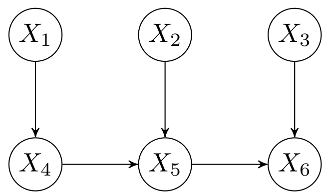


Figure 1: Bayesian network 1

1 Exercises about Variable Elimination Algorithm

Exercise 1.1 *Intermediate Factors. Consider running variable elimination on the Bayesian network of Figure 1. Which of the nodes, if eliminated first, results in the largest intermediate factor? The largest factor is that with the largest number of entries.*

a) X_1

b) X_4

c) X_5

d) X_2

Exercise 1.2 *Elimination Orderings. Which of the following characteristics of the variable elimination algorithm are affected by the choice of elimination ordering? You may select 1 or more options, or none of them.*

a) Which marginals can be computed correctly

b) Size of the largest intermediate factor

c) Ability to handle evidence

d) Runtime of the algorithm

e) Memory usage of the algorithm

Exercise 1.3 *Uses of Variable Elimination. Which of the following quantities can be computed using the sum-product variable elimination algorithm? You may select 1 or more options, or none of them.*

a) $p(X|E = e)$ *in a Bayesian network*

b) $p(X)$ *in a Bayesian network*

c) The most likely assignment to the variables in a Markov network.

d) The partition function for a Markov network

Exercise 1.4 *Marginalization. Suppose we run variable elimination on a Bayesian network where we eliminate all the variables in the network. Which number will the algorithm produce?*

Exercise 1.5 *Marginalization. Suppose we run variable elimination on a Markov network where we eliminate all the variables in the network. Which number will the algorithm produce?*

a) 1

b) A positive number, always between 0 and 1, which depends on the structure of the network.

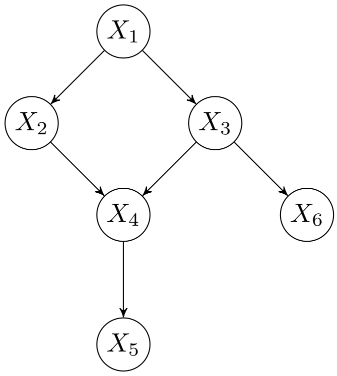


Figure 2: Bayesian network 2

c) Z, *the partition function for the network.*

d) $1/Z$, *where* Z *is the partition function for the network.*

Exercise 1.6 *Intermediate Factors. If we perform variable elimination on the graph of Figure 2 with the variable ordering* $X_2, X_1, X_3, X_6, X_5, X_4$, *which is the intermediate factor produced by the third step (just before summing out* X_3*)?*

a) $\psi(X_1, X_2, X_3, X_4, X_6)$

b) $\psi(X_3, X_4, X_5, X_6)$

c) $\psi(X_3, X_4, X_6)$

d) $\psi(X_4, X_6)$

Exercise 1.7 *Intermediate Factors. If we perform variable elimination on the graph of Figure 2 with the variable ordering* $X_6, X_5, X_4, X_3, X_2, X_1$, *which is the intermediate factor produced by the third step (just before summing out* X_4*)?*

a) $\psi(X_2, X_3, X_4, X_5, X_6)$

b) $\psi(X_2, X_3, X_4, X_5)$

c) $\psi(X_3, X_4, X_5)$

d) $\psi(X_2, X_3, X_4)$

e) $\psi(X_3, X_4)$

f) $\psi(X_2, X_3)$

Exercise 1.8 *Induced graphs. If we perform variable elimination on the graph of Figure 2 with the variable ordering* $X_2, X_1, X_3, X_6, X_5, X_4$, *what is the induced graph for the run?*

