Yitong Wang

CSCI-GA-2590

Homework 5

Mar 25th 2015

1.

1. With chunkPattern.txt provided, the statistics of the scoring are listed below:

# of matching tags = 90; # of tags in key = 114; # of tags in response = 140;

precision = 64.28%; recall = 78.94%; type accuracy = 100.0%

2. With the edited chunk patterns from last homework, the scores are as below:

# of matching tags = 101; # of tags in key = 114; # of tags in response = 124;

precision = 81.45%; recall = 88.59%; type accuracy = 100.0%

3. With the new article, the scores from the original chunk pattern:

# of matching tags = 93; # of tags in key = 130; # of tags in response = 153;

precision = 60.78%; recall = 71.53%; type accuracy = 100.0%

The scores from the modified chunk pattern:

# of matching tags = 105; # of tags in key = 130; # of tags in response = 137;

precision = 76.64%; recall = 80.76%; type accuracy = 100.0%

2.

Using original chunk patterns and article 1:

# of matching tags = 88; # of tags in key = 114; # of tags in response = 133;

precision = 66.16%; recall = 77.19%; type accuracy = 100.0%

Using original chunk patters and article2:

# of matching tags = 94; # of tags in key = 130; # of tags in response = 141;

precision = 66.66%; recall = 72.3%; type accuracy = 100.0%

Using modified chunk patterns and article1:

# of matching tags = 101; # of tags in key = 114; # of tags in response = 117;

precision = 86.32%; recall = 88.59%; type accuracy = 100.0%

Using modified chunk patterns and article2:

# of matching tags = 105; # of tags in key = 130; # of tags in response = 126;

precision = 83.33%; recall = 80.76%; type accuracy = 100.0%