CSE 486/586, Assignment 4

Josh Crail

Due: Tuesday, November 9, 2021, by 11:59 pm.

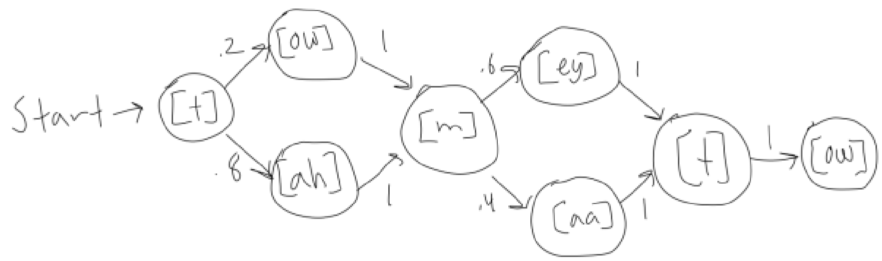
Note 1: The total mark for this assignment is 30. You should *NOT* directly copy anything from slides or other resources. You may get the ideas from slides but what you submit *must be in your own words*. Any help must be acknowledged.

Note 2: Question 5 is only for CSE 586 students (0 point for correct answer, and -5 for wrong answer)

1. Explain in detail how can we develop a Speech Recognition technique. Write your solution mathematically. (10 points)

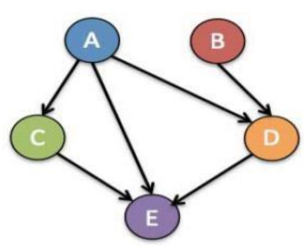
   Speech recognition starts by getting a signal from that seen how likely that given signal is a word which is $P(Words|Signal)$ which can be put into:
   $P(Words|Signal) = \frac{P(Signal|words)P(Words)}{P(Signal)}$

   Where $P(Words)$ is what we call the language mode where the probability is the probability of words being heard. Also $P(Signal|Words)$ is the probability of a certain signal given a word sequence this is called the acoustic model. This is done so that different pronunciations of a word can be accounted for. This is then further broken down where signal is observation of sequence $(O = o_1, o_2, o_3, \dots, o_t)$ and words is the sequence of words $(W = w_1, w_2, w_3, \dots, w_n)$. From this try to find the best match metric $\hat{W}$ where $\hat{W} = \frac{\arg max}{W \in L} \frac{P(O|W)P(W)}{P(O)}$ which can be simplified to $\hat{W} \propto \frac{\arg max}{W \in L} P(O|W)P(W)$ where $P(O|W)$ is the acoustic model and $P(W)$ is the language model. From this the language model can be found by $P(w_1, w_2, w_3, \dots, w_n) = P(w_1) * P(w_2|w_1) * P(w_3|w_1 w_2) \dots P(w_n|w_1 \dots w_{n-1})$. This can result in very complex situations and it can be simplified by the First-Order Markov Assumption. This is done because it can be determined that the probability of a single word mostly is dependent on just the previous word. As such the simplification for the language model using the First-Order Markov Assumption will be $P(w_1, w_2, w_3, \dots, w_n) = P(w_1) * P(w_2|w_1) * P(w_3|w_2) \dots P(w_n|w_{n-1})$. The acoustic model ($P(Signal|Words)$) can be found by splitting it into two parts where one is $P(Phones|Word)$ which is the probability of a sequence of phones given a word and the other is $P(Signal|Phones)$ which is the probability of a sequence of vector quantization values from the acoustic signal given phones. The first one $P(Phones|Word)$ is found by specifying this as a Markov model which shows the possible phone sequences for pronouncing a given word. In the below Markov model, it shows the possible ways tomato can be pronounced.

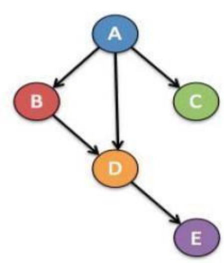


Here the nodes represent the corresponding phone produced at certain points within the pronunciation of a word. The arcs are the probability of that phone being the next one to be pronounced in the word. This can be used to calculate $P(Phones|Word)$, for example $P(\text{[tahmaatow]} \mid \text{tomato}) == 0.8 * 1 * 0.4 * 1 * 1 = 0.32$. The second $P(Signal|Phones)$ is part of the hidden Markov model which is all that needs to be known about it. From this $P(O|W)P(W)$ can be multiplied together which can give us the $\hat{W}$ for the $\frac{\arg max}{W \in L} P(O|W)P(W)$.

2. You are given two different Bayesian network structures 1 and 2, each consisting of 5 binary random variables A, B, C, D, E. Each variable corresponds to a gene, whose expression can be either "ON" or "OFF".



