Inference Rules for Recognizing Textual Entailment

Georgiana Dinu and Rui Wang
Computational Linguistics and Phonetics
Saarland University
{dinu,rwang}@coli.uni-sb.de

February 4, 2009
1 Background
   - DIRT  Discovery of Inference Rules from Text
   - Related work

2 Using DIRT for RTE
   - Observations
   - Extension and refinement
   - Application to RTE
   - Experiments and discussion

3 Future work
Introduction

Paraphrases

Expressions which can be substituted without changing the meaning of the sentences.

(find solution to, solve problem of)
(provide support to, offer aid to)
(has indicated he wants to return to, is considering returning to)
Introduction

Paraphrases
Expressions which can be substituted without changing the meaning of the sentences.

(find solution to, solve problem of)
(provide support to, offer aid to)
(has indicated he wants to return to, is considering returning to)

Textual entailment
Text entails Hypothesis if humans reading T will infer that H is most likely true.

T: *Bush used his weekly radio address to try to build support for his plan to allow workers to divert part of their Social Security payroll taxes into private investment accounts.*

H: *Mr. Bush is proposing that workers be allowed to divert their payroll taxes into private accounts.*
Paraphrases

Expressions which can be substituted without changing the meaning of the sentences.

(find solution to, solve problem of)
(provide support to, offer aid to)
(has indicated he wants to return to, is considering returning to)

Textual entailment

Text entails Hypothesis if humans reading T will infer that H is most likely true.

**T:** *Bush used his weekly radio address to try to build support for his plan to allow workers to divert part of their Social Security payroll taxes into private investment accounts.*

**H:** *Mr. Bush is proposing that workers be allowed to divert their payroll taxes into private accounts.*

Paraphrases for textual entailment?
Outline

1. **Background**
   - DIRT Discovery of Inference Rules from Text
   - Related work

2. **Using DIRT for RTE**
   - Observations
   - Extension and refinement
   - Application to RTE
   - Experiments and discussion

3. **Future work**
Automatic Acquisition of Inference Rules. DIRT

- Automatic acquisition of paraphrases using comparable corpora
  - Barzilay & al, 2001
  - Pang & al, 2003 - multiple translations
  - Shinyama & al, 2003 - news about the same story
Automatic Acquisition of Inference Rules. DIRT

- Automatic acquisition of paraphrases using comparable corpora
  - Barzilay & al, 2001
  - Pang & al, 2003
  - multiple translations
  - Shinyama & al, 2003
  - news about the same story

- DIRT (Discovery of Inference Rules from Text)
  - Lin and Pantel, 2001

Extended Distributional Hypothesis

If two paths tend to occur in similar contexts, the meanings of the paths tend to be similar.

Paraphrase representation

\[ X \leftarrow subst \hspace{1em} prevent \hspace{1em} obj \rightarrow Y \]
\[ X \leftarrow subst \hspace{1em} provide \hspace{1em} obj \rightarrow protection \hspace{1em} mod \rightarrow against \hspace{1em} pcomp \rightarrow n \rightarrow Y \]

- > 12 mil. rules (extracted from 1G of newspaper text)
- Estimated accuracy of most confident rules: \( \approx 50\% \)
- Errors: phrases with opposite meanings are also extracted
Outline

1 Background
   - DIRT Discovery of Inference Rules from Text
   - Related work

2 Using DIRT for RTE
   - Observations
   - Extension and refinement
   - Application to RTE
   - Experiments and discussion

3 Future work
**Using DIRT for RTE**

**RTE3** 45 systems (26 teams), 4 teams use DIRT

- Clark & al.
- Bar-Haim & al.
- Iftene & al.
- Marsi & al.

larger systems

focused on using DIRT
Using DIRT for RTE

**RTE3** 45 systems (26 teams), 4 teams use DIRT

Clark & al.  Bar-Haim & al.  Iftene & al.  larger systems

Marsi & al.  focused on using DIRT

**Inference rule**

\[ \text{pattern}_1(X, Y) \rightarrow \text{pattern}_2(X, Y) \]

**Directional** relation between two text patterns with variables. The left-hand-side template is assumed to entail the right-hand-side template in **certain** contexts, under the same variable instantiation.

