Ling 575 2011 Statistical Machine Translation

Homework 1

Due April 6. 11:59pm Seattle time

Problem 1 (Machine Translation Evaluation)

Pick an online machine translation system, such as http://www.microsofttranslator.com/ or http://translate.google.com/#

Select documents from two different domains (for example, blogs, chat, email, Wikipedia articles, movie titles, poetry ...) and translate them into a language that you know.

a) Select three sentences from each domain and evaluate the translation for fluency (and also adequacy if you speak the source language) [this is a total of six sentences]. List the sentences together with your evaluation.

b) Discuss the results of this evaluation
   a. Does the online system seem better on one of the domains?
   b. What are some problems of the translation -- e.g. are the problems with word order, grammatical mistakes, unknown words, etc.?