MTAT.03.227 Machine Learning
Practice session 11

December 1-3, 2019

1. **Bayesian network example**

   Alarm story: I have a burglar alarm installed at home. It is fairly reliable at detecting a burglary, but also responds on occasion to minor earth quakes. I also have two neighbors, John and Mary. They have promised to call me at work when they hear the alarm. John always calls when he hears the alarm, but sometimes confuses the telephone ringing with the alarm and calls then, too. Mary likes rather loud music and sometimes misses the alarm altogether.

   Five random variables: **A**: Alarm; **B**: Burglary; **E**: Earthquake; **J**: JohnCalls; **M**: MaryCalls (These are binary variables)

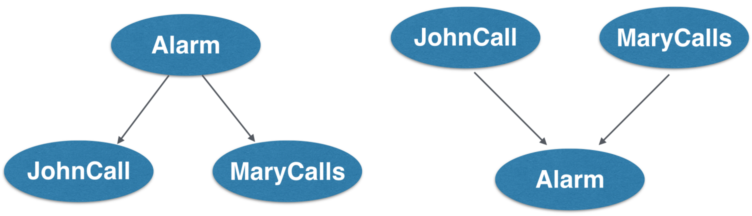
   (a) Draw the BN structure for the given situation. Answer: See the nodes and arrows lecture slide 23.

   (b) We want to define a full probability distribution for this situation.

      i. For every node in the graph, write out the corresponding conditional probability tables (just the structure, without any numbers).
      Answer: See the nodes and arrows lecture slide 32.

      ii. How many probabilities need to be specified for these conditional probability tables? How many probabilities would need to be given if the same joint probability distribution were specified in a joint probability table?
      Answer: There are 5 binary variables, therefore $2^5 = 32$ possible valuations to these variables. To uniquely determine the joint probability table we need one less, that is 31 probabilities, as the last can be calculated as 1 minus the rest.

   (c) Which one of the two Bayesian networks given below makes independence assumptions that are not true?



   Answer: The left one is correct, as also seen in the above tasks. The right one assumes that JohnCall and MaryCalls are unconditionally independent, which is not correct, because knowing that JohnCall is true reveals some information about MaryCalls. The left one assumes that John-Call and MaryCalls are conditionally independent given Alarm, which is correct.

   (d) What is the probability that the alarm has sounded and there was a burglary but no earthquake, and both Mary and John call? (The values for CPTs are shown in the lecture slides)
   Answer: Let us look up the probabilities from lecture slide 47. $P(J, M, A, B, \neg E) = P(J|A)P(M|A)P(A|B, \neg E)P(B)P(\neg E) = 0.9 * 0.7 * 0.94 * 0.001 * (1 - 0.002) = 0.00059$

(e) Given that the alarm sounded and there was no earthquake, and both Mary and John call, would you predict that there is a burglar in your home? *From the lecture we know the joint probability* $P(J, M, A, \neg B, \neg E) = 0.00062$

Answer:

$$P(B|J, M, A, \neg E) = \frac{P(B, J, M, A, \neg E)}{P(J, M, A, \neg E)} = \frac{P(B, J, M, A, \neg E)}{P(B, J, M, A, \neg E) + P(\neg B, J, M, A, \neg E)} =$$

$$\frac{0.00059}{0.00059 + 0.00062} = 0.4876$$

This probability is below 0.5 and therefore it is more probable that there is no burglary.

2. **Hidden Markov Models (HMM)**

We have one Boolean state variable that can have two values: $AtmosphericPressure = \{Low, High\}$. We cannot observe it directly, but we can observe whether it's raining or not – so evidence is also a Boolean variable: $Weather = \{Rain, Dry\}$.

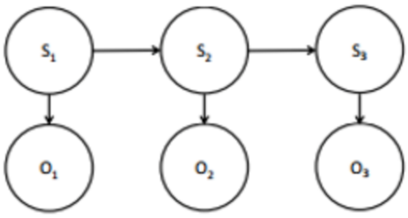
The following tables specify the **initial probabilities** of states, **transition probabilities** and **emission/observation probabilities**, correspondingly:

State	$P(S_1)$
Low	0.4
High	0.6

S_1	S_2	$P(S_2\|S_1)$
Low	Low	0.3
Low	High	0.7
High	Low	0.2
High	High	0.8

S	O	$P(O\|S)$
Low	Rain	0.6
Low	Dry	0.4
High	Rain	0.4
High	Dry	0.6

The HMM (for 3 days) looks like the following graph.



(a) Given an observation sequence (Dry, $Rain$, $Rain$), compute the most likely hidden state sequence.

Hint: Lecture slide 84

Answer: Below is the table of all joint probabilities of P(s1, s2, s3, Dry, Rain, Rain) for all possible hidden state sequences:

s_1, s_2, s_3	$P(s_1, s_2, s_3, Dry, Rain, Rain)$
HHH	0.036864
HHL	0.020736
HLH	0.012096
HLL	0.014336
LHH	0.027216
LHL	0.005376
LLH	0.008064
LLL	0.005184

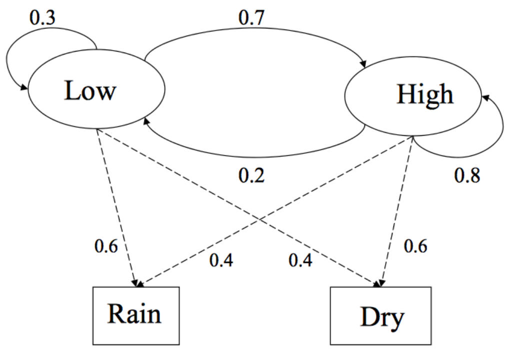
Since $P(s1, s2, s3|Dry, Rain, Rain) = \dfrac{P(s1, s2, s3, Dry, Rain, Rain)}{P(Dry, Rain, Rain)}$ and the denominator is independent of s1, s2, s3, then the most likely sequence is the one with the highest joint probability, that is High,High,High (with joint probability 0.036864). In order to calculate the conditional probability one has to divide this joint probability with the sum of the second column in the above table.

(b) Given the history of evidence, (Dry, $Rain$, $Rain$), what is the probability that the atmospheric pressure was Low on the second day?

Answer: For this one has to add up the probabilities in the above table corresponding to rows with second day having letter L (that is rows 3,4,7 and 8) and divide by the sum of all rows. The result is:

$$\frac{0.012096 + 0.014336 + 0.008064 + 0.005184}{0.036864 + 0.020736 + 0.012096 + 0.014336 + 0.027216 + 0.005376 + 0.008064 + 0.005184} = \frac{0.03968}{0.129872} = 0.30553$$

Note: The same HMM can also be represented as in the figure below.


0.3
0.7
Low
High
0.2
0.8
0.6
0.4
0.4
0.6
Rain
Dry