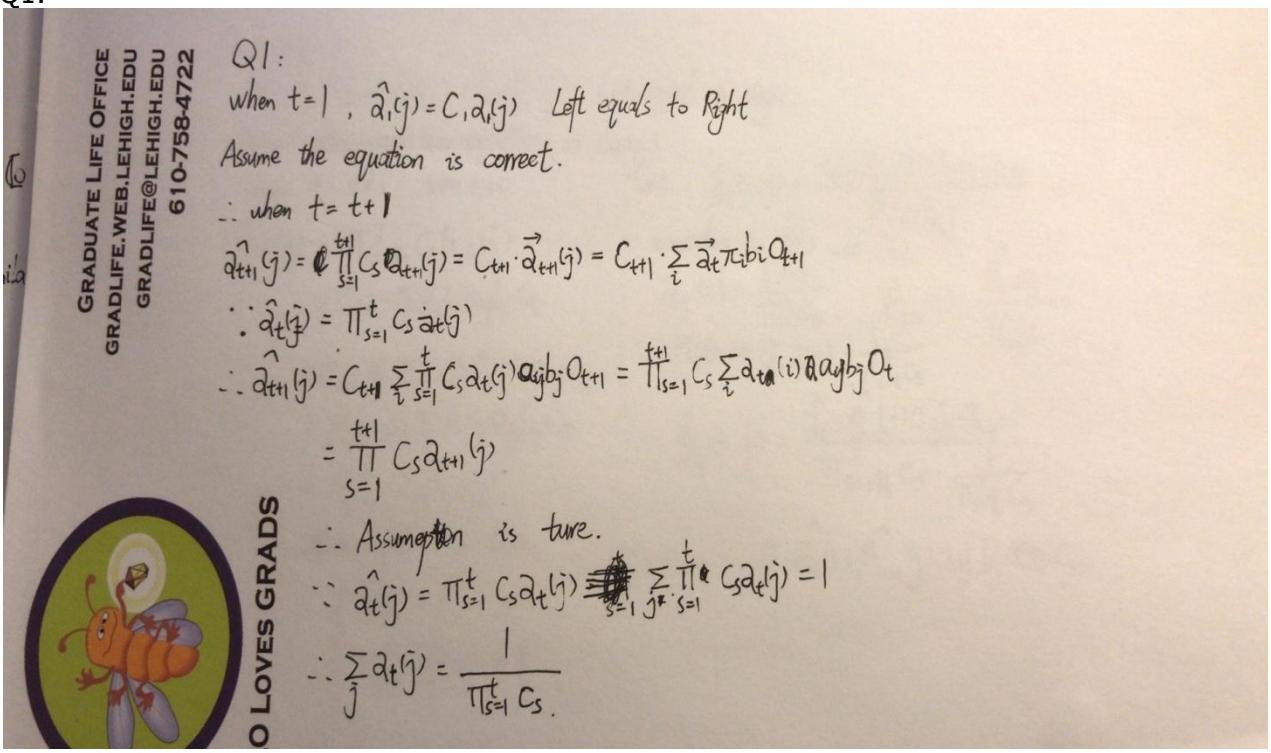
Name: Cun Shi Course: CSE 498-014 (NLP)

Due to plenty of complicated formulas and math characters, I prefer provide my hand-write anwsers than typing though Word. Here are photos of my hand writing answers.

Q1:



Q1:

when $t=1$, $\hat{\alpha}_1(j) = C_1 \alpha_1(j)$ Left equals to Right

Assume the equation is correct.

∴ when $t = t+1$

$$\hat{\alpha}_{t+1}(j) = \prod_{s=1}^{t+1} C_s \alpha_{t+1}(j) = C_{t+1} \cdot \vec{\alpha}_{t+1}(j) = C_{t+1} \cdot \sum_i \vec{\alpha}_t \pi_i b_i O_{t+1}$$

$$\because \hat{\alpha}_t(j) = \prod_{s=1}^{t} C_s \alpha_t(j)$$

$$\therefore \hat{\alpha}_{t+1}(j) = C_{t+1} \sum_i \prod_{s=1}^{t} C_s \alpha_t(j) a_{ij} b_j O_{t+1} = \prod_{s=1}^{t+1} C_s \sum_i \alpha_t(i) a_{ij} b_j O_t$$

