

Report for COMP5211 Project -- Winograd Schema Challenge

LIU Boyu

1. Algorithm:

There are two main categories of features that I used to decide the answer.

I get both ideas from this paper:

<http://www.hlt.utdallas.edu/~vince/papers/emnlp12.html>

(1) Features

Feature 1: Use search engine to search for the matching of two candidates with part of sentence that belong to the pronoun. (The program used Bing)

Feature 2: Use Narrative Chain to get the role of the pronoun in the first part, and find out which candidate satisfy the role.

(2) Detailed Steps

Step 1:

Input data from txt/ xml

Step 2:

Cut the sentence to two parts, with the conjunction as separation.

Step 3:

Use Stanford CoreNLP to analyze the two parts, including dependencies, POS, NER, Lemma, and other needed annotations.

Step 4:

Use Narrative Chain: extract the verb of two parts from result of *Step 3*, with requirements, that their subject should be the first candidate (part 1) or the pronoun (part 2). And use the dependencies to get whether this is a passive mode. Then check the two verbs from two parts (if more than one from a part, check every possible pair). Then compare the numbers of result that pronoun should be first verb's subject or object, adding the information of passive mode, try to decide the candidate. If cannot decide, then go on to *step 5*.

Step5:

Try to match the two candidates with part of sentence that belongs to the pronoun:

[1] Substitute the pronoun with candidates in second part, and get the new sentence with the candidates matched to the whole part of sentence that belong to pronoun

[2] <1> If the second part is Verb + Object, then match "Candidates + Verb"

<2> If the second part is Be + Adj, then match "Adj + Candidates"

<3> If the second part is Be + Noun, then match "Candidates + Be + Noun"

Search [1] on search engine, if one of the candidates has more result than another, we treat it more suitable for the context of pronoun. If cannot be decided, search [2] and try to decide.

(3) Example Analysis (chosen from default dataset)

[1] The Police feared allowing the protesters to organize, so they denied their ability to gather

Explain: From the narrative chain, we find that the subject of deny will be the subject of allow, so Police is treated as the answer, which is right.

[2] Lions eat zebras because they are predators.

Explain: From the search engine, we get more results of “Lions are predators” than “zebras are predators”, so Lions is chosen as the answer, which is right.

2. Usage

[1] Make sure Python 2.7/ 3.4+ is installed

[2] Install libraries by “pip install -r requirements.txt” in the program directory

[3] Install Stanford CoreNLP, and add its directory to \$PATH

[4] Put your data into “input.txt” with same format (can use xml but need to change the program a bit, instructions in the comments, default is from dataset of the paper, in text format)

[5] “Python main.py”, and you can see the result in the console

(* Data Format: Please refer to input.txt and WSCollection.xml)

3. Evaluation

With default input from the dataset:

<http://www.hlt.utdallas.edu/~vince/data/emnlp12/>

The result is:

Total data: 564

Correct: 52.127660%

Wrong: 37.765957%

No Decision: 10.106383%

Error rate: 42.7305%*

Correct rate: 57.2695%*

(*Calculate based on random set “No Decesion” part of data, 50% correct rate)

4. Analysis

The result may greatly depend on the input, as the method implemented focuses on the relationship of (1) Candidates noun and the context of pronoun (2) Verb in the two parts of sentences.

Another problem is that the search engine may return some unexpected results. As most websites are news, this situation may happen: “men bite dogs” are more than “dogs bite men”. And this is not what we want.

So if both relationships make no difference out of two candidates, or the search result is not accurate, this method may fail.

5. Improvement

- <1>We can use FrameNet to boost the relationship between the two words (discussed in the paper)
- <2>We can add the polarity convergence of the two parts of sentence into consideration
- <3>We may design a better search pattern
- <4>If we have more features, we can use machine learning to decide the weight of each feature.

6. Reference

- [1] Paper: <http://www.hlt.utdallas.edu/~vince/papers/emnlp12.html>
- [2] Dataset: <http://www.hlt.utdallas.edu/~vince/data/emnlp12/>
- [3] Stanford CoreNLP: <http://stanfordnlp.github.io/CoreNLP/index.html>
- [4] Engine: www.bing.com
- [5] Narrative Chain: <https://www.usna.edu/Users/cs/nchamber/data/schemas/acl09/>