COMPSCI 276
Homework Assignment 3
Spring 2017

Instructor: Rina Dechter Due: Wednesday, April 26

Relevant reading: (Darwiche chapter 5, Dechter chapter 2, Pearl 4.1-4.3).

1. (Darwiche 5.4) We have two sensors that are meant to detect extreme temperature, which occurs 20% of the time. The sensors have identical specifications, with a false positive rate of 1% and a false negative rate of 3%. If the power is o (dead battery), the sensors will read negative regardless of the temperature. Suppose now that we have two sensor kits: Kit A where both sensors receive power from the same battery, and Kit B where they receive power from independent batteries. Assuming that each battery has a 0.9 probability of power availability, what is the probability of extreme temperature given each of the following scenarios:

   (a) two negative sensor readings

   (b) The two sensor readings are positive?

   (c) One sensor reads positive while the other reads negative.

   Answer the previous questions with respect to each kit.

2. (Darwiche 5.6) Lisa is given a fair coin C1 and asked to flip it eight times in a row. Lisa also has a biased coin C2, with a probability 0.8 of landing heads. All we know is that Lisa flipped the fair coin initially, but we believe that she intends to switch to the biased coin, and that she tends to be 10% successful in performing the switch. Suppose that we observe the outcome of the eight coin flips and want to find out whether Lisa managed to perform a coin switch and when. Describe a Bayesian network and a corresponding query that solves this problem. What is the solution to this problem assuming the flips came out as follows:

   (a) tails, tails, tails, heads, heads, heads, heads, heads.

   (b) tails, tails, heads, heads, heads, heads, heads, heads

3. (Pearl 4.1) You can answer the following question manually. Explain the network that you are solving. Alternatively, you can use one of the software tools, but please explain your steps.

   A language L has a four-character vocabulary $V = \{\epsilon, A, B, C\}$ where ϵ is the empty symbol. The probability that character V_i will be followed by V_j is given by the following matrix:

$P(v_j|v_i) =$

v_i \ v_j	ϵ	A	B	C
ϵ	1/4	1/4	1/4	1/4
A	1/2	0	1/4	1/4
B	1/8	1/2	1/8	1/4
C	1/4	1/8	1/2	1/8

In transmitting messages from L, some characters may be corrupted by noise and be confused with others. The probability that the transmitted character v_j will be interpreted as v_k is given by the following confusion matrix:

$P_c(v_k|v_j) =$

v_j \ v_k	ϵ	A	B	C
ϵ	.9	.1	0	0
A	.1	.8	.1	0
B	0	.1	.8	.1
C	0	.1	.1	.8

The string $\epsilon, \epsilon, B, C, A, \epsilon, \epsilon$ is received, and it is known that the transmitted string begins and ends with ϵ.

(a) Find the probability that the i-th transmitted symbol is C, for $i = 1, 2, \ldots, 7$.

(b) Find the probability that the string transmitted is the one received.

(c) Find the probability that no message (a string of ϵ's) was transmitted.

4. (Darwiche 5.10) Exercise 5.10 (After Jensen). Consider a cow that may be infected with a disease that can possibly be detected by performing a milk test. The test is performed on five consecutive days, leading to five outcomes. We want to find out the state of the cows infection over these days given the test outcomes. The prior probability of an infection on day one is 1/10000; the test false positive rate is 5/1000; and its false negative rate is 1/1000. Moreover, the state of infection at a given day depends only on its state at the previous day. In particular, the probability of a new infection on a given day is 2/10000, while the probability that an infection would persist to the next day is 7/10.

   (a) Describe a Bayesian network and a corresponding query that solves this problem. What is the most likely state of the cow's infection over five days given the following test outcomes.

      i. positive, positive, negative, positive, positive

      ii. positive, negative, negative, positive, positive

      iii. positive, positive, negative, negative, positive

   (b) Assume now that the false negative and false positive rates double on a given day in case the test has failed on any previous day. Describe a Bayesian network and a corresponding query that solves the problem under this assumption.

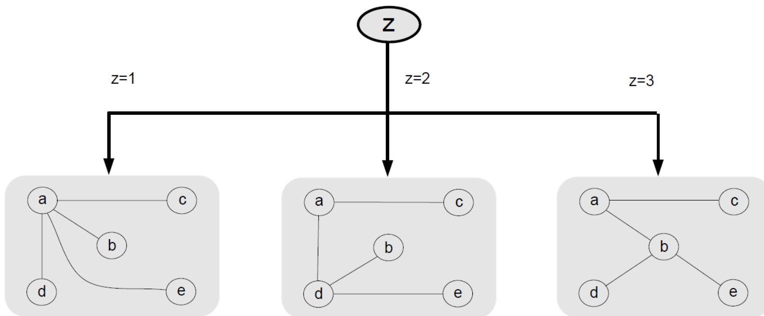

Figure 1: A mixture of trees over a domain consisting of random variables $V = \{a, b, c, d, e\}$, where z is a hidden *choice variable*. Conditional on the value of z, the dependency structure is a tree. A detailed presentation of the mixture-of-trees model is provided in Section 3.

Figure 1: A Mixture network

5. Consider the mixture model in Figure 1 (see attached paper in the class schedule).

   (a) Express the mixture network as a regular Bayesian network in the most sparse manner that you can?

   (b) (extra credit)Provide a general scheme for converting a mixture of Bayesian networks into a regular Bayesian network.

   (c) (extra credit) Discuss the pros and cons of this process