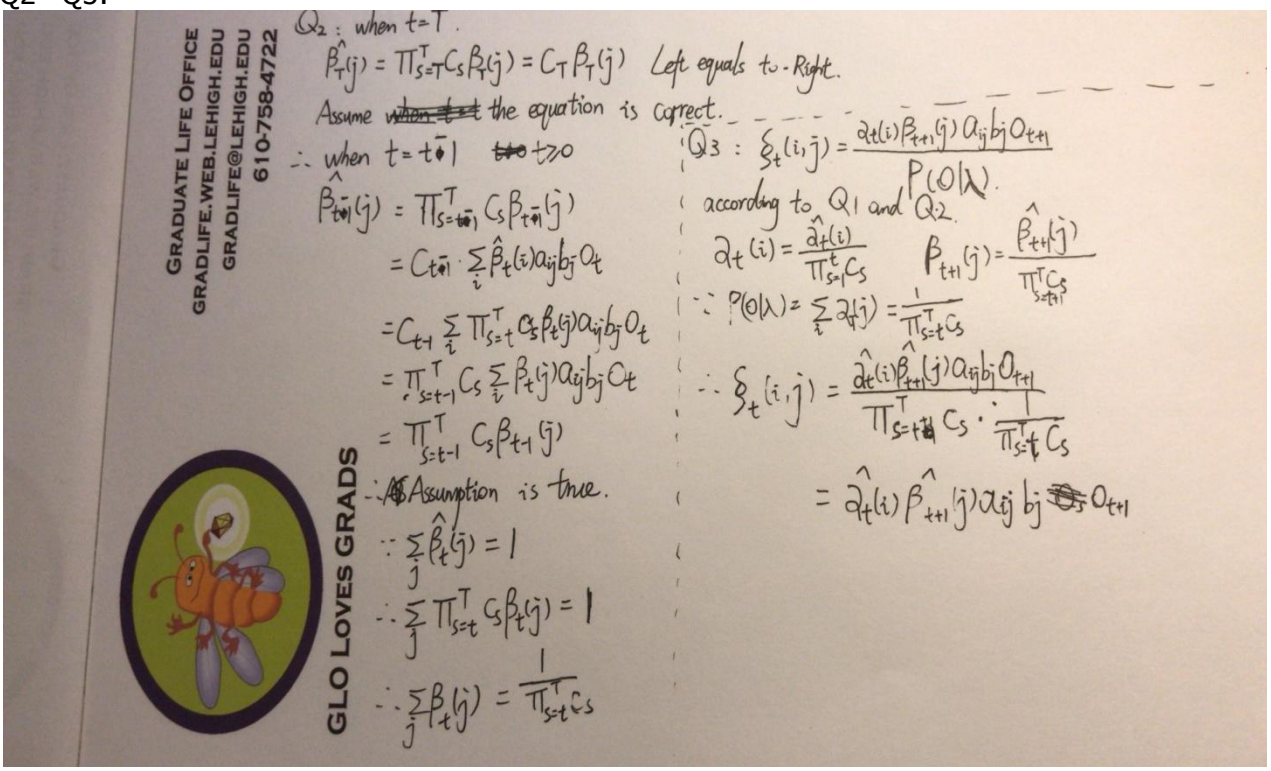
$$= \prod_{s=1}^{t+1} C_s \alpha_{t+1}(j)$$

∴ Assumption is ture.

$$\because \hat{\alpha}_t(j) = \prod_{s=1}^{t} C_s \alpha_t(j) \quad \sum_j \prod_{s=1}^{t} C_s \alpha_t(j) = 1$$

$$\therefore \sum_j \alpha_t(j) = \frac{1}{\prod_{s=1}^{t} C_s}$$

Q2—Q3:



Q2: when $t=T$.

$\hat{\beta}_T(j) = \prod_{s=T}^{T} C_s \beta_T(j) = C_T \beta_T(j)$ Left equals to Right.

Assume the equation is correct.

∴ when $t = t-1$ $t \geq 0$

$$\hat{\beta}_{t-1}(j) = \prod_{s=t-1}^{T} C_s \beta_{t-1}(j)$$

$$= C_{t-1} \cdot \sum_i \hat{\beta}_t(i) a_{ij} b_j O_t$$

$$= C_{t-1} \sum_i \prod_{s=t}^{T} C_s \beta_t(j) a_{ij} b_j O_t$$

$$= \prod_{s=t-1}^{T} C_s \sum_i \beta_t(j) a_{ij} b_j O_t$$

$$= \prod_{s=t-1}^{T} C_s \beta_{t-1}(j)$$

∴ Assumption is true.

$$\because \sum_j \hat{\beta}_t(j) = 1$$

$$\therefore \sum_j \prod_{s=t}^{T} C_s \beta_t(j) = 1$$

$$\therefore \sum_j \beta_t(j) = \frac{1}{\prod_{s=t}^{T} C_s}$$

Q3: $\xi_t(i,j) = \frac{\alpha_t(i) \beta_{t+1}(j) a_{ij} b_j O_{t+1}}{P(O|\lambda)}$

according to Q1 and Q2.

