

Homework #5

In the previous assignment, we addressed the task of extracting relations from the literature in biomedicine and assessed its performance. In this assignment, we will address the problem of coreference resolution, as illustrated in the example text below, where the ambiguous pronoun 'she' is noted in bold, and the two potential coreferent names 'Fujisawa' and 'Mari Motohashi' in italic, with an underline for the correct one:

"In May, Fujisawa joined *Mari Motohashi's* rink as the team's skip, moving back from Karuizawa to Kitami where **she** had spent her junior days."

In this example, 'she' refers to 'Fujisawa' but does not refer to 'Mari Motohashi'. The expected output would be that 'Fujisawa' is 'TRUE' and 'Mari Motohashi' is 'FALSE'. An ambiguous pronoun may also refer to both of the two coreferent names or none.

For this homework assignment, you are asked to go through the following steps for coreference resolution: (1) Download GAP coreference repository from the link: <https://github.com/google-research-datasets/gap-coreference>, (2) develop your program that resolves ambiguous pronouns that refer to names, and (3) assess the performance of your program against 4,000 pairs in gap-test.tsv evaluated by Precision, Recall, F-score, and Bias ratio with two task settings: snippet-context and page-context. Snippet-context indicates that the context (sentences) in the Wikipedia URL cannot be used, and page-context indicates that the context can be used. Bias ratio is calculated by taking the ratio of feminine (F) to masculine (M) F1 scores - F/M. Further information about the data can be found from the link above. Note that your program will be evaluated as gap_scorer.py.

As before, you should use techniques that can be implemented in Python and NLTK, scored as a relevance.

A Write a Python code for coreference resolution based on the repository.

B Discuss your results, to explain how you addressed the goal and to suggest how you can improve the quality of the results further. This document must be in English.

All the requirements as underlined above must be composed by yourself and without help from anyone else. Any similarity of the results will be flagged for plagiarism and, if found sufficiently similar, penalized, up to, but not limited to, a failure to this homework.

Deadline for uploading your homework at KLMS: 26 June (11:59pm, STRICT)

Homework Submission Guidelines

1. Submission file (1 zip file)

CS372_HW5_submission_[your ID].zip

|—— CS372_HW5_code_[your ID].py that saves two outputs - snippet and page

|—— CS372_HW5_report_[your ID].docx and .pdf

|—— CS372_HW5_snippet_output_[your ID].tsv of test data

|—— CS372_HW5_page_output_[your ID].tsv of test data

|—— (optional) CS372_HW5_local_file_[No.]_[your ID] where the local files should be in the same directory of your model code and accessed by a relative path

|—— (optional) CS372_HW5_requirements_[your ID].txt if you import additional packages implemented in Python 3.x

2. Remarks

- Use a maximum of 2 pages for your discussion.
- The code should include comments about your implemented idea.
- For the implementation, you should not use external models, in particular those specialized for coreference resolution, NER, and entity linking.
- In line 160 of gap_scorer.py, replace “iteritems()” to “items()”.
- Your code should be runnable in our environment, where your code is executed based on gap-coreference repository.
- Use 11pt font size and default margin/line spacing for your report.