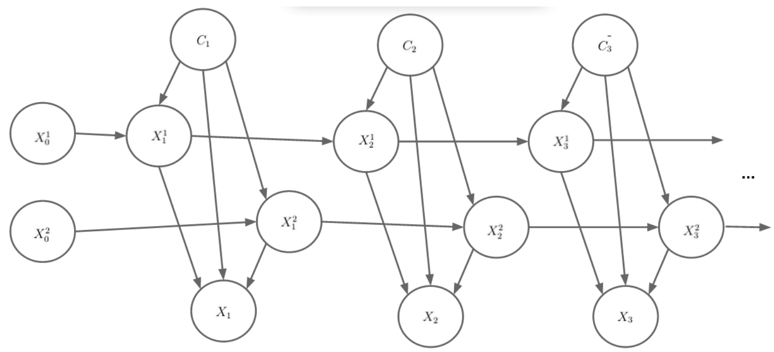
CS 5876 - HW 3

Joshua Campbell (jac677) , Michael Wang (mzw4) , Neil Parker (nwp7)

May 5, 2015

Q1)

1) The parameters of the model are:
   - μ_1 - The first red-apple tree location ($\mu_1 \in \mathbb{R}^2$)
   - μ_2 - The first green-apple tree location ($\mu_2 \in \mathbb{R}^2$)
   - σ - The parameter defining variance (σ^2) for both $\mu_1 and \mu_2$
   - π - A mixture distribution over the $K = 2$ types of trees
2) There O(2N) nodes and $O(N^2)$ edges. Each time step is independent of others, in terms of selecting which type of tree to sprout. However, choosing the location depends on this choice as well as all the previous tree locations and previous tree type choices, since there is no indication (in this model) of which location was the most recent location of the corresponding tree choice. Thus, at each time step, the location at the time step will have one edge from the tree choice at that time test, and an edge from all the previous tree choices and tree locations.
3) There are 7(N-1) edges in the model, where N is the number of iterations. The states looks like below. Each X_t^i depends on X_{t-1}^i and C_t. X_t depends on C_t to choose which value of X_t^i to assume. Each C_t is independent and thus does not depend on any of the variables.



4) $P(C_t = i|\text{parents}) = \pi$
$P(X_t|C_t = i, X_t^1, X_t^2) = 1$
$P(X_t^{(i)}|C_t, X_{t-1}^{(i)}) = \mathbb{1}_{i=C_t} N(X_{t-1}^{(i)}, \sigma^2 I) \,+\, \mathbb{1}_{i \neq C_t}$
$P(X_0^1) = N(\mu_1, \sigma^2 I)$
$P(X_0^2) = N(\mu_2, \sigma^2 I)$

5) For variable elimination (and also the order we will eliminate variables):
Eliminate $X_0^1 : m_{1,\theta}(X_0^1, X_1^1, C_1) = \int_{X_0^1} P(X_0^1)P(X_1^1|C_1, X_0^1)$
Eliminate $X_0^2 : m_{2,\theta}(X_0^2, X_1^2, C_1) = \int_{X_0^1} P(X_0^2)P(X_1^2|C_1, X_0^2)$
Eliminate $X_1^1 : m_{3,\theta}(X_1^1, X_2^1, C_2, C_1) = \int_{X_1^1} P(X_1^1)P(X_2^1|C_2, X_1^1)m_{1,\theta}$
Eliminate $X_1^2 : m_{4,\theta}(X_1^2, X_2^2, C_2, C_1) = \int_{X_1^2} P(X_1^2)P(X_2^2|C_2, X_1^2)m_{2,\theta}$
Eliminate $X_2^1 : m_{5,\theta}(X_2^1, X_3^1, C_3, C_2, C_1) = \int_{X_2^1} P(X_2^1)P(X_3^1|C_3, X_2^1)m_{3,\theta}$
Eliminate $X_2^2 : m_{6,\theta}(X_2^2, X_3^2, C_3, C_2, C_1) = \int_{X_2^2} P(X_2^2)P(X_3^2|C_3, X_2^2)m_{4,\theta}$
Repeat for all X_t^i's, defining up to $m_{2N+2,\theta}$.
Eliminate $C_1 : k_{1,\theta}(C_1, C_2, C_3, ...) = \sum_{C_1} m_{2N+2,\theta}$
Eliminate $C_2 : k_{2,\theta}(C_2, C_3, ...) = \sum_{C_2} k_{1,\theta}$
Eliminate $C_3 : k_{3,\theta}(C_3, C_4, ...) = \sum_{C_3} k_{2,\theta}$
Repeat for all C_t's, defining up to $k_{N-1,\theta}$
Note that all X_t's are deterministic.

Q2) 1) The hidden variables are the set $\{C_t\}$. The set of observed variables are the set $\{X_t\}$.

2) $\log(P_\theta(O, H)) = \log \prod_{t=1}^{N} P(C_t)P(X_t|X_t^1, X_t^2, C_t)P(X_t^1|C_t, X_{t-1}^1)P(X_t^2|C_t, X_{t-1}^2)$
$\log(P_\theta(O, H)) = \log \prod_{t=1}^{N} P(C_t)P(X_t^1|C_t, X_{t-1}^1)P(X_t^2|C_t, X_{t-1}^2)$
$= \sum_{t=1}^{N} \{log(P(C_t)) + log(P(X_t^1|C_t, X_{t-1}^1)\} + log(P(X_t^2|C_t, X_{t-1}^2)\}$

3) $\sum_H Q^i(H) log P_\theta(O,H) =$
$\sum_{t=1}^N \sum_H \{Q^i(H)(log(P(C_t)) + log(P(X_t^1|C_t, X_{t-1}^1)\} + log(P(X_t^2|C_t, X_{t-1}^2))\}$

$Q^i(H) = P(C_1 = h, C_2 = j, C_3 = k, ...)$
$= \prod_{t=1}^N P(C_i = j) = \prod_{t=1}^N Q_t^i(H_t)$

$Q_t^i(C_t = k) = P(C_t = k|X_t, \theta^{(i-1)}) =$
$P(X_t|C_t = k, \theta^{(i-1)})P(C_t = k|\theta^{(i-1)}) = P(C_t = k|\theta^{(i-1)})$

Thus, $\sum_H Q^i(H) log P_\theta(O,H) =$
$\sum_{t=1}^N \sum_{C_1,..,C_t} \{P(C_1, ..., C_t)(log(P(C_t)) + log(P(X_t^1|C_t, X_{t-1}^1)\} +$
$log(P(X_t^2|C_t, X_{t-1}^2))\}$