

Reading Comprehension: Reading, Learning, Teaching

Lynette Hirschman

781-271-7789 • lynette@mitre.org

MITRE Sponsored Research



Problem

- What would it take to build a system that
 - *reads* for understanding,
 - *learns* as it reads, and
 - *teaches* and is taught while reading with a companion?
- How can we test this comprehension skill?
 - Should we use human evaluations, such as graded reading comprehension tests?
 - What other testing options are available or emerging?

Background

Reading comprehension is a “Grand Challenge” for human language technology.

Library of Congress Has Books for Everyone

(WASHINGTON, D.C., 1964) - It was 150 years ago this year that our nation's biggest library burned to the ground. Copies of all the written books of the time were kept in the Library of Congress. But they were destroyed by fire in 1814 during a war with the British. That fire didn't stop book lovers. The next year, they began to rebuild the library. By giving it 6,457 of his books, Thomas Jefferson helped get it started.

1. Who gave books to the new library?
2. What is the name of our national library?
3. When did this library burn down?

Why are some questions harder than others?

What do we need to know to answer these questions?

easy

hard

“burn down” =
“burn to the ground”

“our national library” =
“our nation’s biggest library” =
“the Library of Congress”

Objective

- Establish an ambitious research agenda to build and test computer systems that *read* for understanding, *learn* as they read, and *teach* and are taught while reading
- Explore the consequences of using graded human reading comprehension tests to evaluate these systems
- Pursue other promising evaluation paradigms

FY05: participate in the Recognizing Textual Entailment (RTE) Challenge and apply lessons learned to the reading comprehension problem

Activities

■ FY01–FY04:

- ABCs: Abduction-based architecture for reading comprehension test taking
- Rasch analysis of TREC question answering (QA) evaluations

■ FY05:

- Built two systems to participate in RTE Challenge
- Leveraging structured and unstructured corpora to enhance semantic knowledge
 - RTE, TREC QA, comprehension tests

Highlight

1989 Remedia Publications, Comprehension/5W

Owens Wins 4th Gold Medal!


(MUNICH, GERMANY, 1936) - Jesse Owens is the fastest man in the world. He runs and wins for the United States team. Today, Jesse won his fourth gold medal in the Olympics.


Each time he wins a medal, it is placed around his neck. A band plays the "Star-Spangled Banner."


This is not the first time Jesse has shown his speed. A few years ago, he broke five world records. In fact, he did this in one hour.

Jesse's future was not always bright. As a


Dialog:


 Click on an action to work with this story.

 Please ask a question about the story.

 Where did Jesse win his medal?

 --> Is [the, Olympics] an instance of a location?

 yes

 --> Is [the, Olympics] the correct answer?

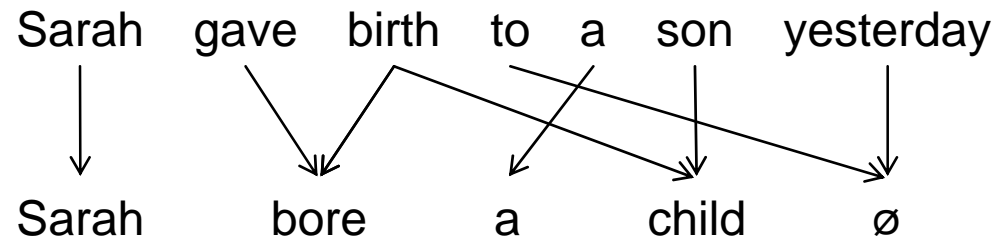
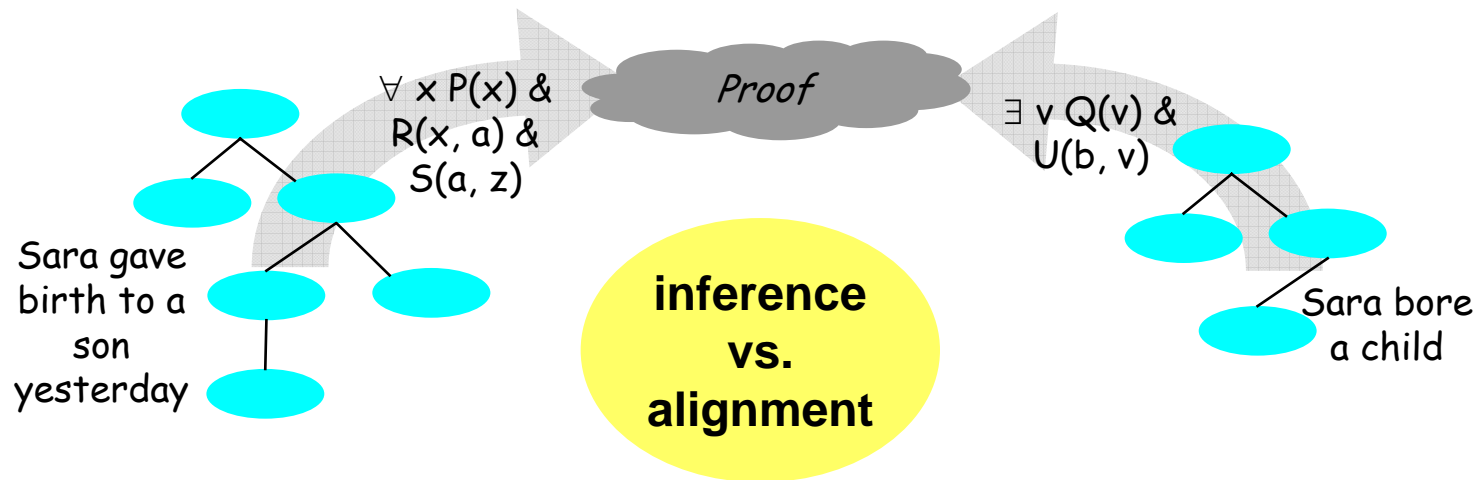
Choose an Action:

