Problem Statement and Motivation

- Currently there is a set of WSC problems that are especially difficult to resolve due to sentiment issues.
- All these sentences have some sort of conditional statement.
- It would be nice to explore models that may or may not improve the performance on these sentences.

10   Fox News is more popular in ratings than CNN since they are more interesting.

11   Fox News is more popular in ratings than CNN since they are boring.

74   Sega lost to Nintendo because they were superior.

75   Sega lost to Nintendo because they were inferior.

158   Obama beat John McCain, because he was the better candidate.

2   Spiderman defeated Magnito because he is a good guy.

3   Spiderman defeated Magnito because he is a bad guy.

8   Federer consistently beat Nadal since he was the better tennis player.

9   Federer consistently beat Nadal since he was the worse tennis player.

16   UPS provides much convenience to the customers with their drop off service since they make the shipping of packages extremely easy.

■ ■ ■

A First Simple Idea for a Solution

1. Identify the sentiment of a pronoun using some kind of heuristic.

2. Parse the sentence and calculate sentiment for its dependencies.

3. Compare the sentiment of the dependencies and the sentiment of the pronoun.

A First Simple Idea for a Solution

1. Identify the sentiment of a pronoun using some kind of heuristic.

2. Parse the sentence and calculate sentiment for its dependencies.

3. Compare the sentiment of the dependencies and the sentiment of the pronoun.

A Selectional Preference Model with Sentiment Values

- I wrote a simple program that takes a **word**, a **dependency** and a **position** as input and produces a "sentiment preference" value as output.

Example: **sp(defeat,dobj,head)**

numPos: 22, numNeg: 24

defeat election vb nn -1.0 0.012336

defeat attacker vb nns -1.0 -0.188084

defeat motion vbd nn -1.0 -0.165506

defeat state vbn nnp -1.0 -0.138598

defeat threat vb nns -1.0 -1.0

defeat blazer vbg nns -1.0 -0.0231835

defeat orleans vbd nnp -1.0 -0.0184837

defeat taliban vb nnp -1.0 0.00165599

defeat people vbp nns -1.0 0.010716

...

Example: **sp(defeat,dobj,argument)**

numPos: 3, numNeg: 2
acknowledge defeat vbg nn 0.106926 -1.0
suffer defeat vbd nn -1.0 -1.0
put defeat vb nn -0.118919 -1.0
admit defeat vbg nn 1.0 -1.0
grade defeat vbn nn 0.0262358 -1.0

The output has the following form:

word_1, word_2, pos_1, pos_2, sent_1, sent_2

Details

- Takes dependency relationships and POS-tags from the *clueweb12 corpus* and builds a file out of these depending on which dependency relationship is passed as argument. Both words are written to the file in its lemma form.

- Extracts sentiment values from the *Takamura et al.* sentiment lexicon and produces a second file.

- Queries are made from the second file.

A Possible Research Question

- Is there a relationship between the sentiment of a pronoun in a sentence and the sentiment of the same sentence's grammatical dependencies?

- *Follow up question:* If there is, is it possible to extract some useful knowledge to improve a pronoun resolution system?

Further Progress

Done: The simple sentiment-level preference function works pretty fast and seemingly bug-free.

Current: Parse the WSC-problems and evaluate the model.

To do: Either improve the model (like using different resources) or do something else.