Sarika Halarnakar - shalarn1
Sindhuula Selvaraju - sselvar4

Machine Translation Assignment 5: Reranking

Contents of the Repository:

1. data : This directory contains the data files given to us and a smaller file we used for testing
2. results: This directory contains result files for the data given to us. The baseline result file is result_original
3. oracle : This has not been modified by us
4. compute-bleu : This has not been modified by us
5. bleu.py: This has not been modified by us
6. add_features.py : This file creates new data files with our added features
7. rerank: This is the original evaluate file given to us
8. rerank_features: This is the original evaluate file given to us with our added features
9. rerank_sim.py: This version of the rerank program has the implementation of the simplex algorithm

Part 1: Adding Features to File

Usage: (In Reranker)

```
python add_features
```

Part 2: Feature Implementation

Usage: (In Reranker)

```
python rerank_features > english.features.out
python compute-bleu < english.features.out
```

Part 3: Simplex Implementation

Usage: (In Reranker)

```
python rerank_simplex > english.simplex.out
python compute-bleu < english.simplex.out
```

Description:

We added 3 features: word count, untranslated words, and the meteor rating. We did this in our add_features.py file which computes the word count, the untranslated words and the meteor rating then adds each value to the original line and saves it in a new file dev+test.100best.added.feats.
We also implemented the simplex algorithm using method 3, moving close to the best point. We calculated the midpoint between the best point and the good point and the midpoint between the best point and the worst point and used those new points. We then compared the bleu scores and selected the best score. We found that the algorithm converges after about 20 iterations.

Results:

```
python compute-bleu < baseline.out 0.273509457562

python compute-bleu < english_features.out 0.285192374952

python compute-bleu < english_simplex.out 0.291046947628
```