1 EM

1.1 E Step

For E step and just for now, all the probabilities P(hidden | observed) are computed using marginalization.

1.2 M Step

For E step we want to find

$$\underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ p(h, y_k|x_k) \ )$$

which can be formulated as

$$\underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(h, y_k, x_k) \ ) - \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(x_k) \ )$$

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(h, y_k, x_k) \ ) - \sum_{k=1}^{N} log(\ b(x_k) \ ) \sum_{h} q(h)$$

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(h, y_k, x_k) \ ) - \sum_{k=1}^{N} log(\ b(x_k) \ )$$

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(h, y_k, x_k) \ ) - \sum_{k=1}^{N} log(\sum_{h', y'} \ b(h', y', x_k) \ )$$

Simplifying it "Assuming nothing will change"

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ b(h, y_k, x_k) \ ) - \sum_{k=1}^{N} \sum_{h', y'} log(\ b(h', y', x_k) \ )$$

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ log(\ e^{\theta_{(h, y_k, x_k)}} \ ) - \sum_{k=1}^{N} \sum_{h', y'} log(\ e^{\theta_{(h', y', x_k)}} \ )$$

Where $\theta_{(h,y,x)} = \sum_{\theta_{i,j}} \theta_{i,j} \qquad such that \qquad i, j \in h \cup y \cup x$

$$= \underset{\theta}{\text{argmax}} \quad \sum_{k=1}^{N} \sum_{h} q(h) \ \theta_{(h, y_k, x_k)} - \qquad \sum_{k=1}^{N} \sum_{h', y'} \theta_{(h', y', x_k)}$$