**NLP – Ex2**

**Q1**

b. #trigrams: 413540

#bigrams: 122930

#unigrams: 2000

#tokens: 1118296

#perplexity for lambda1(left) lambda2(top) grid search:

**000.000 000.100 000.200 000.300 000.400 000.500 000.600 000.700 000.800 000.900**

**000.000** 189.624 115.665 094.730 082.688 074.630 068.860 064.621 061.564 059.628 059.293

**000.100** 104.210 082.635 072.622 066.031 061.319 057.858 055.376 053.845 053.720 -01.000

**000.200** 087.509 071.626 064.313 059.421 055.935 053.475 051.972 051.850 -01.000 -01.000

**000.300** 079.029 065.420 059.467 055.555 052.887 051.272 051.076 -01.000 -01.000 -01.000

**000.400**  074.226 061.614 056.499 053.306 051.411 051.057 -01.000 -01.000 -01.000 -01.000

**000.500** 071.732 059.390 054.885 052.407 051.753 -01.000 -01.000 -01.000 -01.000 -01.000

**000.600** 071.141 058.488 054.581 053.320 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000

**000.700** 072.652 059.060 056.248 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000

**000.800** 077.440 062.138 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000

**000.900** 090.453 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000 -01.000

#best coefficients (lambda1, lambda2) are (0.4, 0.5) with perplexity 51.0567551405

d.

#params: 104550

#train examples: 1118296

training took 0 seconds (re-run for evaluation after training)

dev perplexity : 112.967665327

**Q3**

c.

Perplexity for each epoch:

[598.2366782129152, 258.35566638496755, 186.748500751412, 153.8955940469948, 133.7895546230422, 119.46139157102863, 109.51243179110595, 101.60273449526015, 95.69344691128141, 90.6552240888448, 86.34052603632608, 82.98170482530692, 80.04136627427236, 77.28627441036822, 75.2353598328523, 73.01505584565827, 71.32777081434989, 69.88417041995659, 68.37327530689748, 67.18151719599369]

Finished training

Epoch: 20 Validation Perplexity: 104.119