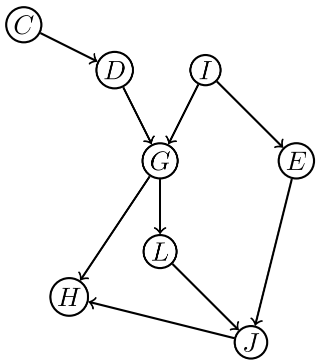
Probabilistic Graphical Models: Problem Set 6 (Solutions)

Svante Linusson, Liam Solus
KTH Royal Institute of Technology

28 September 2021

1. Consider the DAG $\mathcal{G}$ associated with a student's performance in some course:



$\mathcal{G}$

$C =$	Coherence
$D =$	Difficulty
$I =$	Intelligence
$G =$	Grade
$E =$	Exam score
$L =$	Letter of Recommendation
$J =$	Job
$H =$	Happy

Construct a clique tree for this example with respect to an instantance of variable elimination where Φ is the set of conditional probabilities $p(x_i \mid x_{\mathrm{pa}_{\mathcal{G}}(i)})$ associated to $\mathcal{G}$, $\mathbf{Z} = \{C, D, I, G, E, L, H\}$ and $\prec$ is the ordering (C, D, I, H, G, E, L).

Solution:
We first run variable elimination on this DAG for the set of factors

$$\Phi = \{P(C), P(D \mid C), P(I), P(G \mid D, I), P(E \mid I), P(L \mid G), P(J \mid E, L), P(H \mid G, J)\},$$

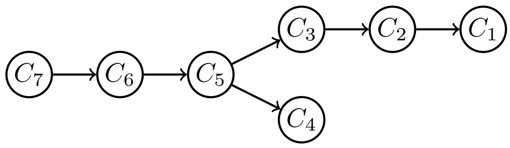
and the ordering $\prec\ = (C, D, I, H, G, E, L)$. We then compute

(a) $\psi_1(C, D) = P(C)P(D \mid C)$, $\tau_1(D) = \sum_C \psi_1(C, D)$,
(b) $\psi_2(D, G, I) = \tau_1(D)P(G \mid D, I)$, $\tau_2(G, I) = \sum_D \psi_1(D, G, I)$,
(c) $\psi_3(G, E, I) = \tau_2(G, I)P(I)P(E \mid I)$, $\tau_3(G, E) = \sum_I \psi_3(G, E, I)$,
(d) $\psi_4(H, G, J) = P(H \mid G, J)$, $\tau_4(G, J) = \sum_H \psi_4(H, G, J)$,
(e) $\psi_5(G, J, E, L) = \tau_4(G, J)\tau_3(G, E)P(L \mid G)$, $\tau_5(J, E, L) = \sum_G \psi_5(G, J, E, L)$,
(f) $\psi_6(J, E, L) = \tau_5(J, E, L)P(J \mid E, L)$, $\tau_6(J, L) = \sum_E \psi_6(J, E, L)$,
(g) $\psi_7(J, L) = \tau_6(J, L)$, $\tau_7(J)$.

We then create one clique for each ψ_i consisting of the variables in Scope$[\psi_i]$:

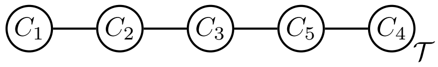
(a) $C_1 = \{C, D\}$,
(b) $C_2 = \{D, G, J\}$,
(c) $C_3 = \{G, E, I\}$,
(d) $C_4 = \{H, G, J\}$,
(e) $C_5 = \{G, J, E, L\}$,
(f) $C_6 = \{J, E, L\}$, and
(g) $C_7 = \{J, L\}$.

We then create a clique tree by connecting cliques C_i and C_j by edges $C_i \leftarrow C_j$ if ψ_j is defined using τ_i. Here we include already the directions, but the clique tree could be taken as undirected and the choice of edges specifies that our root node is C_7:



---

2. Let Φ denote the set of factors used in problem 1. Consider the clique tree $\mathcal{T}$



where

$$C_1 = \{C, D\}, C_2 = \{D, I, G\}, C_3 = \{G, I, E\}, C_4 = \{G, H, J\}, C_5 = \{G, J, L, E\},$$

and we assign each factor in $\phi \in \Phi$ a clique $\alpha(\phi)$ as

$$\alpha(p(C)) = C_1, \alpha(p(D \mid C)) = C_1, \alpha(p(G \mid I, D)) = C_2, \alpha(p(I)) = C_3, \alpha(p(E \mid I)) = C_3,$$
$$\alpha(p(H \mid G, J) = C_4, \alpha(p(L \mid G)) = C_5, \text{ and } \alpha(p(J \mid L, E)) = C_5.$$

(a) What are the initial potentials of $\mathcal{T}$?

(b) Fix the root C_r to be C_5 and perform sum-product message passing to compute $\beta_r(C_r)$.

(c) Fix the root C_r to be C_3 and perform sum-product message passing to compute $\beta_r(C_r)$. Were you able to reuse anything from parts (a) or (b) in this computation?

