

Watson and Jeopardy

Lecture 23: November 27, 2013

CS886-2 Natural Language Understanding
University of Waterloo

Watson at Jeopardy



Jeopardy

- Host reads a **clue in the form of an answer**
 - But it is really a question
- Contestants **respond with a question**
 - But it is really an answer
- **Clue:** When hit by electrons, a phosphor gives off electromagnetic energy in this form.
 - What form of electromagnetic energy does a phosphor give when hit by electrons?
- **Response:** What is a photon?
 - Photon

CS886 Lecture Slides (c) 2013 P. Poupart

3

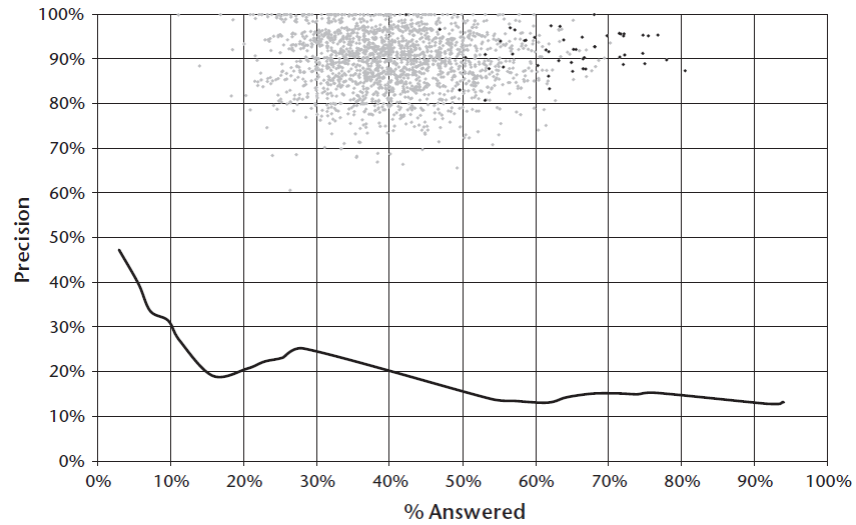
QA Systems in 2007

- Designed for TREC (not Jeopardy)
- Two state of the art QA systems
 - IBM: **PIQUANT** (Practical Intelligent QUestion ANswering Technology)
 - CMU: **OpenEphyra** (Open source QA framework)

CS886 Lecture Slides (c) 2013 P. Poupart

4

PIQUANT vs Jeopardy Champions



CS886 Lecture Slides (c) 2013 P. Poupart

5

Jeopardy vs TREC

Jeopardy

- No specific corpus
- No internet access
- 1-6 seconds per question
- Complex questions
- Confidence is critical

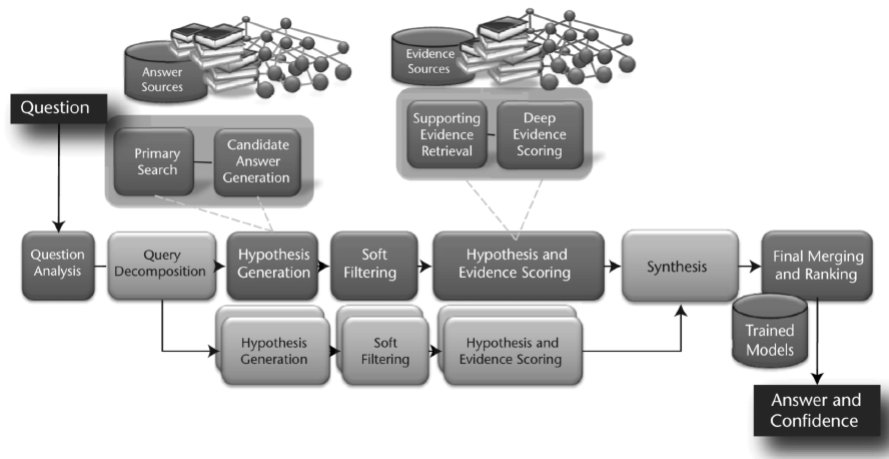
TREC

- Corpus: 1 million docs
- Internet access
- 1 week: answer 500 quest.
- Simple questions
- Confidence not measured

