 2 pairs, so 4 basic types:

1. Elaboration: Filling a slot to connect two entities
   - John got piggybank for REASON have money for REASON buy present
2. Reference Resolution: Hey – it's coreference!!!
3. View Application: *The Red Sox killed the Yankees*
   - KILLED is not animal; KILLING is viewed as a DEFEAT-CONVINCINGLY
4. Concretization: Infer more specific
   - TRAVELLING in an AUTOMOBILE is an instance of DRIVING

Basic NLP: Progress has been made!

"Arens and Wilensky's PHRAN program was used where possible [to convert input sentences to KODIAK knowledge representations]. For some input, PHRAN was not up to the task, so a representation was constructed by hand instead." (p. 4)
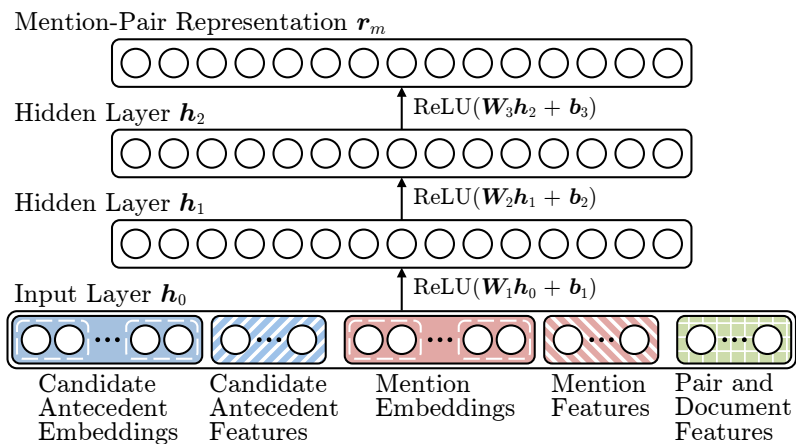
CoreNLP
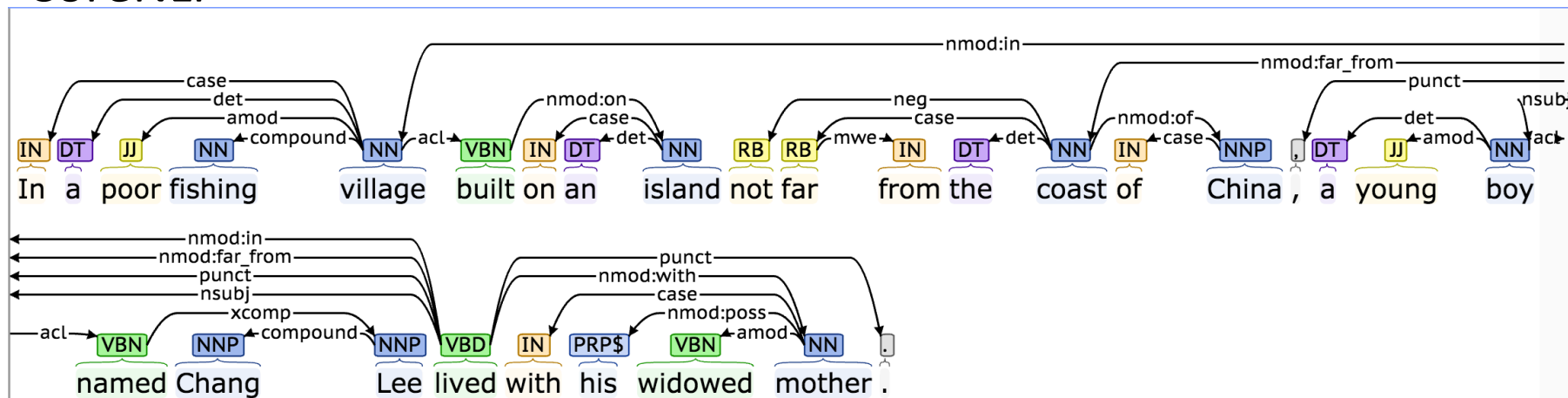
2. Reference Resolution

[Clark & Manning ACL 2016, EMNLP 2016]

Neural networks improve mention semantic compatibility scores

Model	Eng	Chi
Chen & Ng (2012)	54.52	**57.63**
Fernandes (2012)	**60.65**	51.46
Stanford DeepRL (2016)	**65.73**	**63.88**

