Rina Dechter, Spring-2018
Assigned: May 16
Due: May 30

COMPSCI 276: Reasoning with Graphical Models, Problem Set 4 part 1

This homework has 2 parts. The first is on bounded inference and the second is on sampling.

Part 1 Answer questions 1 and 2, and one of 3 and 4 in part 1. The Questions are based on the chapter 8 and 9 on bounded inference.

1. Read chapters 8 and 9 and provide comments.

2. Consider the Bayes network DAG in Figure 1:

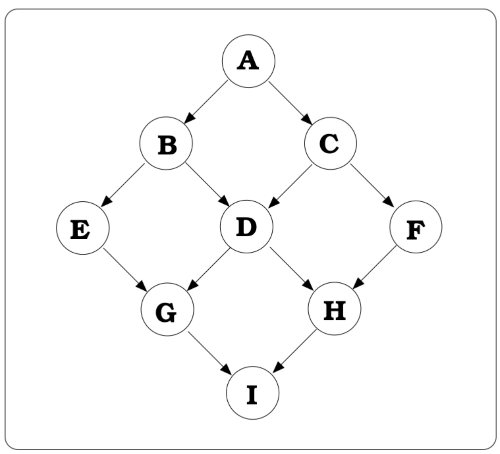


Figure 1: A Bayesian network

   (a) (10) Discuses the performance of the bucket-elimination algorithm for finding the belief of $P(A|I = 0)$. Demonstrate its performance schematically (describe algebraically what function is computed in each bucket). What would be the complexity of the algorithm?

   (b) (10) Apply the approximation algorithm mbe-bel(i=3) for the task of finding the belief in A. Trace the algorithm's performance schematically (show functions, no numbers). What is the time and space complexity of the algorithm?

   (c) (10) Apply mbe-mpe(i=3) to find and upper bound for the mpe of the network given $I = 0$. Trace the algorithms. Show how you construct an approximate mpe tuple.

   (d) (10) Propose a node duplication simplification for the network that will correspond to the mini-bucket scheme you used in your answer above.

(e) (extra credit, 10) Run mini-bucket approximation, and IBP to compare the approximation quality on this network. (You can use any software tool.)

3. (20) Consider the coding networks in the class notes in chapter 8,

   (a) what is the relaxed networks, generated by node duplication, that would correspond to the mini-bucket execution of this example. Draw the relaxed network.

   (b) Apply schematically, mini-clustering to find the belief of each variable. Show the tree-decomposition over which the mini-clustering algorithm executes.

4. (20) Assume you are given a 4x4 directed grid (like in Figure 1 but having dimension 4 rather than 3).

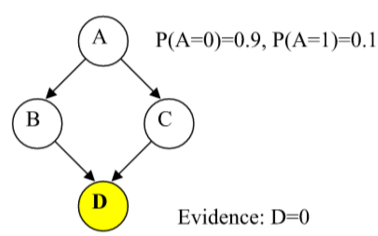
   (a) Provide an arc-labeled minimal dual graph for the 4x4 grid network.

   (b) Generate a join-graph whose maximal cluster size is 4. Show the functions in each cluster and the variables in each cluster. Label the arcs with the appropriate separators.

   (c) Show the schematic messages that would pass in one iteration of IJGP on your join graph.

Part 2: For sampling consult class slides, the "Cutset sampling in Bayesian networks" and Darwiche chapter 15. Answer the following 2 questions. See also Link to the cutset sampling: `https://www.ics.uci.edu/ csp/r137.pdf`

5. (20) Consider a Bayesian network with variables A, B, C, D, and evidence D=0:



A	P(B=0)	P(B=1)
0	0.9	0.1
1	0.1	0.9

A	P(C=0)	P(C=1)
0	0.9	0.1
1	0.1	0.9

BC	P(D=0)	P(D=1)
00	0.99	0.01
01	0.1	0.9
10	0.1	0.9
11	0.05	0.95

   (a) Compute the exact posterior probabilities $P(A|D=0)$, $P(B|D=0)$, $P(C|D=0)$.

(b) Show how to generate a sample using forward sampling. Generate 10 samples using forward sampling. How many samples do you get where D=0 and how many did you have to reject? Estimate posterior probabilities for B, C, A based on those 10 samples?

(c) Show how you would generate a sample using ordered Gibbs sampler. Generate 10 samples and compute the posterior probabilities (beliefs).

(d) Show how you would generate a sample using Gibbs cutset sampling by sampling only cutest B,C. Estimate the posterior probabilities for B, C, A by 1) counting the fraction of samples with desired variable value; 2) using mixture estimator. Which estimate was better? How do Gibbs estimates compare to forward sampling?

(e) Write a few sentences to outline the main trade-offs (the good and the bad) between using forward sampling and Gibbs sampling. Note: You can hard-code the sampling process in the language of your choice or you may perform calculations by hand. Mainly, you need a random number generator that generates the numbers in the range [0, 1] to draw a random value. You can generate a sequence of random numbers using C/C++/Java random number generator or even Excel (just type =rand() in the Excel cell). You can build the network and compute exact posterior marginals as well as necessary sampling probabilities for Gibbs sampler using a Bayesian Network applet:

http://www.aispace.org/bayes/index.shtml (click the button to start the applet at the link above)

6. (20) Consider again the network in Figure 1.

Assume all nodes are bi-valued with domains $D(X) = \{0, 1\}$.

Assume we know the CPT of $P(I|G, H)$ shown in Table 1.

G	H	$P(I = 0\|G, H)$	$P(I = 1\|G, H)$
0	0	1	0
0	1	0.91	0.09
1	0	0.91	0.09
1	1	0.95	0.95

Table 1: CPT of $P(I|G, H)$

(a) Is this network ergodic? Explain.

(b) Assume node I is observed: $\{I = 1\}$. We say that assignment of values to the variables is consistent with evidence if they do not conflict with the evidence values. Assume we want to apply Gibbs sampling to this network and we assigned initial values to the network nodes as follows: $S^0 = A = 0, , G = 0, H = 0$.

Is this instantiation consistent with evidence? Explain.

(c) (extra credit) Assume node I is observed: $\{I = i\}$. Find a cycle-cutset of the network C and trace cutset sampling algorithm (Gibbs sampling on cutset C): Show step-by-step how the first two samples would be computed. Discuss what kind of algorithm you can use to compute $P(C_i|C\ Ci)$ and what the complexity would be.

(d) (10) Given $I = 1$ show how to generate a sample using LW (likelihood weighting)