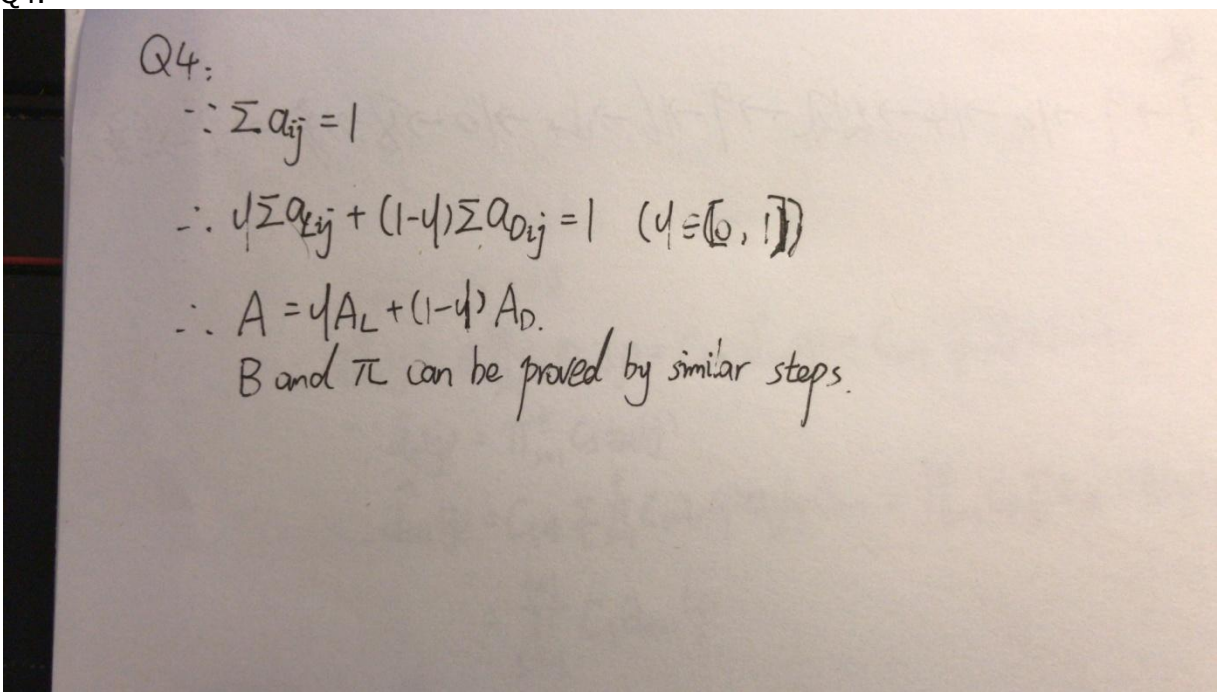
$$\alpha_t(i) = \frac{\hat{\alpha}_t(i)}{\prod_{s=1}^{t} C_s} \qquad \beta_{t+1}(j) = \frac{\hat{\beta}_{t+1}(j)}{\prod_{s=t+1}^{T} C_s}$$

