From Textual Entailment to Knowledgeable Machines

Peter Clark
Allen Institute for Artificial Intelligence (AI2)
Mission: achieve scientific breakthroughs by constructing AI systems with reasoning, learning, and reading capabilities.
Overall Goals

- **Vision: The Digital Aristotle**
  - Large volumes of general and scientific knowledge, stored in a "computable" form that supports reasoning and explanation.

  "Explainable Reasoning"

- **Intermediate Focus:**
  - Pass *science exams* as written
  - Heavy emphasis on semi-automated knowledge acquisition
  - Textual entailment at every step
    - ≈ the “modus ponens” of reasoning
The Task

- Current focus:
  - 4th grade, multiple choice science questions
  - Wide variety of question types
  - Requires general, lexical, and scientific knowledge
The 4th Grade NY Regents’ Science Exam

- What types of questions are there?
- What would it take to answer them?
The 4th Grade NY Regents’ Science Exam

- What types of questions are there?
- What would it take to answer them?
Multiple Choice to Textual Entailment

12 Which food is a fruit?
   A  a potato
   B  an onion
   C  a carrot
   D  a pumpkin

1. Convert to 4 true/false questions

   $Q_A$: A potato is a fruit?
   $Q_B$: An onion is a fruit?
   $Q_C$: A carrot is a fruit?
   $Q_D$: A pumpkin is a fruit?

2. Convert each true/false question to an entailment problem

   $T$: potato $\text{ entails? }$ $H$: potato is a fruit?

   OR: (for $X$ isa $Y$ questions, and questions with a setup)

   $T$: potato $\text{ entails? }$ $H$: fruit?
Multiple Choice to Textual Entailment

12 Which food is a fruit?
   A  a potato
   B  an onion
   C  a carrot
   D  a pumpkin

1. Convert to 4 true/false questions

<table>
<thead>
<tr>
<th>Question</th>
<th>Confidence</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q_A: A potato is a fruit?</td>
<td>0.04</td>
</tr>
<tr>
<td>Q_B: An onion is a fruit?</td>
<td>0.12</td>
</tr>
<tr>
<td>Q_C: A carrot is a fruit?</td>
<td>0.21</td>
</tr>
<tr>
<td>Q_D: A pumpkin is a fruit?</td>
<td>0.64</td>
</tr>
</tbody>
</table>

Answer is D

2. Convert each true/false question to an entailment problem

T: potato entails? H: potato is a fruit?

OR: (for X isa Y questions, and questions with a setup)
T: potato entails? H: fruit?

3. Pick highest confidence answer
The 4th Grade NY Regents’ Science Exam

- What types of questions are there?
- What would it take to answer them?

“Basic”
- Taxonomies
- Dictionaries
- Fact databases
- Properties of an object
- Examples of situations
- Example database, rules
- Causal rules
- Causality
- Processes
- Process descriptions
- Specific domain representations
  - Energy Transfer Life Cycle
- Specific domain models
- Diagrams
- Definitions
1. Taxonomic

3 Sleet and hail are forms of
   A clouds
   B energy
   C evaporation
   D precipitation

- Simple **lexical entailment**
  - e.g., T: “sleet” entails? H: “precipitation”
  - Several good sources of simple “isa” knowledge
    - WordNet, Cyc, Wikipedia
  - Is a basic operation for more complex entailment tasks
2. Definitions

4. The movement of soil by wind or water is called

A. condensation
B. evaporation
C. erosion
D. friction

- Search for best entailing definition

Dictionary Resources

erosion: The process of being eroded by wind, water, or other natural agents.
erosion: The wearing away of rocks and other deposits on the earth's surface ...
erosion: The gradual wearing away of land surface materials, especially rocks, ...
2. Definitions

4. The movement of soil by wind or water is called

A. condensation
B. evaporation
C. erosion
D. friction

- Search for best entailing definition

**Dictionary Resources**

erosion: The process of being eroded by wind, water, or other natural agents.
erosion: The wearing away of rocks and other deposits on the earth's surface ...
erosion: The gradual wearing away of land surface materials, especially rocks, ...