Paraphrases: bidirectional inference rules.
Outline

1 Background
   - DIRT Discovery of Inference Rules from Text
   - Related work

2 Using DIRT for RTE
   - Observations
   - Extension and refinement
   - Application to RTE
   - Experiments and discussion

3 Future work
Using DIRT for recognizing textual entailment

- sell Y to X ↔ X buy Y

**T:** The sale was made to pay Yukos’ US$ 27.5 billion tax bill, Yuganskneftegaz was originally sold for US$ 9.4 billion to a little known company Baikalfinansgroup which was later bought by the Russian state-owned oil company Rosneft.

**H:** Baikalfinansgroup was sold to Rosneft.

- ≈ 2% of RTE sets
- > 80% correct entailment rules ( >60% positive entailment)
Using DIRT for recognizing textual entailment

- **sell Y to X ↔ X buy Y**
  
  **T:** The sale was made to pay Yukos’ US$ 27.5 billion tax bill, Yuganskneftegaz was originally sold for US$ 9.4 billion to a little known company **Baikalfinansgroup** which was later **bought** by the Russian state-owned oil company **Rosneft**.
  
  **H:** **Baikalfinansgroup** was **sold to Rosneft**.

  - ≈ 2% of RTE sets
  - > 80% correct entailment rules (>60% positive entailment)

- **X concern Y ↔ X involve Y**

  **T:** **Libya’s case against Britain and the US concerns** the dispute over their demand for extradition of Libyans charged with blowing up a Pan Am jet over Lockerbie in 1988.
  
  **H:** **One case involved** the extradition of Libyan suspects in the Pan Am Lockerbie bombing.

  - Upper bound ≈ 20% of RTE sets
Using DIRT for recognizing textual entailment

- RTE pairs require knowledge which can be encoded as inference rules

- X write Y $\leftrightarrow$ X author Y
- X founded in Y $\leftrightarrow$ X opened in Y
- X launch Y $\rightarrow$ X produce Y
- X represent Y $\rightarrow$ X work for Y
- X faces menace from Y $\leftrightarrow$ X endangered by Y
- death relieved X $\leftrightarrow$ X died
- X, peace agreement for Y $\rightarrow$ X is formulated to end war in Y
- X passed the leadership of Y to Z $\rightarrow$ X belongs to Y
Outline

1 Background
  • DIRT Discovery of Inference Rules from Text
  • Related work

2 Using DIRT for RTE
  • Observations
  • Extension and refinement
  • Application to RTE
  • Experiments and discussion

3 Future work
Extending and refining DIRT

Add extra lexical knowledge to deduce new rules?

1. Allow every word in a rule to be replaced by a WordNet synonym
   \( X \text{ face threat of } Y \)
   \[ \approx X \text{ at risk of } Y \]
   \[
   \begin{align*}
   \text{face} \quad & \approx \text{confront, front, look, face up} \\
   \text{threat} \quad & \approx \text{menace, terror, scourge} \\
   \text{risk} \quad & \approx \text{danger, hazard, jeopardy, endangerment, peril}
   \end{align*}
   
   Problems: Incorrect rules added due to sense ambiguity, propagation of erroneous rules

2. Post-processing DIRT. Remove rules containing antonyms:
   - \( X \text{ have confidence in } Y \leftrightarrow X \text{ lack confidence in } Y. \)
Outline

1 Background
   - DIRT Discovery of Inference Rules from Text
   - Related work

2 Using DIRT for RTE
   - Observations
   - Extension and refinement
   - Application to RTE
     - Experiments and discussion

3 Future work
Tree skeletons

Dependency-based structures  

1. Identify two pairs of anchor nodes (in T and H)
2. Extract the dependency tree chains connecting the anchor nodes

**T:** For their discovery of ulcer-causing bacteria, Australian doctors **Robin Warren** and Barry Marshall have received the 2005 **Nobel Prize** in Physiology or Medicine.