$$\because P(O|\lambda) = \sum_i \alpha_t(j) = \frac{1}{\prod_{s=t}^{T} C_s}$$

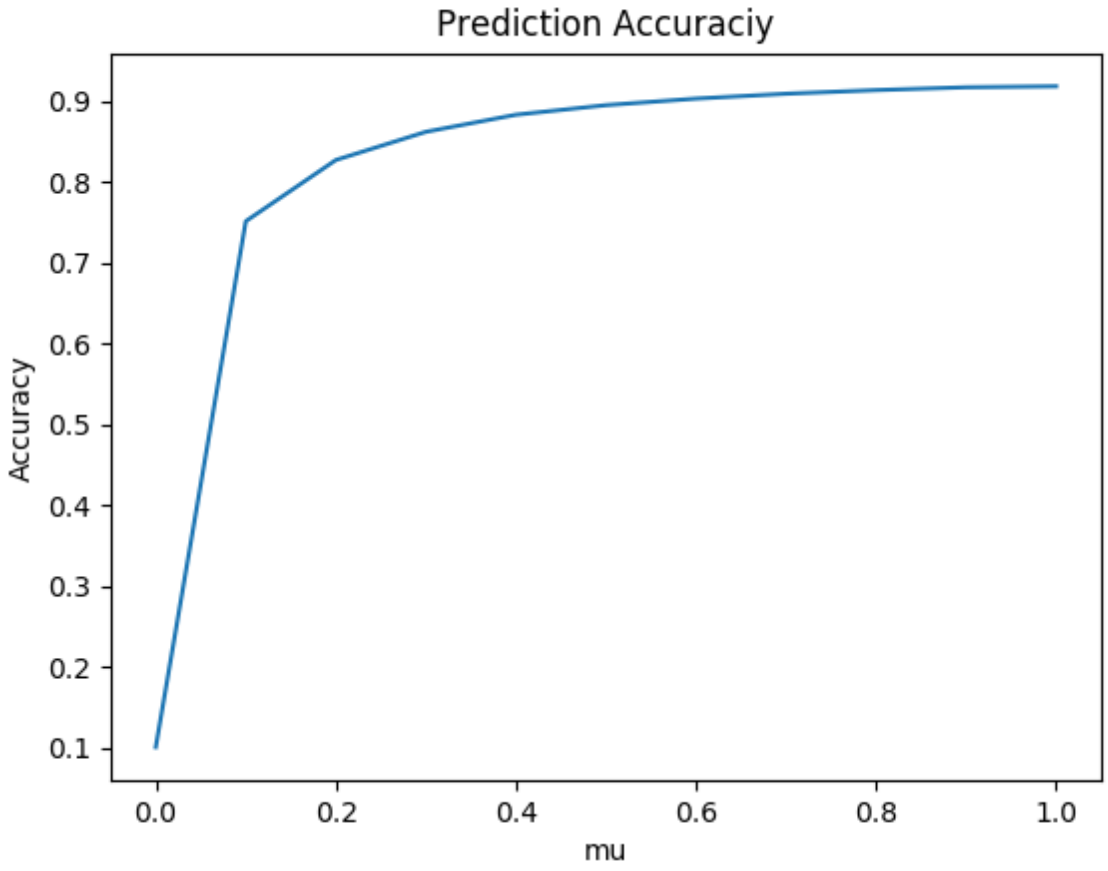
$$\therefore \xi_t(i,j) = \frac{\hat{\alpha}_t(i) \hat{\beta}_{t+1}(j) a_{ij} b_j O_{t+1}}{\prod_{s=t+1}^{T} C_s \cdot \frac{1}{\prod_{s=t}^{T} C_s}}$$

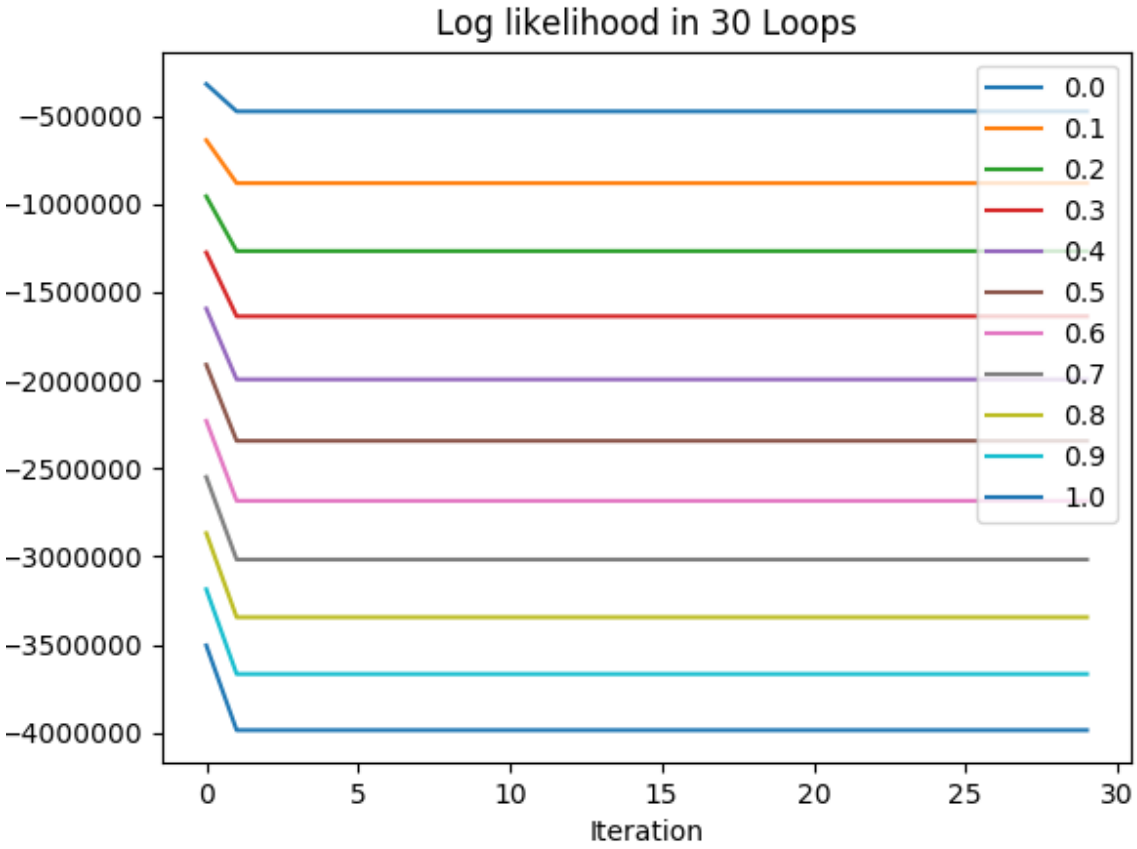
$$= \hat{\alpha}_t(i) \hat{\beta}_{t+1}(j) a_{ij} b_j O_{t+1}$$