T: The *gradual wearing away of land surface materials*, especially rocks, sediments, and soils, *by the action of water, wind*, or a glacier

H: The *movement of soil by wind or water*
T: The *gradual wearing away of land surface materials*, especially rocks, sediments, and soils, *by the action of water, wind*, or a glacier

entails?

H: The *movement of soil by wind or water*

| # words in common | weighted #words | # with hypernyms | ∑ bi-grams | paras-phrases | ... | p(H|T) |
|-------------------|-----------------|------------------|------------|--------------|-----|-------|
| 3.0               | 4.31            | 3.45             | 2.0        | 1.20         |     | 0.76  |

**TRAINING DATA**

|                  | weighted #words | # with hypernyms | ∑ bi-grams | paras-phrases | p(H|T) |
|------------------|-----------------|------------------|------------|--------------|-------|
| H TRUE           | 1.23            | 2.12             | 5.0        | 1.98         | 0.20  |
| H FALSE          | 5.43            | 1.11             | 1.0        | 0.24         | 0.24  |
| H TRUE           | 1.12            | 3.45             | 2.0        | 1.76         | 0.76  |
2. Definitions

- The movement of soil by wind or water is called:
  - A. condensation
  - B. evaporation
  - C. erosion
  - D. friction

Answer is erosion (C)

- Search for best entailing definition

Dictionary Resources

- erosion: The process of being eroded by wind, water, or other agents. 0.54
- erosion: The wearing away of rocks and other deposits on the earth. 0.43
- erosion: The gradual wearing away of land surface materials, ... 0.76

- friction: The rubbing of surfaces against each other 0.21
- friction: a resistance encountered when one body moves relative to another body with which it is in contact 0.11
- friction: surface resistance to relative motion, as of a body sliding 0.13
3. within-question entailments

1. Which example describes an organism taking in nutrients?
   A. a dog burying a bone
   B. a girl eating an apple
   C. an insect crawling on a leaf
   D. a boy planting tomatoes in a garden

Q_B: A girl eating an apple *is an example of* an organism taking in nutrients?

T: A girl eating an apple

entails?

H: an organism taking in nutrients
3. within-question entailments

1. Which example describes an organism taking in nutrients?
   A. a dog burying a bone
   B. a girl eating an apple
   C. an insect crawling on a leaf
   D. a boy planting tomatoes in a garden

Q_B: A girl eating an apple is an example of an organism taking in nutrients?

T: A girl eating an apple

H: an organism taking in nutrients

\[ p(H|T) = 0.91 \]
The 4th Grade NY Regents’ Science Exam

- What types of questions are there?
- What would it take to answer them?

"Entailment from Corpus"
4. Entailment from a corpus

24 Which part of a plant produces the seeds?

A flower
B leaves
C stem
D roots

$H_A$: The flower is the part of a plant that produces the seeds
$H_B$: The leaves are the part of a plant that produces the seeds
$H_C$: The stem is the part of a plant that produces the seeds
$H_D$: The roots are the part of a plant that produces the seeds

T: 

$H_A$? $H_B$? $H_C$? $H_D$?
4. Entailment from a corpus

**24 Which part of a plant produces the seeds?**

A. flower
B. leaves
C. stem
D. roots

Hₐ: The **flower** is the part of a **plant** that produces the seeds
Hₐ: The leaves are the part of a plant that produces the seeds
Hₚ: The stem is the part of a plant that produces the seeds
Hₚ: The roots are the part of a plant that produces the seeds

*Entailment Confidence?*

Plants can grow from a seed into a flower, tree, or bush. **Plants** reproduce by producing **flowers** and fruits that have **seeds**. The seeds then grow into plants.
The 4th Grade NY Regents’ Science Exam

- What types of questions are there?
- What would it take to answer them? 

“Diagrams”

“Models”
When a baby shakes a rattle, it makes a noise. Which form of energy was changed to sound energy?

