Assignment 2 (Part A)

Akshay Kumar Gupta (2013CS50275)

Implementation language: Python

Implementation:

Convergence criterion used: Average log-likelihood of the predicted words (eps = 1.0e-3), calculated after every 100 iterations post 1000 burn-in iterations.

(1) Comparison of Gibbs sampling and Loopy BP:

Data-Tree.dat (1)

	Character Accuracy	Word Accuracy	Log - Likelihood	Time Taken (s)
Gibbs Sampling	0.69436997	0.19047619	-13.18708561	73.927
Loopy BP	0.67560322	0.16666667	-8.94724050	8.656

Data-TreeWS.dat (1)

	Character Accuracy	Word Accuracy	Log - Likelihood	Time Taken (s)
Gibbs Sampling	0.67770419	0.17391304	-14.41166223	196.859
Loopy BP	0.66556291	0.15760870	-10.06001206	43.588

Data-Loops.dat (1)

	Character Accuracy	Word Accuracy	Log - Likelihood	Time Taken (s)
Gibbs Sampling	0.59712230	0.10714286	-15.07931821	30.102
Loopy BP	0.56834532	0.07142857	-10.63556823	6.680

Data-LoopsWS.dat (1)

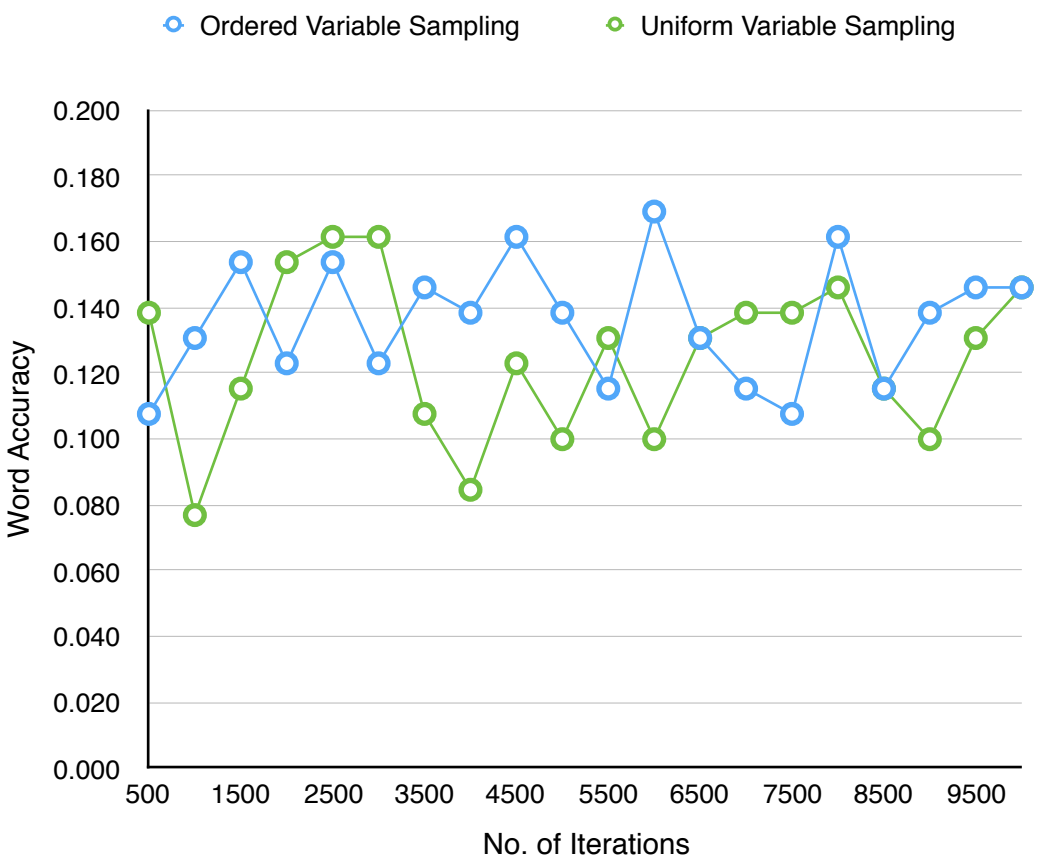
	Character Accuracy	Word Accuracy	Log - Likelihood	Time Taken (s)
Gibbs Sampling	0.67901235	0.13076923	-13.88341900	165.224
Loopy BP	0.68209877	0.16153846	-10.07198178	81.696

Observations:

1. The average log-likelihood of the predicted words is an acceptable convergence criterion as it always converges within 10000 iterations, and often in much less.
2. Different runs of gibbs sampling (with the same convergence criterion) give different word accuracies. For instance, on the Tree dataset, it varies from 0.13 to 0.19.
3. According to the results used above (which are not necessarily typical), gibbs sampling outperforms loopy BP on three datasets, while loopy BP does better on the LoopsWS dataset.

Varying the number of iterations:

(2) Comparison of variables sampled in a fixed order vs. variables sampled uniformly at random (on loopsWS dataset):



Observations:

1. After a burn-in of 1000 iterations, the number of iterations was varied from 500 to 10000 (using a skip of 500). As seen in the plot above, no clear trend emerges with the maximum accuracy occurring at 3000 iterations (for ordered sampling).

2. Running the same experiment again (with a different random number seed) generates a completely different graph, showing that at least until 10000 iterations no convergence in word accuracy occurs.

Observations (Extra Credit):

1. For uniform variable sampling too, there is no clear trend in the plot.
2. It generally performs poorer than ordered sampling, at least in the experiment performed here.
3. A possible reason for poorer results is that some variables are not sampled enough. Sampling for even more iterations than 10000 might increase its accuracy.