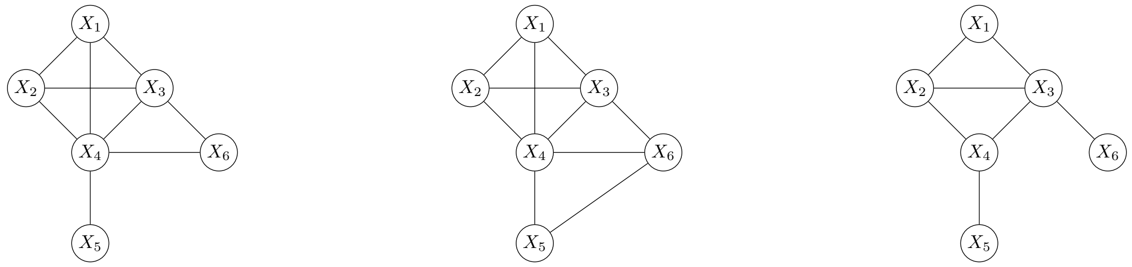
a) Figure 3a

b) Figure 3b

c) Figure 3c

d) None of Figure 3

Exercise 1.9 *Time complexity of Variable Elimination. Consider a Bayesian network taking the form of a chain of variables* $X_1 \to X_2 \to \ldots X_n$, *where each variable* X_i *takes on* k *values. Which is the computational cost of variable elimination on this network if we follow the natural order* $X_1, X_2, \ldots$*?*



(a) Induced graph 1 (b) Induced graph 2 (c) Induced graph 3

Figure 3: Different possible induced graphs for BN of Fig. 2

a) $O(kn^2)$

b) $O(nk^2)$

c) $O(nk^3)$

d) $O(nk)$

Exercise 1.10 *Time complexity of Variable Elimination. Consider again a Bayesian network taking the form of a chain of variables* $X_1 \to X_2 \to \dots X_n$*, where each variable* X_i *takes on* k *values. Which is the computational cost of variable elimination on this network if we follow the order* $X_2, X_3, \dots, X_n, X_1$*?*

a) $O(kn^2)$

b) $O(nk^2)$

c) $O(nk^3)$

d) $O(nk)$

Exercise 1.11 *Time complexity of Variable Elimination. Suppose we eliminate all the variables in a Markov network. What could affect the runtime of the algorithm? You may select 1 or more options, or none of them.*

a) *Number of variables in the network*

b) *Number of values each variable can take*

c) *Whether there is evidence*

2 Exercises about Clique Tree Algorithm

Exercise 2.1 *Message Ordering. In the clique tree of Figure 4, which of the following starting message passing orders is/are valid? (Note: These are not necessarily full sweeps that result in calibration). You may select 1 or more options, or none of them.*

a) $C_1 \to C_2$, $C_2 \to C_3$, $C_3 \to C_4$, $C_3 \to C_5$

b) $C_4 \to C_3$, $C_3 \to C_2$, $C_2 \to C_1$

c) $C_4 \to C_3$, $C_5 \to C_3$, $C_2 \to C_3$

d) $C_1 \to C_2$, $C_2 \to C_3$, $C_5 \to C_3$, $C_3 \to C_4$

Exercise 2.2 *Message Passing in a Clique Tree. In the clique tree of Figure 4, what is the correct form of the message from clique 3 to clique 2,* $\delta_{3\to 2}$*, where* $\psi_i(C_i)$ *is the initial potential of clique* i*?*