A. electrical
B. light
C. mechanical
D. heat

Requires a specific computation over a representation

baby shake rattle \rightarrow \text{rattle make noise} \rightarrow \text{movement} \rightarrow \text{mechanical energy} \rightarrow \text{sound} \rightarrow \text{sound energy}

(C) Mechanical Energy
Performance (excluding diagrams)

- Works okay… ~55% score (vs. 25% random guessing)

<table>
<thead>
<tr>
<th>Date</th>
<th>Baseline</th>
<th>2013.07.29</th>
<th>2013.08.02</th>
<th>2013.09.04</th>
<th>2013.10.31</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run nickname (if any)</td>
<td>Baseline</td>
<td>SOLR_V1</td>
<td>SOLR_V2</td>
<td>SOLR_V3</td>
<td>Monthly</td>
</tr>
<tr>
<td>MC, Non-Diagrams Only</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TRAINING SET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th grade</td>
<td>45%</td>
<td>44%</td>
<td>55%</td>
<td>54%</td>
<td>57%</td>
</tr>
<tr>
<td>8th grade</td>
<td>45%</td>
<td>48%</td>
<td>55%</td>
<td>58%</td>
<td>56%</td>
</tr>
<tr>
<td>12th grade</td>
<td>37%</td>
<td>28%</td>
<td>26%</td>
<td>28%</td>
<td>30%</td>
</tr>
<tr>
<td>AP Biology</td>
<td>42%</td>
<td>48%</td>
<td>46%</td>
<td>49%</td>
<td>47%</td>
</tr>
<tr>
<td>TEST SET</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4th grade</td>
<td>47%</td>
<td>53%</td>
<td>60%</td>
<td>57%</td>
<td>56%</td>
</tr>
<tr>
<td>8th grade</td>
<td>36%</td>
<td>55%</td>
<td>52%</td>
<td>54%</td>
<td>51%</td>
</tr>
<tr>
<td>12th grade</td>
<td>43%</td>
<td>49%</td>
<td>49%</td>
<td>50%</td>
<td>43%</td>
</tr>
<tr>
<td>AP Biology</td>
<td>38%</td>
<td>49%</td>
<td>44%</td>
<td>48%</td>
<td>49%</td>
</tr>
</tbody>
</table>
29. What is one way that humans can maintain good health?

- A. sleep all day
- B. exercise daily
- C. eat snacks often
- D. stay up late

System answer: (B) 😊

If horses are kept inside in a barn, they require regular daily exercise for their physical health and mental well-being.

12. Which food is a fruit?

- A. a potato
- B. an onion
- C. a carrot
- D. a pumpkin

System answer: (C) 😞

Carrots can also be used alone or with fruits in jam and preserves.
1. If an object is attracted to a magnet, the object is most likely made of:

A. wood
B. plastic
C. cardboard
D. metal

System answer: (D) 😊

In experiments in which statoliths were replaced with metal shavings, researchers "tricked" crayfish into swimming upside down by using magnets to pull the shavings to the upper end of the statocysts located at the base of their antennae.

27. Fourth graders are planning a roller-skate race. Which surface would be the best for this race?

A. gravel
B. sand
C. blacktop
D. grass

System answer: (A) 😞

Graders are commonly used in the construction and maintenance of dirt roads and gravel roads.
What is going on?

- Largely "smart guessing" based on word associations
- BUT:
  - doesn’t give us meaningful explanations
  - doesn’t get us closer to machine reading
  - performance max’es out at ~60%

- What is missing?
  - Richer representation of meaning
    - in the question
    - in the corpus
Entailment

Text

"textual" entailment

Query

?- has-part(ribosome, ?x).

Logic

logical entailment
Entailment

Text

Logic

Query

"textual" entailment

?- has-part(ribosome,?x).
Textual Entailment = The Science of Semi-Formal Representations
Semi-formal representations

- **Lexical** – good baseline
- **Dependency trees** – precise but complex to manipulate
- **Full logic** – very hard to translate into
- **SVO units ("depth 1 parses", "tuples")?**
  - top-level: syntactic structure
  - lower level: phrasal

Text: The vibrations from sound move tiny bones in our ears.