CS886 Lecture Slides (c) 2013 P. Poupart

6

DeepQA Architecture



CS886 Lecture Slides (c) 2013 P. Poupart

7

Key Aspects

1. **Ensemble framework**
 - Multiple techniques for each component
 - Combine/rank hypotheses produced by each technique
2. **Pervasive confidence measures**
 - All algorithms produce a hypothesis and a score

CS886 Lecture Slides (c) 2013 P. Poupart

8

Content Acquisition

- No set corpus and no internet access
- Acquisition of relevant content
 - Manual and automated steps
 - encyclopedias, dictionaries, thesauri, newswire articles, literary works
 - Freebase, WordNet, DBPedia, etc.
 - Passages of some web pages

CS886 Lecture Slides (c) 2013 P. Poupart

9

Question Analysis

- Compute shallow parses, deep parses, logical forms, semantic role labels, coreference, relations, named entities
- Question Classification:
 - puzzle question, math question, definition question, named entity, lexical answer type detection

CS886 Lecture Slides (c) 2013 P. Poupart

10

Lexical Answer Type

- When hit by electrons, a phosphor gives off electromagnetic energy in this form.
 - Answer type:
- This title character was the crusty and tough city editor of the *Los Angeles Tribune*
 - Answer type:

CS886 Lecture Slides (c) 2013 P. Poupart

11

Hypothesis Generation

- Generate candidate hypotheses from content sources
- text search engines with different approaches
- document search as well as passage search
- knowledge base search
- named entity recognition
- **Focus on recall:** generate lots of possible hypotheses

CS886 Lecture Slides (c) 2013 P. Poupart

12

Hypothesis scoring

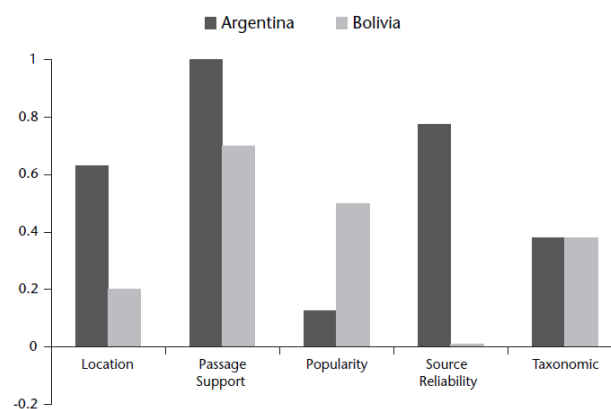
- **Focus on precision:** filter and rank hypotheses
- Many scoring techniques to verify different dimensions
 - Taxonomic, Geospatial (location), Temporal, Source Reliability, Gender, Name Consistency, Relational, Passage Support, Theory Consistency

CS886 Lecture Slides (c) 2013 P. Poupart

13

Example

- Chile shares its longest land border with this country.



CS886 Lecture Slides (c) 2013 P. Poupart

14

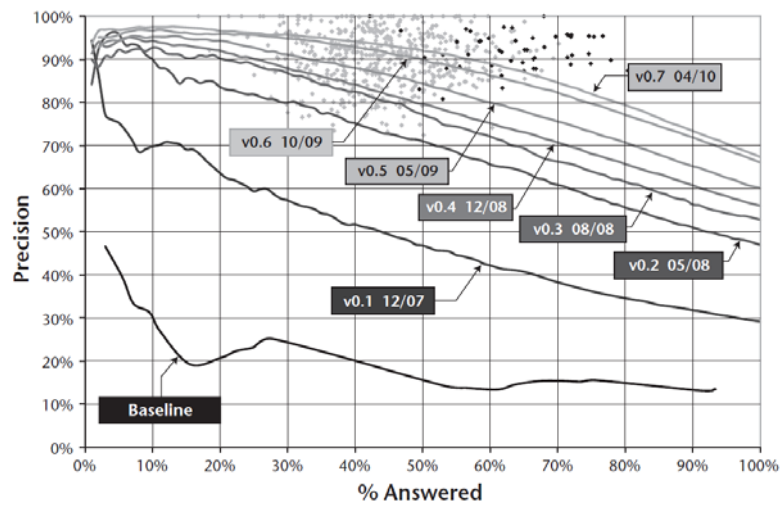
Ranking

- Combine scores to rank hypotheses
 - Supervised learning
 - Ensemble and hierarchical models

CS886 Lecture Slides (c) 2013 P. Poupart

15

Improvements



CS886 Lecture Slides (c) 2013 P. Poupart

16