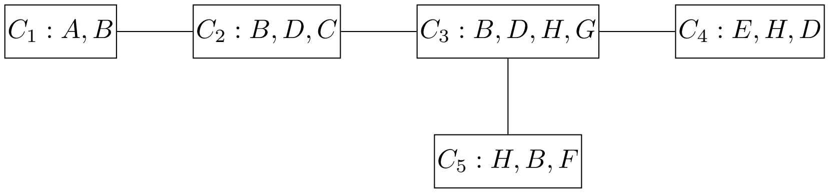


Figure 4: Clique tree 1

(a) Markov network (b) Proposed clique tree

Figure 5: MN and proposed clique tree

a) $\sum_{G,H} \psi_3(C_3) \times \delta_{4\to3} \times \delta_{5\to3}$

b) $\sum_{B,D} \psi_3(C_3) \times \delta_{4\to3} \times \delta_{5\to3}$

c) $\sum_{B,D,G,H} \psi_3(C_3) \times \delta_{4\to3} \times \delta_{5\to3}$

d) $\sum_{G,H} \psi_3(C_3) \times \delta_{2\to3}$

Exercise 2.3 *Clique Tree properties. Consider the Markov Network of Figure 5a over potentials* $\phi_{A,B}, \phi_{B,C}, \phi_{A,D}, \phi_{B,E}, \phi_{C,F}, \phi_{D,E}$ *and* $\phi_{E,F}$*. Which of these properties are necessary for a valid clique tree for the network of Figure 5a, but are **not** satisfied by the graph of Figure 5b? You may select 1 or more options, or none of them.*

a) No loops

b) Running intersection

c) Family preservation

d) Node degree lower than or equal to 2.

e) It is a valid clique tree.

Exercise 2.4 *Cluster Graphs vs. Clique Trees. Suppose that we ran sum-product message passing on a cluster graph G for a Markov network M and that the algorithm converged. Which of the following statements is true only ifG is a clique tree and is not necessarily true otherwise?*

a) All the options are true for cluster graphs in general.

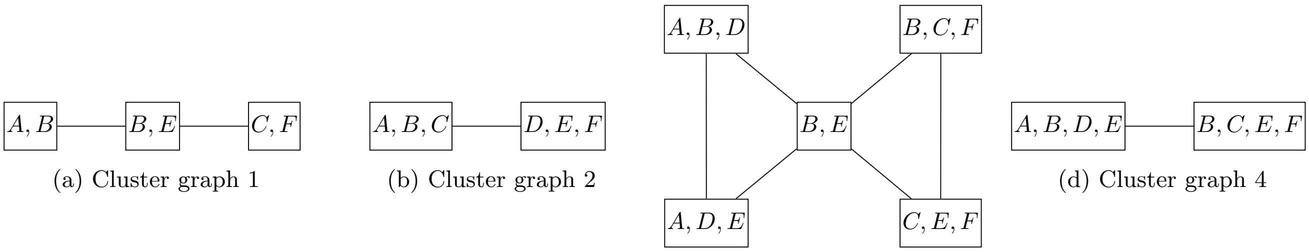
b) If there are E edges in G, there exists a message ordering that guarantees convergence after passing 2E messages.

c) The sepsets in G are the product of the two messages passed between the clusters adjacent to the sepset.

d) G is calibrated.

e) The beliefs and sepsets of G can be used to compute the joint distribution defined by the factors of M.

Exercise 2.5 *Clique Tree calibration. Which of the following is true? You may select 1 or more options, or none of them.*



(a) Cluster graph 1 (b) Cluster graph 2 (c) Cluster graph 3 (d) Cluster graph 4

Figure 6: Possible cluster graphs for MN in Figure 5a

a) If a clique tree is max-calibrated, then all pairs of cliques are max-calibrated.

b) If there exists a pair of adjacent cliques that are max-calibrated, then a clique tree is max-calibrated.

c) After we complete one upward pass and one downward pass of the max-sum message passing algorithm, the clique tree is max-calibrated.

d) If a clique tree is max-calibrated, then within each clique, all variables are max-calibrated with each other.