Imitation Learning or Reinforcement Learning permits improved coreference decision making
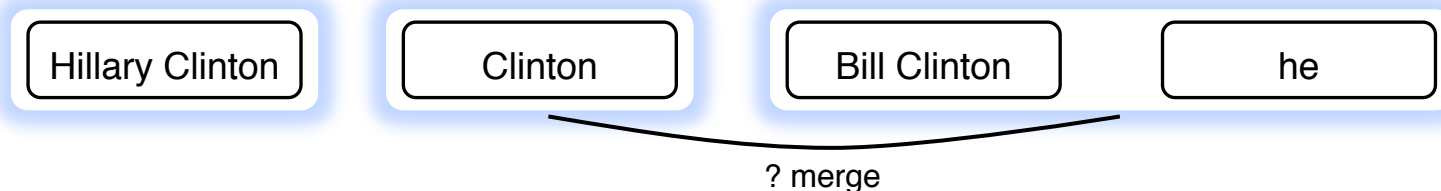
Building elaborations a la Norvig (1986)
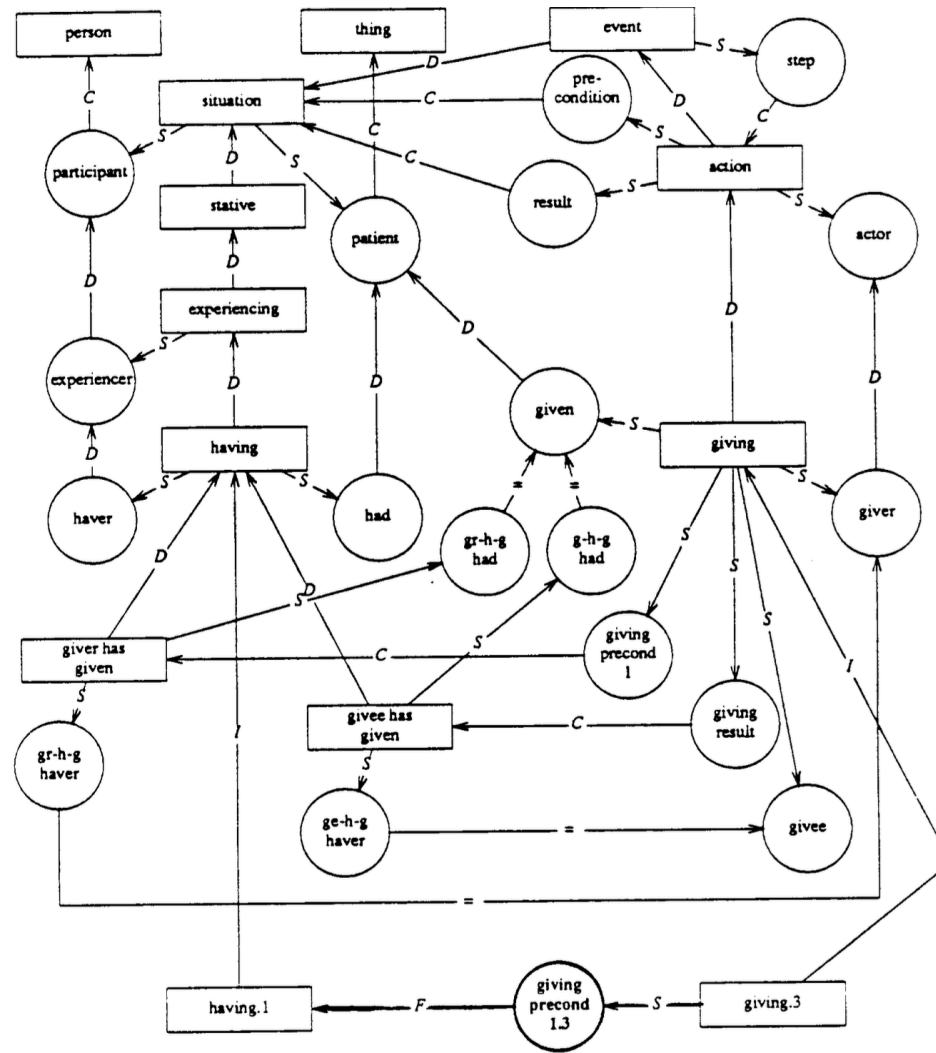
Understanding event relationships
(Berant et al. 2014)

