THU at TAC 2008 RTE track

Minlie Huang

Assistant professor

aihuang@tsinghua.edu.cn

Joint work with F Li, Z Zheng,
Y Tang, F Bu, R Ge, X Zhu

Tsinghua University, Beijing, China

Outline

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

Outline

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

Introduction

清华大学
Tsinghua University

- Recognizing Textual Entailment (RTE)
  - To decide whether the text can entail the hypothesis.
- Hypothesis to solve the problem
  - To recognize TRUE entailment and FALSE entailment is different.
  - We should design different strategy for TRUE Entailment Recognition and FALSE Entailment Recognition

Outline

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

TRUE Entailment Recognition

清华大学
Tsinghua University

- Hypothesis on True Entailment
  - We believe that the hypothesis (H) can be transferred to the text (T) by some transformation rules
  - H and T have high similarities in TRUE Entailments
- Strategy to Recognize TRUE Entailment
  - Compare the similarity between the hypothesis and the text in several levels of linguistic representation
    - Word similarity
    - Phrase (Named Entity) similarity
    - Syntactic similarity
    - Semantic similarity

Word Match Similarity

清华大学
Tsinghua University

- We designed an extended LLM algorithm *(Tutorial @ ACL 07)* based on WordNet relations
  - Remove stop-words
  - Lemmatize words in Hypothesis and Text
  - Compute the number of matches by using WordNet Relation Comparison
  - Return #Matches / HYP_Lemmas as a feature value

Word Match Similarity

清华大学
Tsinghua University

- WordNet Relation Comparison
  - textWord and HypothesisWord has the same lemma
  - MemberOfSameSynset(textWord, hypothesisWord)
  - SynonymOf(textWord, hypothesisWord)
  - HypernymDistanceFromTo(textWord, hypothesisWord) <= 3
  - EntailedBy(textWord, hypothesisWord) or CausedBy(textWord, hypothesisWord)
  - MeronymyDistanceFromTo(textWord, hypothesisWord) <= 3
  - Deviated(textWord, hypothesisWord))

Named Entity Similarity

清华大学
Tsinghua University

- Two kinds of NER tools are used
  - Stanford NER, and Sharp
- Eight types of NEs are identified
  - Person, Location, Organization, Misk, Time, Money, Date, Percent
- Five predefined match relations
  - Full Match
  - NE words match, but Type not match (NER wrong)
  - Type Match, but NE words not match
  - MisMatch
  - Only appears in H

Syntactic Similarity

清华大学
Tsinghua University

- The sentence is first translated to the parse tree
- Two methods are used
  - The normalized alignment of dependency tree algorithm (Marsi, 2006)
  - We also design a simple path similarity calculation method based on the syntactic tree

Semantic Similarity

- A Semantic Role Labeler is designed to label predicate and arguments for hypothesis
- The semantic similarity is calculated as

$$SRL\_score = predicate\_score * arg\_score$$

- where predicate score is approximated as

$$predicate\_score = \begin{cases} 1 & verb\ match \\ 0.5 & otherwise \end{cases}$$

- argument score:

$$arg\_score = \sqrt[n]{\prod_{i=1}^{n} Sim(H(i), \arg_i)}$$

  - n is the number of *args* and arg_i is the *ith* arg for predicate

Integration

- For each type of similarity, we can set a threshold, and consider the pair with higher score as true entailment
- We also use machine learning tool *weka* to combine different similarity scores.

Outline

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

FALSE Entailment Recognition

清华大学
Tsinghua University

- Hypothesis on FALSE Entailment
  - We believe that there must be some mismatches between the text and the hypothesis.
- Strategies to Recognize FALSE Entailment
  - Exact entity and relation mismatches are detected to recognize false entailment

Number Mismatch

清华大学
Tsinghua University

- Number is an “exact” entity
  - If the number appears in the Hypothesis, but doesn’t appear in the Text, we predict that the entailment is false
  - Also consider simple number calculation
    - if the hypothesis contains “over 1000” and the text contains a numeric value above 1000, like 1024, we don’t predict it as false entailment.

Time & Date Mismatch

清华大学
Tsinghua University

- Time & Date are exact entities
  - If the time and date mismatch, we predict it is false entailment.

Location Mismatch

清华大学
Tsinghua University

- Location is an exact entity
  - For each location entity in H, if there is no corresponding entity in T, we predict it is false entailment.
  - Wikipedia is used to extend the location relations
    - country-nationality (China--Chinese) list and country-capital (China--Beijing) list are acquired

Quantifier Mismatch

清华大学
Tsinghua University

- Quantifier is an exact entity
  - Including “all”, “every”, “each”, “none”, “no”
  - If a noun (or noun phrase) appears both in H and T and this noun is modified by quantifier in H but not in T, then we predicate a false entailment

“Say” relation mismatch

清华大学
Tsinghua University

- An exact relation
  - This relation mismatch means that somebody says something happens in Text, but in Hypothesis, it is said that something happens.
  - We predicate it as false entailment

“Locate” relation mismatch

清华大学
Tsinghua University

- An exact relation
  - For some special location description words, like “locate”, “base”, “from” exist in Text.
  - We first align the objects, and then compare the subjects for these location markers in text and hypothesis.
  - If the subjects mismatch, we predict it is false entailment.

Negation and subjunctive mismatch

- The text is first split into several segment sentences.
- Choose a segment sentence, whose similar score with hypothesis is higher than a predefined threshold
- If the negation or subjunctive words only appear in the segment sentence or hypothesis, we predicate it as false entailment

Outline

清华大学
Tsinghua University

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

Submissions & Results

清华大学
Tsinghua University

- Two Submissions
  - The first (QUANTA1) uses the word match score, named entity match score and task description as features to recognize true entailment and use all the mismatch rules to recognize false entailment.
  - The second (QUANTA2) uses all the four match scores as features to recognize true entailment and use all the mismatch rules to recognize false entailment

Submissions & Results

清华大学
Tsinghua University

Two-way evaluation (6)	Accuracy	Average precision
QUANTA1	0.659	0.6225
QUANTA2	0.623	0.5926

Three-way evaluation (5)	Accuracy	Average precision
QUANTA	0.588	0.6332

Outline

- Introduction
- TRUE Entailment Recognition
- FALSE Entailment Recognition
- Evaluation Results
- Conclusion

Conclusion

清华大学
Tsinghua University

- A simple framework for recognizing textual entailment
- We design different strategies for TRUE Entailment Recognition and FALSE Entailment Recognition.
- The evaluation results show that our approach is suitable for RTE task

THU at TAC 2008 RTE track

清华大学
Tsinghua University

Thank you!