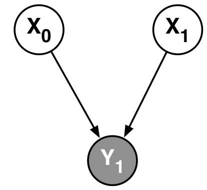
CSE 150A/250A. Assignment 3

Out: *Tues Oct 15*
Due: *Mon Oct 21* (by 11:59 PM, Pacific Time, via gradescope)
Grace period: 24 hours

3.1 Inference in a chain



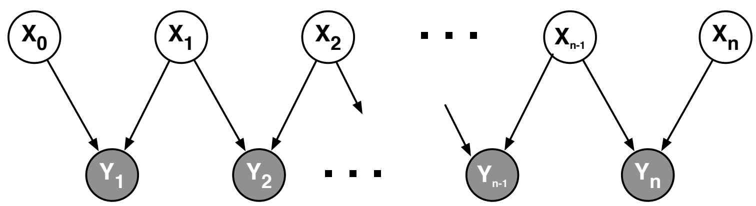
Consider the simple belief network shown to the right, with nodes X_0, X_1, and Y_1. To compute the posterior probability $P(X_1|Y_1)$, we can use Bayes rule:

$$P(X_1|Y_1) \;=\; \frac{P(Y_1|X_1)\,P(X_1)}{P(Y_1)}$$

(a) Show how to compute the conditional probability $P(Y_1|X_1)$ that appears in the numerator of Bayes rule from the CPTs of the belief network.

(b) Show how to compute the marginal probability $P(Y_1)$ that appears in the denominator of Bayes rule from the CPTs of the belief network.

Next you will show how to generalize these computations when the basic structure of this DAG is repeated to form a chain. Like the example we saw in class, this is another instance of efficient inference in polytrees.



Consider how to efficiently compute the posterior probability $P(X_n|Y_1, Y_2, \ldots, Y_n)$ in the above belief network. One approach is to derive a recursion from the *conditionalized* form of Bayes rule

$$P(X_n|Y_1, Y_2, \ldots, Y_n) \;=\; \frac{P(Y_n|X_n, Y_1, Y_2, \ldots, Y_{n-1})\,P(X_n|Y_1, Y_2, \ldots, Y_{n-1})}{P(Y_n|Y_1, \ldots, Y_{n-1})}$$

where the nodes $Y_1, Y_2, \ldots, Y_{n-1}$ are treated as background evidence. In this problem you will express the conditional probabilities on the right hand side of this equation in terms of the CPTs of the network and the probabilities $P(X_{n-1}\!=\!x|Y_1, Y_2, \ldots, Y_{n-1})$, *which you may assume have been computed at a previous step of the recursion*. Your answers to (a) and (b) should be helpful here.

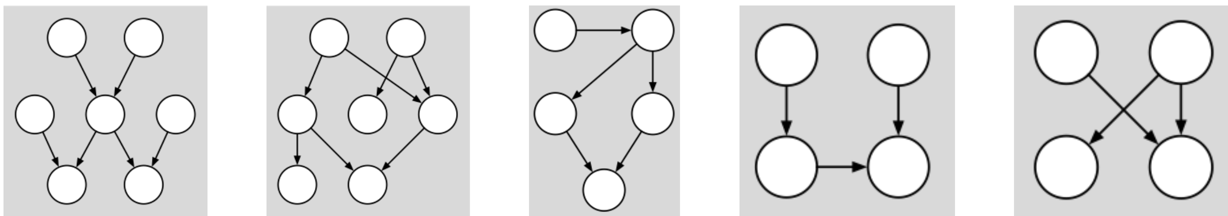
(c) Simplify the term $P(X_n|Y_1, Y_2, \ldots, Y_{n-1})$ that appears in the numerator of Bayes rule.

(d) Show how to compute the conditional probability $P(Y_n|X_n, Y_1, Y_2, \ldots, Y_{n-1})$ that appears in the numerator of Bayes rule. Express your answer in terms of the CPTs of the belief network and the probabilities $P(X_{n-1}\!=\!x|Y_1, Y_2, \ldots, Y_{n-1})$, which you may assume have already been computed.

(e) Show how to compute the conditional probability $P(Y_n|Y_1, Y_2, \ldots, Y_{n-1})$ that appears in the denominator of Bayes rule. Express your answer in terms of the CPTs of the belief network and the probabilities $P(X_{n-1}\!=\!x|Y_1, Y_2, \ldots, Y_{n-1})$, which you may assume have already been computed.

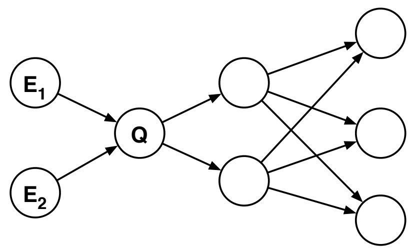
3.2 Node clustering and polytrees

In the figure below, *circle* the DAGs that are polytrees. In the other DAGs, shade **two** nodes that could be *clustered* so that the resulting DAG is a polytree.



