

2009 Solutions

(E) Summary

An extractive summarizer scores each sentence according to some criteria that are correlated with being a good summary sentence. Then it picks the top 3 sentences from each story based on the sum of its scores on the different criteria.

The first criterion measures primacy. The first sentence gets 3 points, the second 2, and the third 1, to account for the likely increased importance of the initial sentences. Then multiples of .1 are added, starting with .0 for the last sentence, to break the other criteria's ties in favor of earlier sentences.

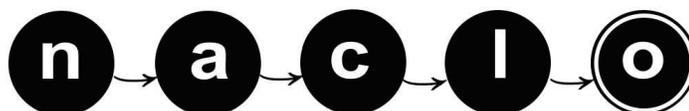
The second criterion measures recency. The last sentence gets 3 points, the second-last 2, and the antepenultimate 1, to account for that they're likely summaryish themselves.

The third criterion counts named entities in the sentence, since they're likely to be the most important actors.

The fourth criterion counts words from the title that appear in the sentence (after reducing each word to a stem; e.g. struck = strike). These sentences are likely to pertain most immediately to the topic of the story.

The fifth criterion counts named entities introduced for the first time in this sentence. The first mention of a named entity is probably important for understanding its role.

The sixth criterion counts past-tense verbs. Current information is probably more important, and past-tense verbs are slightly less likely to give new information.



2009 Solutions

(E) Summary

Solutions

Story 1, sentence 2, criterion 3 - change to 1

Story 1, sentence 2, criterion 5 - change to 1

Story 1, sentence 2, total - change to 2.3

Story 1, sentence 5, criterion 3 - change to 0

Story 1, sentence 5, total - change to 3.0

Story 2, sentence 1, criterion 1 - change to 3.9

Story 2, sentence 1, total - change to 8.9

Story 2, sentence 2, criterion 1 - change to 2.8

Story 2, sentence 2, total - change to 4.8

Story 2, sentence 3, criterion 1 - change to 0.6 (similarly change the next 6 numbers down from this one to be 0.1 less than before)

Story 2, sentence 2, total - change to 2.6 (this affects the totals for the next 6 sentences)

Story 2, sentence 5, criterion 5 - change to 2

Story 2, sentence 5, total - change to 4.4

Story 2, sentence 8, criterion 5 - change to 1

Story 2, sentence 8, total - change to 4.1

Part 2:

1.7 0 2 3 2 -4 4.7