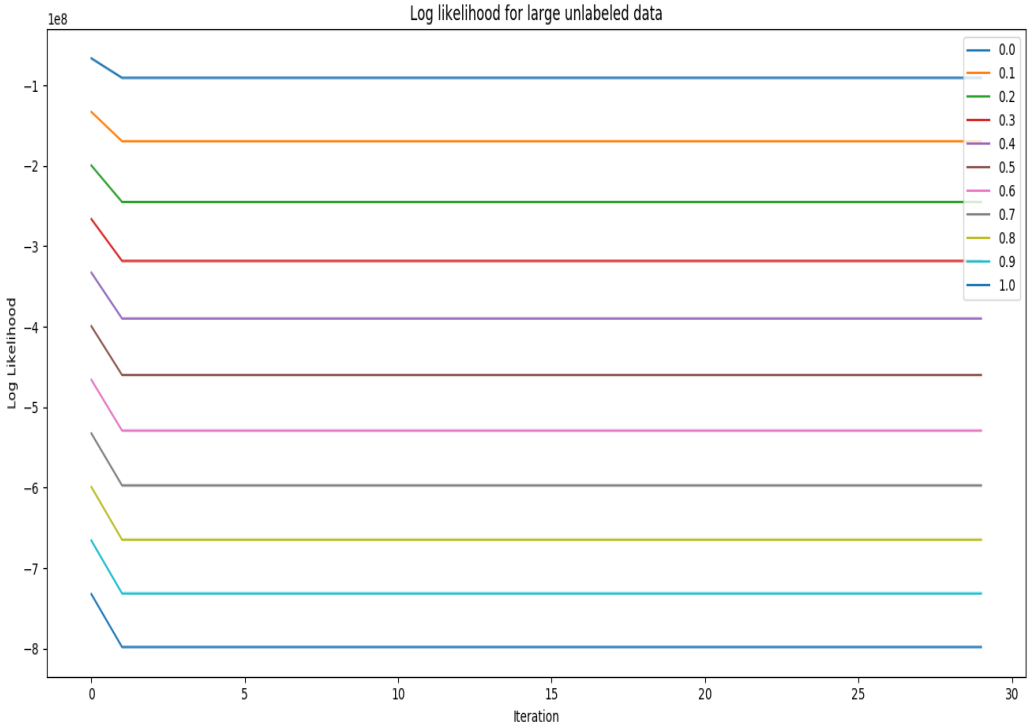
Q4:



Project Results




Log likelihood in 30 Loops
−500000
−1000000
−1500000
−2000000
−2500000
−3000000
−3500000
−4000000
0
5
10
15
20
25
30
Iteration
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0


Log likelihood for large unlabeled data
1e8
−1
−2
−3
−4
−5
−6
−7
−8
Log Likelihood
0
5
10
15
20
25
30
Iteration
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1.0

Note:

1. In the likelihood graph, each line seems like parallel. However, if looking at data carefully, you will find very small difference between data after second iteration. If there still don't have any difference 10 numbers after dot, that because the results have converged.

2. The likelihood for unlabeled large data is really hard to collect. It toke me 6 hours to complete the all 11*30 loops. To improve the efficiency, parallel computing, distribute computing or cloud computing must be the better approach.