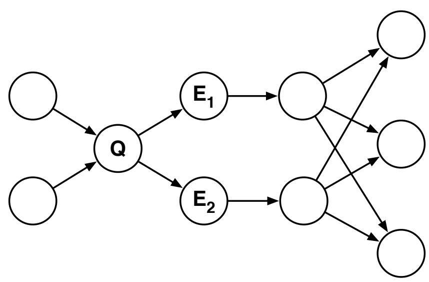
3.3 Cutsets and polytrees

Clearly not all problems of inference are intractable in loopy belief networks. As a trivial example, consider the query $P(Q|E_1, E_2)$ in the network shown below:

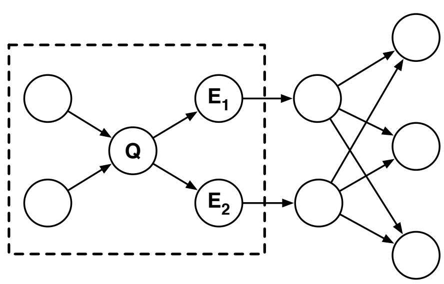


In this case, because E_1 and E_2 are the parents of Q, the query $P(Q|E_1, E_2)$ can be answered directly from the conditional probability table at node Q.

As a less trivial example, consider how to compute the posterior probability $P(Q|E_1, E_2)$ in the belief network shown below:



In this belief network, the posterior probability $P(Q|E_1, E_2)$ can be correctly computed by running the polytree algorithm on the subgraph of nodes that are enclosed by the dotted rectangle:



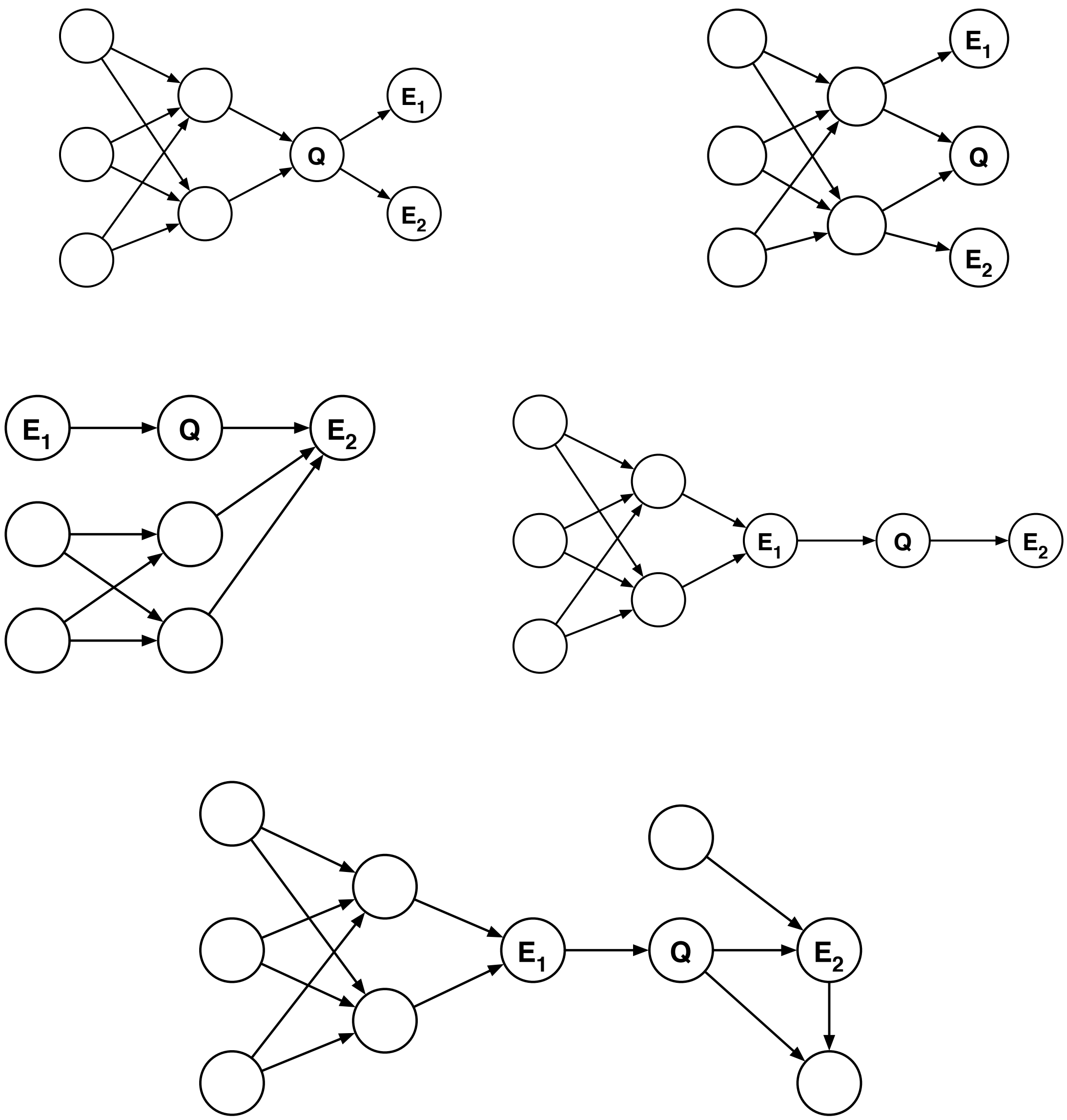
In this example, the evidence nodes d-separate the query node from the loopy parts of the network. Thus for this inference the polytree algorithm would terminate before encountering any loops.

