**ProbMod Homework 4 – January 2020**

**Bar Dagan 303026454**

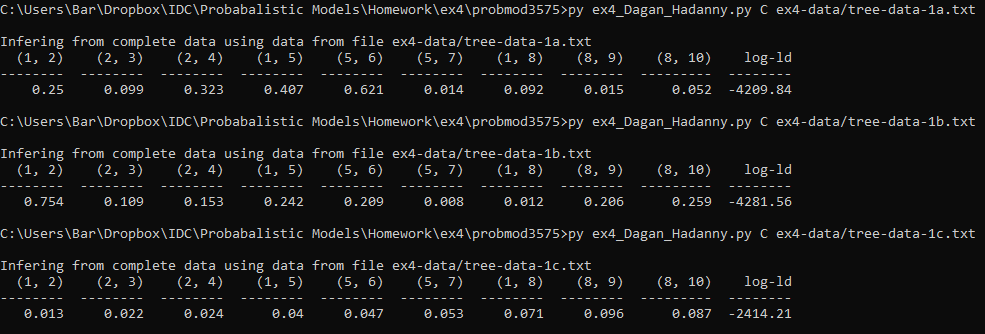
**Ido Hadanny 034537969**

Question 1:

The program invokes function “inference\_from\_complete\_data” which calls functions:

* Get\_table – read file and return table.
* Get\_parameters – calculates new parameters by normalizing counts of changes (0->1 or 1->0) for each edge.
* Get\_log\_likelihood – calculates log likelihood of complete data given a set of parameters.

The sufficient statistics in this case are the counts of changes (0->1 or 1->0) and counts of unchanged (1->1 or 0->0).



Question 2:

The program invokes function maximum\_probability\_inference which iterates until convergence:

1. For each observation in the data:

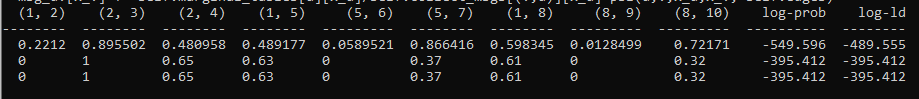
* collect\_distribute\_sum is called to calculate the likelihood
* collect\_distribute\_max is called to get a max probability assignment.

1. Get\_parameters and get\_log\_likelihood is used like in question 1 to calculate new parameters and log\_prob.

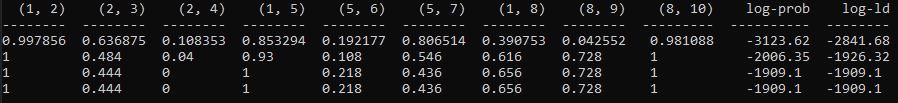
*\*collect\_distribute\_sum and collect\_distribute\_max are from Homework 3 so they are not explained here.*

The sufficient statistics in this case are expected sufficient statistics: counts of changed and unchanged observation of the max probability assignment.

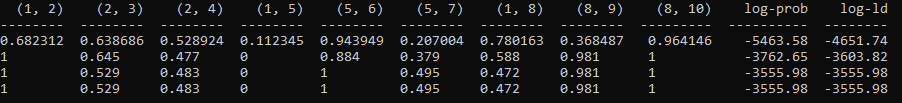
To find the best initial parameters, we randomly sampled a uniform distribution to select the 9 parameters. This process was repeated 1000 times and the initial parameters which yielded the best final log probability were used. This process yielded the output:  
For tree-data-2a.txt:



For tree-data-2b.txt:



For tree-data-2c.txt:



Log-likelihood is monotonic non-decreasing, therefore it always gets better from iteration to iteration. Increasing the halting threshold will therefore decrease the number of iterations while yielding a final log-likelihood that is not as high. Decreasing it will cause more iterations and a higher log-likelihood.

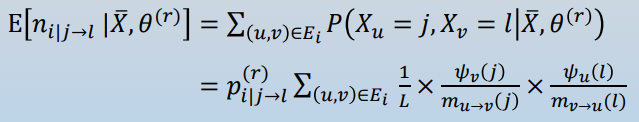
Question 3:

The program invokes function expectation\_maximization\_inference which iterates until convergence of log likelihood:

1. For each observation in the data:

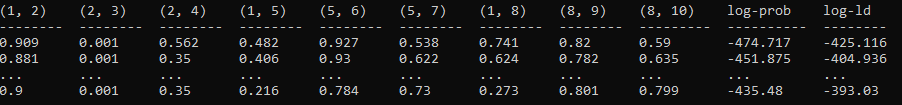
* collect\_distribute\_sum is called to calculate Ψ, m and likelihood (L) for all edges.
* collect\_distribute\_max is called to get a max probability assignment.

1. Sufficient statistics are calculated using equation below.
2. Log probability is calculated using the max probability assignment as in question 2.
3. New parameters are calculated from the ESS.

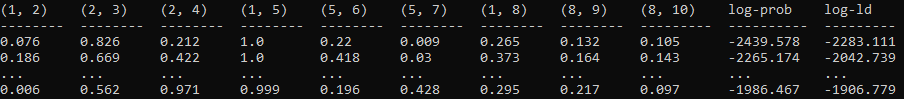
Expected sufficient statistic calculation: 

To find the best initial parameters we ran 100 different experiments with random uniform selection of parameters. Since EM requires much more iterations to converge, we only performed 100 and not 1000 different runs. Data presented below is a subset of the iterations for the best initial parameters because it took a lot of iterations for it to converge.

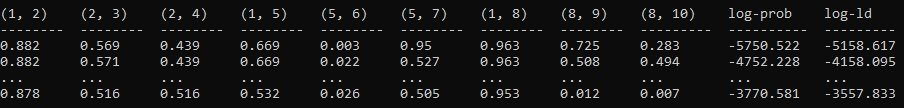
For tree-data-2a.txt:



For tree-data-2b.txt:



For tree-data-2c.txt:



Comparing Question2 and Question3. Generally, EM is better at reaching the best parameters for maximizing the log likelihood, but EM generally takes more iterations and is therefore slower and more expensive to run.