Your input:

yes

Highlight

Recognizing Textual Entailment (RTE)



Impacts

- **Extended the state of the art in using reading comprehension tests as a grand challenge for natural language processing**
 - **Other groups evaluating their systems using our data, including U Edinburgh**
 - **DARPA interested in using reading comprehension as a grand challenge**
- **Developed methodology for decomposition of reading comprehension system performance**
- **Exploited psychometric/educational testing methodology for assessing test difficulty in TREC QA**

Future Plans

2. Clay soil forms a fairly effective barrier against the movements of water. It also swells and shrinks significantly as its water content changes. Sandy soil, in contrast, allows water to move freely and does not change shape as the water content varies.

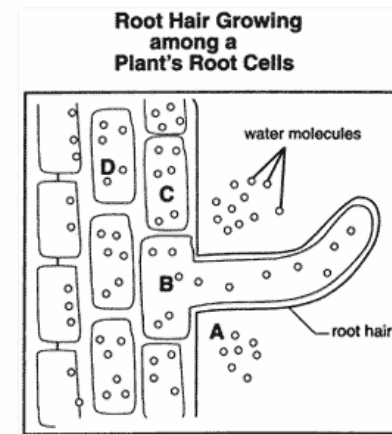
In which statement is the appropriate soil selected for its intended site?

- 1. Sandy soil would make a good lining for a toxic waste site.
- 2. Clay soil would work well in a drain field.
- 3. Clay soil would be a good foundation for a large building.
- 4. Clay soil would form a good liner if a person built a pond.
- 5. A sandy lake bottom would prevent water from seeping out of the lake.

3. The root hair of a plant, shown in the diagram below, is the most efficient way for the plant to absorb water from surrounding soil.

DARPA IPTO Candidate Grand Challenge

Build a system that will pass an appropriate standardized test (GED, SAT, NY Regents)



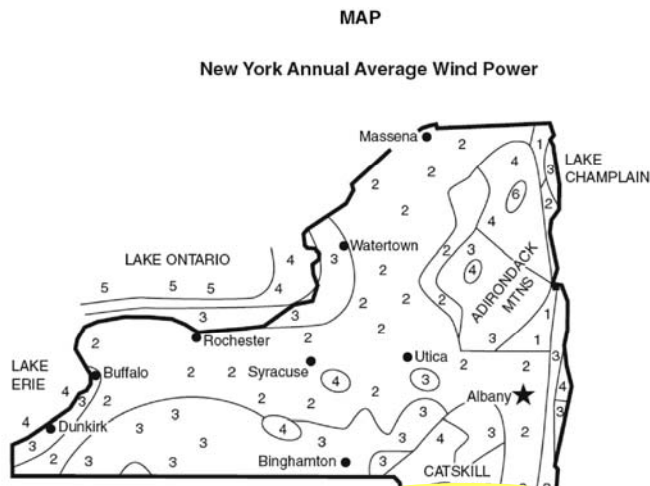
At what point is the flow of water the GREATEST?

- 1. C to B

MITRE

© 2005, The MITRE Corporation

Future Plans



DARPA IPTO Candidate Grand Challenge

Build a system which will pass an appropriate standardized test (e.g., NY Regents), integrating visual and textual understanding

14 According to the map, the location where the highest level of wind power in New York State occurs is

- (1) in the Catskill Mountains
- (2) in the Adirondack Mountains
- (3) around Syracuse
- (4) around Buffalo

15 The curved lines on the map are used to

- (1) define areas of similar wind power
- (2) indicate sites of conventional power plants
- (3) indicate the direction of prevailing winds
- (4) designate areas of power shortages

16 Wind power is *least* likely in the area that is immediately

- (1) east of Binghamton
- (2) south of Massena
- (3) west of Watertown
- (4) north of New York City