Enable

Cause

Water is split, providing a source of electrons and protons (hydrogen ions, H^+) and giving off O_2 as a by-product. ***Light absorbed*** by chlorophyll drives a ***transfer of the electrons and hydrogen ions*** from water to an acceptor called $NADP^+$.
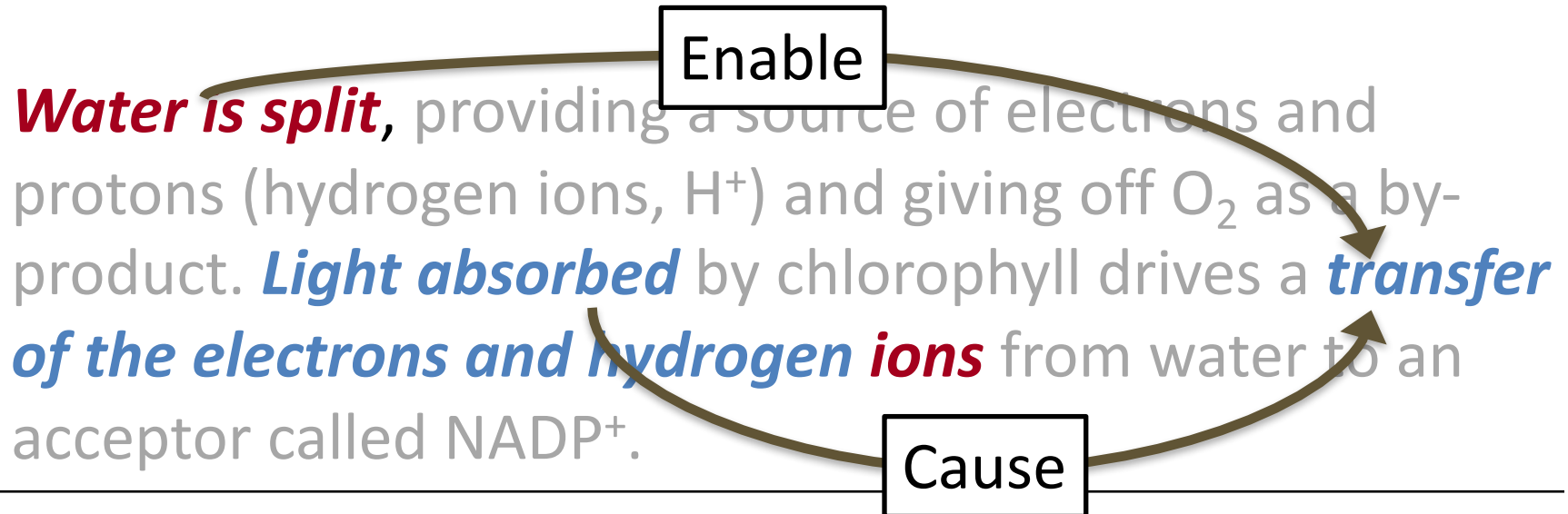
What can the splitting of water lead to?
A: Light absorption
B: Transfer of ions

Natural logic inference … using text as the meaning representation

(Angeli and Manning 2016)

Machine Comprehension of Language
[Chen, Bolton, & Manning, ACL 2016]