(continued)

For each of the five loopy belief networks shown below, consider how to compute the posterior probability $P(Q|E_1, E_2)$.

If the inference can be performed by running the polytree algorithm on a subgraph, enclose this subgraph by a dotted line as shown on the previous page. (The subgraph should be a polytree.)

On the other hand, if the inference cannot be performed in this way, shade **one** node in the belief network that can be instantiated to induce a polytree by the method of cutset conditioning.

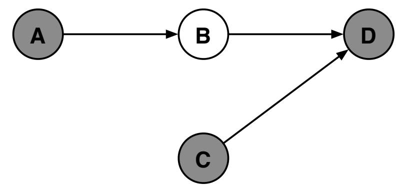


3.4 Even more inference

Show how to perform the desired inference in each of the belief networks shown below. Justify briefly each step in your calculations.

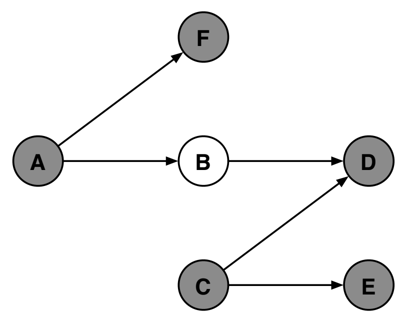
(a) **Markov blanket**

Show how to compute the posterior probability $P(B|A, C, D)$ in terms of the CPTs of this belief network—namely, $P(A)$, $P(B|A)$, $P(C)$, and $P(D|B, C)$.



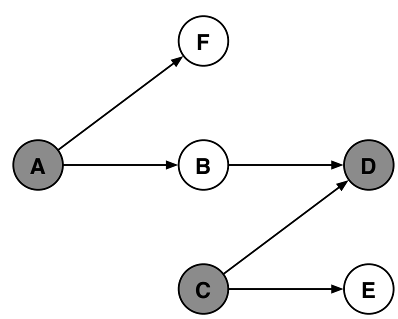
(b) **Conditional independence**

This belief network has conditional probability tables for $P(F|A)$ and $P(E|C)$ in addition to those of the previous problem. Assuming that all these tables are given, show how to compute the posterior probability $P(B|A, C, D, E, F)$ in terms of these additional CPTs and your answer to part (a).



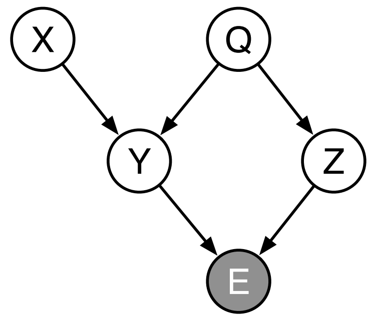
(c) **More conditional independence**

Assuming that all the conditional probability tables in this belief network are given, show how to compute the posterior probability $P(B, E, F|A, C, D)$. Express your answer in terms of the CPTs of the network, as well as your earlier answers for parts (a) and (b).



3.5 More likelihood weighting

(a) **Single node of evidence**



Suppose that T samples $\{q_t, x_t, y_t, z_t\}_{t=1}^{T}$ are drawn from the CPTs of the belief network shown above (with fixed evidence $E = e$). Show how to estimate $P(Q\!=\!q|E\!=\!e)$ from these samples using the method of likelihood weighting. Express your answer in terms of sums over indicator functions, such as:

$$I(q, q') = \begin{cases} 1 & \text{if} \quad q = q' \\ 0 & \text{otherwise} \end{cases}$$

In addition, all probabilities in your answer should be expressed in terms of CPTs of the belief network (i.e., probabilities that do not require any additional computation).

(b) **Multiple nodes of evidence**

Suppose that T samples $\{q_{1t}, q_{2t}, x_t, y_t, z_t\}_{t=1}^{T}$ are drawn from the CPTs of the network shown below (with fixed evidence $E_1\!=\!e_1$ and $E_2\!=\!e_2$). Show how to estimate

$$P(Q_1\!=\!q_1, Q_2\!=\!q_2|E_1\!=\!e_1, E_2\!=\!e_2)$$

from these samples using the method of likelihood weighting. Express your answer in terms of indicator functions and CPTs of the belief network.