3 Exercises about Message Passing Algorithm in Cluster Graphs

Exercise 3.1 *Family Preservation. Suppose we have a factor $P(A|C)$ that we wish to include in our sum-product message passing inference. We should:*

a) Assign the factor to all cliques that contain A or C

b) None of these

c) Assign the factor to one clique that contain A and C

d) Assign the factor to all cliques that contain A and C

Exercise 3.2 *Cluster Graph construction. Consider again the Markov Network of Figure 5a over potentials $\phi_{A,B}, \phi_{B,C}, \phi_{A,D}, \phi_{B,E}, \phi_{C,F}, \phi_{D,E}$ and $\phi_{E,F}$. Which is valid cluster graph for H? You may select 1 or more options, or none of them.*

a) Figure 6a

b) Figure 6b

c) Figure 6c

d) Figure 6d

Exercise 3.3 *Message Passing in a Cluster Graph. Suppose we wish to perform inference in Figure 5a using the Cluster Graph of Figure 7. Assuming that the variables in the sepsets are equal to the intersection of the variables in the adjacent cliques, which expression correctly represents the message $\delta_{3\to6}$ that cluster C_3 sends to cluster C_6 in belief propagation?*

a) $\delta_{3\to6}(E) = \sum_B \phi_{B,E}(B,E) \cdot \delta_{2\to3}(B) \cdot \delta_{4\to3}(B) \cdot \delta_{7\to3}(E) \cdot \delta_{6\to3}(B)$

b) $\delta_{3\to6}(E) = \sum_B \phi_{B,E}(B,E) \cdot \delta_{2\to3}(B) \cdot \delta_{4\to3}(B) \cdot \delta_{7\to3}(E)$

c) $\delta_{3\to6}(B,E) = \phi_{B,E}(B,E) \cdot \delta_{2\to3}(B) \cdot \delta_{4\to3}(B) \cdot \delta_{7\to3}(E)$

d) $\delta_{3\to6}(E) = \sum_B \delta_{2\to3}(B) \cdot \delta_{4\to3}(B) \cdot \delta_{7\to3}(E)$

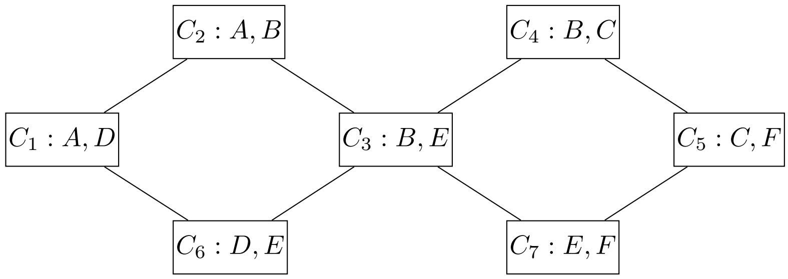


Figure 7: Cluster graph for MN in Figure 5a

X_i	X_j	$\phi(X_i, X_j)$
1	1	10
1	0	1
0	1	1
0	1	10

Table 1: Table of factor ϕ

Exercise 3.4 *Message Passing computation. Consider the Markov network from Figure 5a. If the initial factors in the Markov network are of the form as shown in Table 1, regardless of the specific value of* X_i *and* X_j *(we are encouraging connected variables to share the same assignment), compute the first message* $\delta_{3\to 6}$ *passed during in loopy belief propagation. Assume that the messages are all initialized to 1, i.e. all the entries are initially set to 1.*

Exercise 3.5 *Extracting marginals. Given that you can renormalize the messages at any point during belief propagation and still obtain correct marginals, consider the message that you computed in the previous exercise. Use this observation to compute the (approximate) probability* $p(D = 1, E = 1)$*.*

Answers

Ex. 1.1: c

Ex. 1.2: b,d,e

Ex. 1.3: a,b,d

Ex. 1.4: 1

Ex. 1.5: c

Ex. 1.6: c

Ex. 1.7: d

Ex. 1.8: a

Ex. 1.9: b

Ex. 1.10: c

Ex. 1.11: a,b,c

Ex. 2.1: d

Ex. 2.2: a

Ex. 2.3: a,c

Ex. 2.4: b

Ex. 2.5: a,c

Ex. 3.1: c

Ex. 3.2: c,d

Ex. 3.3: b

Ex. 3.4: 11, 11

Ex. 3.5: 0.45