Solution:

a) The initial potential ψ_i for the clique C_i is given by multiplying all factors in Φ that have been assigned to C_i. Hence, they are

(a) $\psi_1 = P(C)P(D \mid C)$,

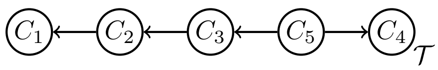
(b) $\psi_2 = P(G \mid I, D)$,

(c) $\psi_3 = P(I)P(E \mid I)$,

(d) $\psi_4 = P(H \mid G, J)$, and

(e) $\psi_5 = P(L \mid G)P(J \mid L, E)$.

b) We fix the root $C_r = C_5$ and orient our clique tree accordingly:



The sepsets labeling the edges are $S_{1,2} = \{D\}, S_{2,3} = \{G, I\}, S_{3,5} = \{G, E\}, S_{4,5} = \{G, J\}$. The message $\delta_{i \to j}$ is computed by summing out the variables in C_i that are not in the sepset $S_{i,j}$ labeling the edge between C_i and C_j. These messages are

(a) $\delta_{1\to 2}(D) = \sum_C \psi_1(C, D)$,

(b) $\delta_{2\to 3}(G, I) = \sum_D \psi_2(G, I, D)\delta_{1\to 2}(D)$,

(c) $\delta_{3\to 5}(G, I) = \sum_I \psi_3(E, I)\delta_{2\to 3}(G, I)$, and

(d) $\delta_{4\to 5}(G, J) = \sum_H \psi_4(G, H, J)$.

Hence, the beliefs $\beta_5(G, J, L, E)$ are

$$\beta_5(G, J, L, E) = \psi_5(G, J, L, E)\delta_{3\to 5}(G, I)\delta_{4\to 5}(G, J).$$

c) Now we fix the root $C_r = C_3$:

$$C_1 \leftarrow C_2 \leftarrow C_3 \rightarrow C_5 \rightarrow C_4 \; \mathcal{T}$$

If we computed the messages that get passed along the arrows that did not get reversed we would get exactly the same messages as in part (b). Hence, we only need to compute the messages of the arrows that changed direction, of which there is exactly one:

(a) $\delta_{5\to 3}(G, E) = \sum_{J,L} \psi_5(G, J, L, E)\delta_{4\to 5}(G, J)$.

Hence the beliefs $\beta_3(G, I, S)$ are

$$\beta_3(G, I, E) = \psi_3(G, I, E)\delta_{2\to 3}(G, I)\delta_{5\to 3}(G, E).$$

We were able to re-use all but one message from part (b), so we only needed to compute the new message $\delta_{5\to 3}(G, E)$ and the beliefs $\beta_3(G, I, E)$.

---

3. The clique tree we constructed in problem (1) has more nodes than the clique tree we used for inference in problem (2). In particular, the clique tree in problem (2) used only the maximal cliques from the tree in problem (1). We now show that this is always possible:

   Let $\mathcal{T}$ be a clique tree over a set of factors Φ. Show that there exists a clique tree $\mathcal{T}'$ such that

   (1) each clique in $\mathcal{T}'$ is also a clique in $\mathcal{T}$, and

   (2) there is no pair of distinct cliques C_i, C_j in $\mathcal{T}'$ such that $C_j \subset C_j$.

4. Consider the Hidden Markov Model from Lecture 6 where $m = 6$ and $n = 4$.

   (a) Use the clique-tree algorithm to compute the marginal distributions

   $$f_{Y_1}(y_1), \quad f_{Y_2}(y_2), \quad f_{Y_3}(y_3), \quad f_{Y_4}(y_4), \quad \text{and } f_{Y_5}(y_5).$$

   (b) We observe the random sample $\mathbf{z} = [\mathbf{y}_1, \mathbf{y}_2, \mathbf{y}_3, \mathbf{y}_4]^T$ from $\mathbf{Y} = [Y_1, Y_2, Y_3, Y_4, Y_5]^T$:

   $$\begin{aligned} \mathbf{y}_1 &= [2, 4, 4, 2, 3]^T, \\ \mathbf{y}_2 &= [1, 0, 1, 3, 3]^T, \\ \mathbf{y}_3 &= [2, 2, 3, 4, 2]^T, \\ \mathbf{y}_4 &= [1, 3, 2, 1, 3]^T. \end{aligned}$$

   Compute the marginal posteriors

   $$f_{X_1|\mathbf{Z}}(x_1|\mathbf{z}), \quad f_{X_2|\mathbf{Z}}(x_2|\mathbf{z}), \quad f_{X_3|\mathbf{Z}}(x_3|\mathbf{z}), \quad f_{X_4|\mathbf{Z}}(x_4|\mathbf{z}), \quad \text{and } f_{X_5|\mathbf{Z}}(x_5|\mathbf{z}).$$

Solution:

Since the conditional probabilities needed to define the joint distribution are specified by the choice of $m = 6$ and $n = 4$, one can implement the clique-tree algorithm to compute the desired marginals in part (a). For part (b), we can apply VE as in problem 5 from the previous practice session.