**H:** Robin Warren was awarded a Nobel Prize.

**Figure:** Dependency structure of text. Tree skeleton in bold
Tree skeletons and inference rules

Figure: Dependency structure of hypothesis. Tree skeleton in bold

Rule matched in tree skeleton

\[ X \leftarrow^{subj} receive \rightarrow^{obj} Y \rightarrow X \leftarrow^{obj1} award \rightarrow^{obj2} Y \]
Outline

1 Background
   • DIRT Discovery of Inference Rules from Text
   • Related work

2 Using DIRT for RTE
   • Observations
   • Extension and refinement
   • Application to RTE
   • Experiments and discussion

3 Future work
Experiments

If a pair contains a **tree skeletons** and an **inference rule** is matched, decide it is a case of positive entailment.

- Collection: Dirt, Top-40 rules ( > 4 mil. rules)
- Data sets: RTE2 (1600 pairs), RTE3 (1600)
- Tree skeleton coverage: ≈ 30%
- Rule collections: Dirt, Dirt+WN, Id (identity rules), Dirt+Id+WN

| Set   | Dirt<sub>TS</sub> | Dirt+WN<sub>TS</sub> | Id<sub>TS</sub> | Dirt+Id+WN<sub>TS</sub> | Dirt+Id+WN |
|-------|-------------------|----------------------|-----------------|--------------------------|------------|
| RTE2  | 49/0.69           | 94/0.67              | 45/0.66         | **130/0.65**             | 673/0.50   |
| RTE3  | 42/0.69           | 70/0.70              | 29/0.79         | **93/0.72**              | 661/0.55   |

**Table:** results with various rule collections. No of pairs covered/Precision on these pairs
BoW: Baseline overlap system. (Counts word overlap and is trained to learn a threshold)

BoW&Main: Our system with BoW backup on the rest of the pairs

| RTE Test (# pairs) | BoW      | BoW&Main |
|--------------------|----------|-----------|
| RTE2 (85)          | 51.76%   | 60.00%    |
| RTE3 (64)          | 54.68%   | 62.50%    |
| RTE2 (800)         | 56.87%   | 57.75%    |
| RTE3 (800)         | 61.12%   | 61.75%    |
Error Analysis

25 pairs (RTE3 test errors)

| Source of error      | % pairs |
|----------------------|---------|
| Incorrect rules      | 16%     |
| Rule application     | 32%     |
| Other errors         | 52%     |

1. X generate Y ↔ X earn Y, X issue Y ↔ X hit Y
2. ... founded the Institute of Mathematics at the University of Milan → University of Milan was founded by ...
3. Other errors
   - could be managed in a profitable manner → is managed in a profitable manner
   - rains, create flooding, devastate → floods are ravaging
Future work

- Combine various resources to obtain more lexical and world knowledge.
- Use more complex inference rules (e.g., inference rules with selectional preferences, directional inference rules, Basili et al., 2007, Szpektor et al., 2008, Bhagat et al., 2008).
- Develop a paraphrase-oriented annotation of the RTE data.
Future work

- Combine various resources to obtain more lexical and world knowledge
- Use more complex inference rules (e.g. inference rules with selectional preferences, directional inference rules, Basili et al., 2007, Szpektor et al., 2008, Bhagat et al., 2008)
- Develop a paraphrase-oriented annotation of the RTE data
- Dependency parsing with richer annotation (NE recognition, anaphora resolution)

We thank Dekang Lin and Patrick Pantel for providing the DIRT collection.
Bibliography

Basili, Roberto and Diego De Cao and Paolo Marocco and Marco Pennacchiotti

Learning Selectional Preferences for Entailment or Paraphrasing Rules.
In Proceedings of RANLP, 2007

Bhagat, Rahul and Pantel, Patrick and Hovy, Eduard

LEDIR: An Unsupervised Algorithm for Learning Directionality of Inference Rules.
In Proceedings of the 2007: EMNLP-CoNLL, 2007

Dagan, I., Glickman, O., and Magnini, B..

The Pascal Recognizing Textual Entailment Challenge.
In Lecture Notes in Computer Science, Vol. 3944, Machine Learning Challenges, 2006

Dekang Lin and Patrick Pantel

DIRT - Discovery of Inference Rules from Text
Proceedings of the ACM SIGKDD Conference on Knowledge Discovery and Data Mining, 2001

Szpektor, Idan and Dagan, Ido and Bar-Haim, Roy and Goldberger, Jacob

Contextual Preferences.
In Proceedings of ACL-08: HLT, 2008