The University of Texas
at Austin

CS388 Natural Language Processing

Programming Assignment 02

Edited by LaTeX

Department of Computer Science

STUDENT
Jimmy Lin
xl5224

COURSE COORDINATOR
Prof. Raymond J. Mooney

UNIQUE NUMBER
51360

RELEASE DATE
Feb. 20 2016

DUE DATE
Mar. 08 2016

TIME SPENT
10 hours

February 21, 2016

1. Problem Statement

2. Experiment

Comparisons of word perplexity between various Bigram Models can be found at Table 1.

Table 1: Word Perplexity Comparisons between Various Bigram Models

Datasets	atis		wsj		brown	
Bigram Models	training	testing	training	testing	training	testing
Forward	10.59	24.05	88.89	275.12	113.36	310.67
Backward	9.11	21.71	71.04	205.75	89.30	218.79
Bidirectional	9.37	17.54	48.60	130.48	63.74	172.75

Note that the Bidirectional Model employs even interporlation from Forward Model and Backward Model in the normal probability space (not log probability space).

From the Table 1, it can be observed that Backward Model and Bidirectional Model reach lower word perplexity (which is better) than Forward Model in any training/testing set derived from three experimental datasets. Furthermore, Bidirectional Model outperforms Forward Model in any scenario except the case of the training set from atis. This could be out of the overfitting phenomenon.

3. Discussion

How does the overall test accuracy of CRF and HMM differ (when using only tokens) and why?

How does the test accuracy for OOV items for CRF and HMM differ (when using only tokens) and why?

How does the training accuracy of HMM and CRF differ and why?

How does the run time of HMM and CRF differ and why?

How does adding orthographic features affect the accuracy (both overall and OOV) and runtime of the CRF and why?

Which features helped the most?