Probabilistic Models – Homework Exercise 3

Publication date: *Thursday, April 28*
Due date: *Sunday, May 15*

The purpose of this exercise is to implement and experiment with message passing algorithms in trees. We will consider a **digital transmission network** defined using an undirected tree $T = (V, E)$, where vertices correspond to transmission nodes / routers and edges represent transmission channels. Each edge $e \in E$ in the network is associated with a noise probability p_e, which defines the probability that a bit transmitted along that edge will be flipped (0→1 or 1→0). Thus the network is fully defined by its structure $T = (V, E)$ and the edge noise probabilities $\{p_e\}_{e\in E}$. Transmission then proceeds as follows:

1. A certain source node $v_0, \in V$ generates a bit (0 or 1) and transmits it to all its neighbors.
2. Once a node receives a transmission, it transmits that bit to all neighbors other than the one it received the transmission from.
3. Transmission along an edge $(u, v) \in E$ in the network is noisy and a bit $b \in \{0,1\}$ in v is received as $1 - b$ in u with probability $p_{(u,v)}$.
4. Once the entire process ends, every node in the network holds a bit, which can be different from the one transmitted at the source node v_0.

Problem 1: General properties of the transmission model

a) Define a Bayesian network that represents a given transmission network. Your model should have a RV for each node in the network, which represents the bit held at that node. Define all the necessary conditional distributions whose product gives the joint distribution over all bits in the network. Assume that transmission originates from a certain source node v_0, and that the transmitted bit is 0 or 1 with equal probability (½).

b) Prove that the marginal distribution over the identity of every bit in the network is uniform, i.e. $P(X_v = 0) = P(X_v = 1) = 1/2$ for every $v \in V$.

   Hint: use induction on the distance of v from the source node v_0.

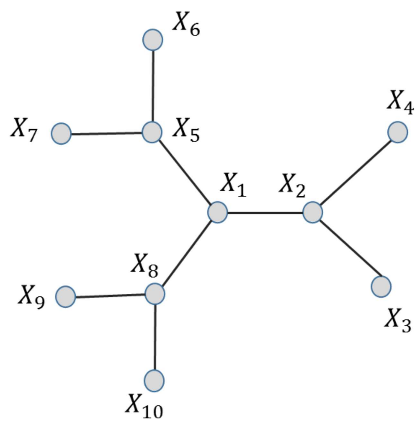
c) Prove that the joint probability over bit identities that you determined in (a) does not change when you change the identity of the source node v_0.

   Hint: again, use induction on the distance between the two source nodes.

The conclusion from the claims you proved in Problem 1 is that the transmission process is "reversible" in the sense that it does not provide any information on the direction or source of transmission. The practical implication of this is that you may use the same Bayesian network regardless of the actual source of transmission.

Problem 2: Inference in a specific transmission network

Consider the transmission network defined by the following tree:



e	p_e
(1,2)	0.1
(2,3)	0.1
(2,4)	0.2
(1,5)	0.1
(5,6)	0.1

e	p_e
(5,7)	0.4
(1,8)	0.1
(8,9)	0.5
(8,10)	0.3

We assume that after the transmission process has completed, we can observe the value of the bits only at the leaves of the tree (X_A for $A = \{3,4,6,7,9,10\}$). The values of bits at the internal nodes are hidden (X_1, X_2, X_5, X_8)

a) Write a computer program in your <u>language of choice</u> that implements a message passing procedure for this network for computing the marginal conditional distribution over all bits given an assignment of bits at the leaves ($P(X_i|X_A)$). Your program should contain two procedures – $collect$ and $distribute$, as outlined in class (see Lecture #6). Invoking $collect$ from a given source node and then $distribute$ from the same source node should result in calculation of the marginal conditional distribution tables $P(X_i|X_A)$ for all internal nodes i in the network (each table has two values). The calculations performed by the message passing procedure should also include the data likelihood - $P(X_A)$. Note that your implementation should allow you to run the message passing algorithm from <u>any source node</u>. Running it from different source nodes should result in the same conditional distributions (see conclusions from Problem 1).

Write your code clearly and document it well. Submit it as an appendix to the exercise.

b) Use your program from (a) to compute the marginal conditionals and data likelihood under the following three data settings:

	X_3	X_4	X_6	X_7	X_9	X_{10}
Setting 1	0	1	1	0	0	1
Setting 2	0	0	1	0	0	1
Setting 3	1	1	1	1	1	1

In each setting run your message passing procedure from three different source nodes X_1, X_2, X_6 to validate that you obtained the same result regardless of the source node. Submit the results of the 3x3 runs. For each run you should specify the likelihood $P(X_A)$ and the conditionals $P(X_i = b|X_A)$ for $b \in \{0,1\}$ and $i \in \{1,2,5,8\}$.

c) Write a computer program in your language of choice that implements a message passing procedure for this network for computing the most probable assignment for all bits, given an assignment of bits at the leaves ($\max\{P(X_1, X_2, X_5, X_8, X_A)\}$). Your program should contain two procedures – $collect$ and $distribute$, as outlined in class (see Lecture #6). Invoking $collect$ from a given source node and then $distribute$ from the same source node should result in calculation of the optimal assignment for all bits in the internal nodes in the network. The calculations performed by the message passing procedure should also include the probability of this optimal assignment. Note that your implementation should allow you to run the message passing algorithm from any source node. Running it from different source nodes should result in the same conditional distributions (see conclusions from Problem 1).

Write your code clearly and document it well. Submit it as an appendix to the exercise.

d) Use your program from (c) to compute the optimal bit assignments to internal nodes under the three data settings specified in (b). In each setting run your message passing procedure from three different source nodes X_1, X_2, X_6 to validate that you obtained the same result regardless of the source node. Submit the results of the 3x3 runs. For each run you should specify the bit assignment with highest probability to the internal nodes X_1, X_2, X_5, X_8, as well as the joint probability of this assignment ($P(X_1, X_2, X_5, X_8, X_A)$).

e) For each of the three data settings, compare the best bit assignment (d) to the bit assignment obtained by using for each internal node the bit that maximizes its posterior probability $P(X_i = b|X_A)$ (b). Are they identical in all cases? Describe briefly what is the difference between the two types of inferences.

Submission Instructions:

- Submit your work on the course **Moodle website** by Sunday, May 15 @23:59.
- Type your solutions or write legibly and scan. If you scan, make sure the scan came out fine. Submit your code as appendix to the submitted work.
- **Submit your work individually!** I do encourage you to discuss ideas with others up to a certain point. Ideas ≠ solutions. If you collaborate with someone on ideas (more than a couple of words before/after class), specify your collaborator's name clearly on the first page and make sure that they acknowledge your collaboration as well.
- **You are allowed to code in pairs.** This means that you may pair up to divide the burden of writing and testing your code. Specify who your partner is (by name or id) in your submission. The code may be submitted by only one of you, but the rest of the assignment (written answers and execution results) should be written and submitted individually.
- You have two weeks to complete the assignment. **Plan your time wisely**. If you would like to ask for an extension due to special circumstances, send me your request by e-mail at least 48 hours before the deadline. I will not grant last minute extensions!
- Please post any questions that you have on the course Piazza website: https://piazza.com/idc.ac.il/spring2016/cs3575/.