

# 11-731: Machine Translation

## Homework Assignment #9:

Out: Wednesday, April 08, 2009

Due: 2:00 pm Wednesday, April 15, 2009

This homework consists of two parts. In part 1, you will evaluate different reordering models that were discussed in the class. In Part 2, you will use several metrics (BLEU, METEOR and TER) to evaluate translation hypotheses. Use the setup provided at the following location: [/afs/cs.cmu.edu/project/cmt-55/lti/Courses/731/homework/HW9](http://afs/cs.cmu.edu/project/cmt-55/lti/Courses/731/homework/HW9).

### Part 1

MOSES toolkit provides several options to train reordering models<sup>1</sup>. The default setup uses distance-based reordering model, which you used in HW7. Use it as the baseline system A. Now train lexicalized reordering models for the following configurations:

- B. Monotone (`monotonicity-f`)
- C. Monotone, Swap and Discontinuous (`msd-bidirectional-fe`)

For each of the configurations A through C, run MERT on the development set `./dev/dev.es` with reordering windows 4 and 6 (use `-distortion-limit` parameter to set the reordering window size). Report the resulting BLEU scores.

Compare the results and state your observations.

### Part2

- 1) Use the final translation hypotheses from MERT in each of the 6 configurations above and evaluate them using the following additional metrics (case insensitive):
  - a. METEOR
  - b. TER

The evaluation scripts are available in `./score` directory. Use `GenerateSGML.sh` to convert the output into SGML format as required by the evaluation scripts.

Do you see the same pattern as BLEU? State your observations.

- 2) Test for the BLEU Metric if the differences are statistically significant. You may refer to <http://projectile.sv.cmu.edu/research/public/tools/bootStrap/tutorial.htm> for details. Scripts are locally available in `./score/sig-test/`.

---

<sup>1</sup> <http://www.statmt.org/moses/?n=FactoredTraining.BuildReorderingModel>