Repn. (the vibrations from sound, move, tiny bones in our ears)

subject   verb   object [pps]

- These simple propositions = a basic “unit of meaning”
- Inference is a mixture of structural and phrasal matching
Sentences may contain one or more related units

Some animals grow thick fur in winter to stay warm.

(some animals, grow, thick fur, in winter) \textbf{EFFECT} (some animals, stay, , warm)

Types of relations between units:

- P1 \textbf{AND} P2
- P1 \textbf{IMPLIES} P2
- P1 \textbf{EFFECT} P2
- P1 \textbf{CAUSES} P2
- P1 \textbf{PURPOSE} P2
- Can define extraction patterns for these units

  S V O “to help” V O \[\rightarrow\] P1 EFFECT P2

  Some animals grow thick fur in winter to help maintain body heat

  (some animals, grow, thick fur, in winter) EFFECT (some animals, maintain, body heat)
Pattern-based Extraction

- Can transduce text into this form

\[\begin{align*}
S &\ V &\ O \ "to\ help" \ S &\ V &\ O \\
S &\ V &\ O \ "in\ order\ to" \ V &\ O \\
S &\ V \ "cause" \ S &\ V &\ O \\
\ldots &\ \ldots &\ \ldots \\
\end{align*}\]

\[\begin{align*}
P1 \\
P2 \ \text{AND} \ P3 \\
P4 \\
P5 \ \text{CAUSES} \ P6 \\
P7 \ \text{ENABLES} \ P8 \ \text{AND} \ P9 \\
P10 \\
P11 \\
P12 \ \text{IMPLIES} \ P13 \\
P14 \ \text{CAUSES} \ P15 \\
P16 \\
\end{align*}\]
Can transduce text into this form:

<table>
<thead>
<tr>
<th>CAUSE</th>
<th>EFFECT</th>
</tr>
</thead>
<tbody>
<tr>
<td>(gravity, pull)</td>
<td>(objects, fall, towards Earth)</td>
</tr>
<tr>
<td>(scientists, using, a model)</td>
<td>(scientists, understand, better)</td>
</tr>
<tr>
<td>(dividing, a single cell)</td>
<td>(form, two daughter cells)</td>
</tr>
<tr>
<td>(Animals, use, saturated fatty acids)</td>
<td>(store, energy)</td>
</tr>
<tr>
<td>(humans, get, regular rest)</td>
<td>(be, healthy)</td>
</tr>
<tr>
<td>(Fish, have, fins)</td>
<td>(move, )</td>
</tr>
<tr>
<td>(animals, move, to warmer climate)</td>
<td>(avoid, change in seasons)</td>
</tr>
<tr>
<td>(cactus, hold, water)</td>
<td>(survive, in the desert)</td>
</tr>
</tbody>
</table>
9 Which characteristic can a human offspring inherit?

A. facial scar
B. blue eyes
C. long hair
D. broken leg

System Answer: (A)
Some traits that can be inherited are color of hair, color of skin, color of eyes, and height.

System Answer: (B)

Blue eyes is a characteristic that a human offspring can inherit?

(human offspring, can inherit, the characteristic of blue eyes)

,, can inherit, color of hair
,, can inherit, color of skin
,, can inherit, color of eyes
,, can inherit, height

Some traits that can be inherited are color of hair, color of skin, color of eyes, and height.
Where does this break down?

1. Quality of extractions is low
2. Sentence-level units are too small, context independent

(create a diagram) **EFFECT** (show the cells of multicellular organisms may be organized at different levels)

(Say, thanks to the authors) **EFFECT** (access a customizable version of this book)
(All cells are small, very) **EFFECT** (one or more cells need)

- Need larger-sized structures
  - spanning several sentences
Textual entailment is less about “matching text” and more about the science of semi-formal representations. Those representations explicate the “world knowledge” the text encodes, can be sharable “knowledge resources” in their own right, and take us a step closer to “knowledgeable machines.”

Thank you!