CS 228, Winter 2011-2012
Problem Set #2

This assignment is due at 12 noon on Feburary 6. Submissions should be placed in the filing cabinet labeled CS228 Homework Submission Box located in the lobby outside Gates 187.

Suppose we wish to perform exact inference over a chain Markov Random Field given by $X_1 - X_2 - \cdots - X_n$. We construct a clique tree of the form $\boldsymbol{C}_1 - \cdots - \boldsymbol{C}_{n-1}$ where $Scope[\boldsymbol{C}_i] = \{X_i, X_{i+1}\}$, and calibrate it, resulting in a calibrated clique tree $\mathcal{T}$ over unnormalized factors $\mathcal{F}$. Assume that each variable X_i has $|Val(X_i)| = d$.

a) [**5 points**] Describe how to compute the marginal of a single variable, $P_\Phi(X_i)$ for a given X_i using the calibrated clique tree. What is the running time cost of this operation? Does the result depend on which clique we use to extract the marginal from? Why?

Now, consider the problem of answering a query $P(X_1, X_3)$ in our calibrated clique tree. We cannot extract this marginal from any clique in $\mathcal{T}$ because the query variables are not within the scope of any of our cliques. It would be costly to have to construct a different clique tree with a clique containing X_1 and X_3 and to re-calibrate it. Fortunately, the query can be done by performing variable elimination over a subtree of the calibrated clique tree, which is shown in Algorithm 1 (this is Algorithm 10.4 in the textbook). We can see from the algorithm that we need only perform variable elimination over the portion of the tree that contains the variables that constitute our query, which greatly accelerates the process compared to the naive method of running variable elimination on the entire graph.

Algorithm 1 Out-of-clique inference in clique tree

Procedure CTree-Query (
  $\mathcal{T}$, // Clique tree over Φ
  $\{\pi_i\}, \{\mu_{i,j}\}$, // Calibrated clique and sepset beliefs for $\mathcal{T}$
  $\boldsymbol{Y}$ // A query
)
  Let $\mathcal{T}'$ be a subtree of $\mathcal{T}$ such that $\boldsymbol{Y} \subseteq Scope[\mathcal{T}']$
  Select a clique $r \in \mathcal{V}_{\mathcal{T}'}$ to be the root
  $\Phi \leftarrow \pi_r$
  for each $i \in \mathcal{V}_{\mathcal{T}'} - \{r\}$
    $\phi \leftarrow \frac{\pi_i}{\mu_{i,p_r(i)}}$
    $\Phi \leftarrow \Phi \cup \{\phi\}$
  $\boldsymbol{Z} \leftarrow Scope[\mathcal{T}'] - \boldsymbol{Y}$
  Let $\prec$ be some ordering over $\boldsymbol{Z}$
  return Sum-Product-VE$(\Phi, \boldsymbol{Z}, \prec)$

b) [**5 points**] What is the total running time complexity of computing marginals over all $\binom{n}{2}$ variable pairs, which we can do by running Algorithm 1 for each pair? Describe the time complexity in terms of n and d.

Note: In line 5 of Algorithm 1, the function $p_r(i)$ is the parent of node i in the tree rooted at r.

c) **[5 points]** Since we are computing marginals for all variable pairs, we can store computations done for the previous pairs and use them to save time for the computations of the remaining pairs. The key recursive relationship that makes this work is the following equation (for $i < j - 1$):

$$P(X_i, X_j) \quad = \quad \sum_{X_{j-1}} P(X_i, X_{j-1}) P(X_j \mid X_{j-1})$$

Prove that the equation above holds.

d) **[10 points]** Construct a dynamic programming algorithm that computes marginals over all $\binom{n}{2}$ variable pairs based on the recursive relation in part (c), and achieves a running time that is asymptotically at least n or d times faster than part (b). Describe the time complexity of your algorithm in terms of n and d.

Note: Make sure to clearly specify how each of the probabilities P you use are computed.