1. (Extra credit, 5 pts). Read Dechter chapter 5 and provide comments on clarity and typos.

2. (30) Given the directed graph $G$ in Figure 1 (used also in homework 2),
   
   (a) Show the bucket-tree associated with the ordering $d_1 = F, C, A, G, D, H, E, B$ and display all the messages ($\pi$s and $\lambda$s) along the tree.
   
   (b) Assuming you performed all the computation without any evidence. How can you extract the marginal probability of $D$? Explain.
   
   (c) Assuming you observed $F = 1$ and $B = 1$, explain how you would compute (update) $BEL(D) = P(D|F = 1, B = 1)$.
   
   (d) Give a bound on the time and space complexity for solving this problem using $O$ notation.
   
   (e) Assume you have evidence over $F$. Describe how the loop-cutset scheme can find the belief for every variable. What is its time and space complexity?
   
   (f) Assume you compute the beliefs using join-tree clustering. What would be the time and space complexity? Explain.
   
   (g) Suggest an efficient scheme for solving the network without recording more than unary functions. Discuss your proposals.

   \[ \text{Figure 1: A directed graph} \]

3. (20 pts) (revised question 4.5 in Pearl’s book) Consider the network in Figure 2.
   
   (a) What is the dual graph of this network
Figure 2: A two layer network

(b) Find a join-tree representation for the network and show how you would compute $Bel(D_1|M_2 = false, M_3 = true, M_4 = false)$ schematically (demonstrating the type of messages that would be passed).

4. (10) (extra credit) Let $(G,P)$ be a Bayesian network where $G$ is an directed acyclic graphs over variables $X$ and let $C \subseteq X$ be a subset of variables that form a loop-cutset. Prove that $P(C = c)$ can be computed in linear time and space.

5. (10) (extra credit) Which method has better time complexity, the loop-cutset method or join-tree clustering? Prove your claims.

6. (10) Consider the network in Figure 3.

(a) What will be the complexity of loop-cutset conditioning on the network?
(b) How would it compare with BE-bel?
(c) How would it compare with join-tree clustering?

Figure 3: A chain directed graph

7. Consider the Bayes network DAG in Figure 4:

(a) (5) Generate a bucket-tree for this network.
(b) (5) Suppose you want to compute the probability of evidence. What would be the time and space complexity of doing so using Bucket-elimination? a. when the evidence is on variable $E$? b. when the evidence is $I$? c. when you want to compute the belief of $P(A|I = 0)$?

(c) (5) Assume that you have a Markov network grid when the potential are pair-wise (same as the above figure but remove the arrows). Answer the above question relative to such a grid markov network.

(d) (10) Assume you want to solve the problem in Figure 4 using algorithm cycle-cutset conditioning, what is the smallest cycle-cutset you can find (1-cutset) relative to the moral graph? What is the smallest 2-cutset you can find?

(e) (10) Describe how would 2-cutset conditioning can be applied to this problem. What would be the time and memory of 2-cutset conditioning on this problem?