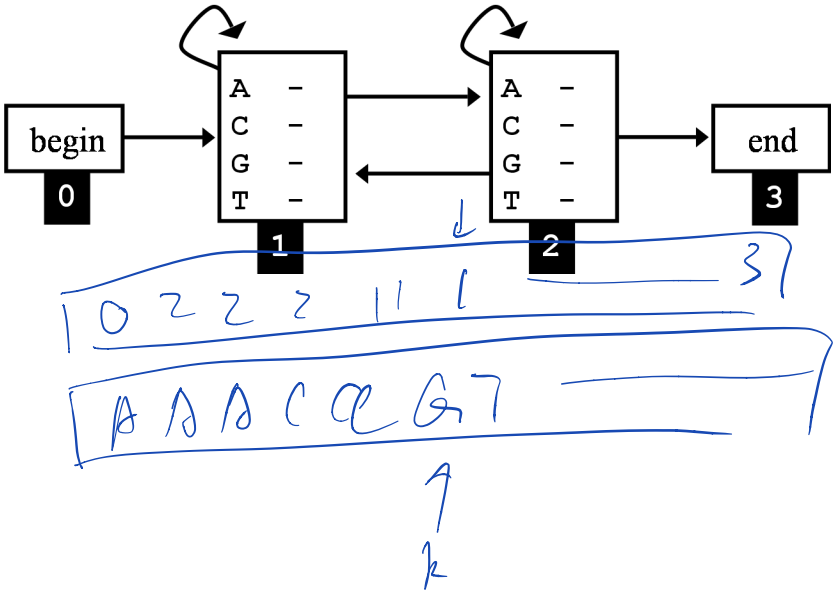
1. Parameter Learning in Hidden Markov Models: Consider a situation in which we are learning the parameters of the HMM shown below. We are given training sequences for which the correct state is unknown (i.e. hidden) for most characters in the sequence. However, for some sequences we know that the k^{th} character must be emitted by *state 1*. Describe how we should adjust standard Baum-Welch learning to handle this situation. Be specific in your description.



2. *K*-Means Clustering: Show how *k*-means would cluster the following genes represented as vectors:

$x_1 = \langle 3, 1 \rangle$
$x_2 = \langle 3, 3 \rangle$
$x_3 = \langle 5, 2 \rangle$
$x_4 = \langle 7, 8 \rangle$

Assume that $k = 2$, the initial coordinates of the cluster centers are $\langle 2, 2 \rangle$ and $\langle 6, 3 \rangle$ and we are using Manhattan distance: $\text{dist}(x_i, x_j) = \sum_e | x_i[e] - x_j[e] |$

$C_1 = (3,1), (3,3) \Rightarrow (3,2) \Rightarrow (3,1)(3,3),(5,2)$

$C_2 = (5,2), (7,8) \Rightarrow (6,5) \Rightarrow (7,8)$

$\Rightarrow \left(\frac{11}{3}, 2\right)$

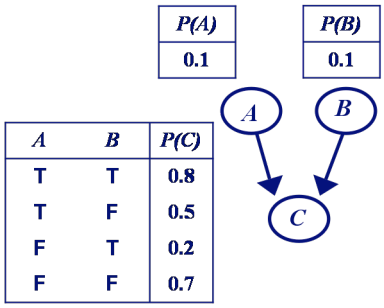
$\Rightarrow (7,8)$

3. EM Clustering (10 points): Consider using EM clustering to cluster instances that are represented three by Boolean values:

$x_1 = \langle \neg a, b, \neg c \rangle$
$x_2 = \langle \neg a, b, c \rangle$
$x_3 = \langle a, \neg b, \neg c \rangle$
…

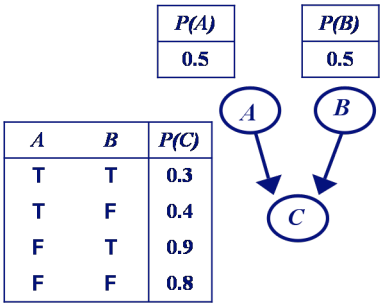
Suppose that we have two clusters, and each represented by a Bayes net as shown below.

cluster 1



A	B	P(C)
T	T	0.8
T	F	0.5
F	T	0.2
F	F	0.7

cluster 2



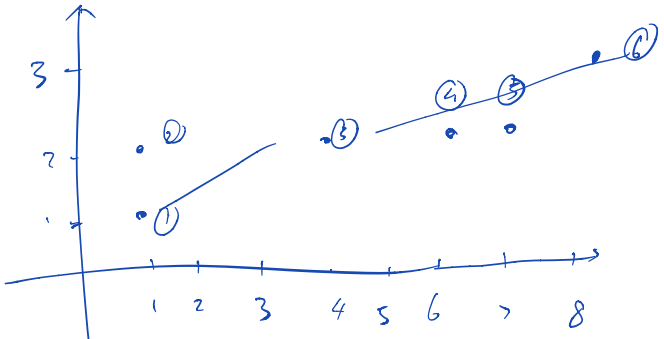
A	B	P(C)
T	T	0.3
T	F	0.4
F	T	0.9
F	F	0.8

(a) Show the calculations that would be done in the E-step for the three instances shown.

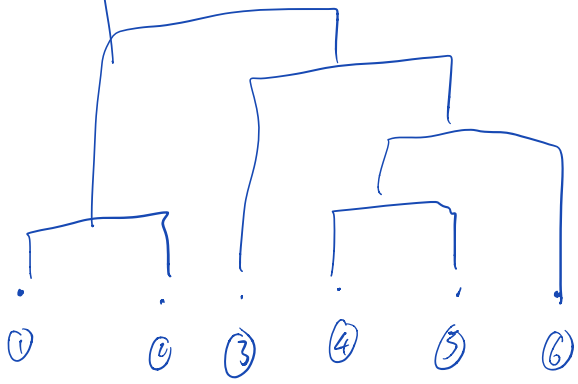
(b) Briefly describe how you would do the M step.

4. Hierarchical Clustering: Given the following set of two-dimensional vectors, show the hierarchical clusterings that would be produced by the single-link and complete-link methods using Euclidean distance.

$x_1 = \langle 1, 1 \rangle$
$x_2 = \langle 1, 2 \rangle$
$x_3 = \langle 4, 2 \rangle$
$x_4 = \langle 6, 2 \rangle$
$x_5 = \langle 7, 2 \rangle$
$x_6 = \langle 8, 3 \rangle$



Single :



Complete