Bidirectional RNNs

Q — characters in " @placeholder " movies have gradually become more diverse

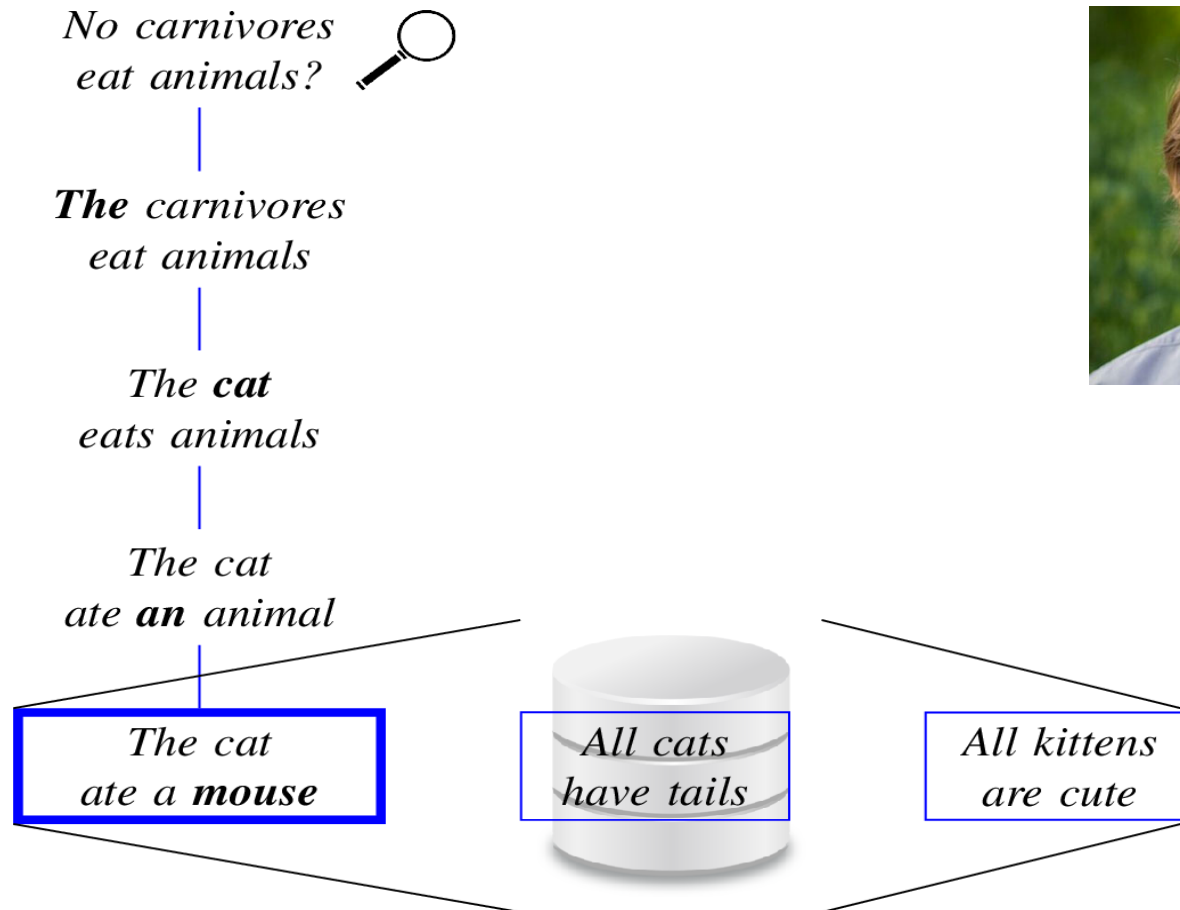**q**

P →   …   …   … **p̃ᵢ**

(@entity4) if you feel a ripple in the force today , it may be the news that the official @entity6 is getting its first gay character . according to the sci-fi website @entity9 , the upcoming novel " @entity11 " will feature a capable but flawed @entity13 official named @entity14 who " also happens to be a lesbian . the character is the first gay figure in the official @entity20 the movies , television shows , comics and books approved by @entity6 franchise owner @entity22 -- according to @entity24 , editor of " @entity6 "

Attention

A — entity6

$$\alpha_i \;=\; \boxed{\operatorname*{softmax}_i \left(\mathbf{q}^\top \mathbf{W}_s \tilde{\mathbf{p}}_i \right)} \qquad\qquad \mathbf{o} \;=\; \sum_i \alpha_i \tilde{\mathbf{p}}_i$$

What do we still need?

BiLSTMs with attention seem to be taking over the field and improving our ability to do **everything**

Neural methods are leading to a **renaissance** for all language generation tasks

There's a real scientific question of where and whether we need explicit, localist language and knowledge representations and inferential mechanisms

What do we still need?

However:

We still have very primitive methods for building and accessing **memories** or **knowledge**

Current models have almost nothing for developing and executing **goals** and **plans**

We still have quite inadequate abilities for understanding and using **inter-sentential relationships**

We still can't at a large scale do **elaborations** from a **situation** using **common sense knowledge**