Network 1



Network 2

i) Let A, B, C, D, E follow the independence assumptions of Network #1 above. Simplify P(A, B, C, D, E). (3 points)

P(A) P(B) P(C|A) P(D|A) P(E|A, C, D)

ii) Let A, B, C, D, E follow the independence assumptions of Network #2 above. Simplify P(A, B, C, D, E). (3 points)

P(A) P(B|A) P(C|A) P(D|B,A) P(E|D)

3. Consider the Bayesian network shown below. Calculate *P*(*b, i,¬m, g, j*). (6 points)

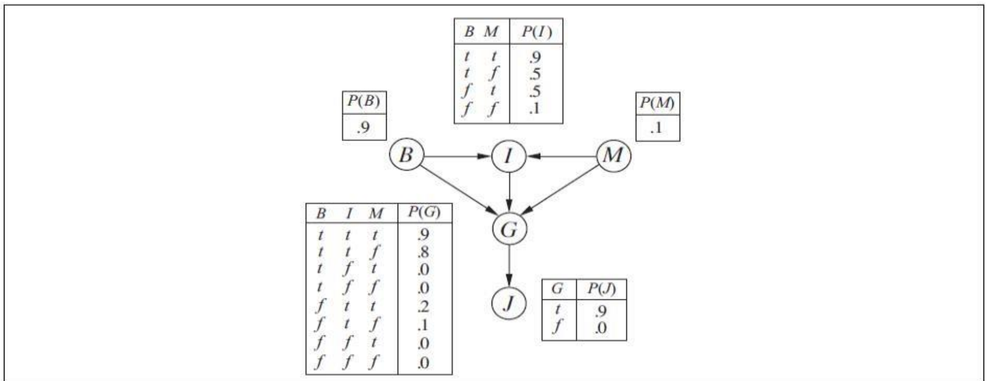


Figure 14.23 A simple Bayes net with Boolean variables $B = BrokeElectionLaw$, $I = Indicted$, $M = PoliticallyMotivatedProsecutor$, $G = FoundGuilty$, $J = Jailed$.

$P(b, i, \neg m, g, j)$

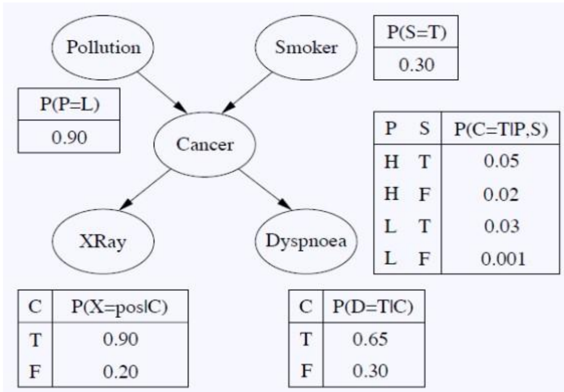
$P(b) = .9$
$P(i|\, b, \neg m) = .5$
$P(\neg m) = .9$
$P(g|b, i, \neg m) = .8$
$P(j|g) = .9$

$P(b, i, \neg m, g, j) = (.9)(.5)(.9)(.8)(.9) = 0.2916$

4. The following problem is known as *medical diagnostic*. This is an application of AI in the medical area. It shows how we can solve *Big Data* problems using small data.

   A patient has been suffering from shortness of breath (called dyspnoea) and visits the doctor, worried that he has lung cancer. The doctor knows that other diseases, such as tuberculosis and bronchitis, are possible causes, as well as lung cancer. She also knows that other relevant information includes whether or not the patient is a smoker (increasing the chances of cancer and bronchitis) and what sort of air pollution he has been exposed to. A positive X-ray would indicate either TB or lung cancer.

   Assume the patient has dyspnoea, is not a smoker, and has not been exposed to high pollution. If the x-ray comes positive, what is the probability that the patient has cancer? The Bayesian network and corresponding conditional probability tables for this situation are shown below. (8 points)



$$P(C=T \mid D=T, S=F, P=L, X=pos) = \frac{P(C=T, D=T, S=F, P=L, X=pos)}{P(C=T, D=T, S=F, P=L, X=pos) + P(C=F, D=T, S=F, P=L, X=pos)}$$

$P(C \mid P=L, S=F) = 0.001$
$P(D \mid C=T) = 0.65$
$P(S) = .70$
$P(P=L) = .90$

$$P(C=T \mid D=T, S=F, P=L, X=pos) = \frac{(.001)(.65)(.70)(.90)(.90)}{(.001)(.65)(.70)(.90)(.90) + (.999)(.30)(.70)(.90)(.20)} = \frac{0.00036855}{0.00036855 + 0.0377622}$$

$P(C=T \mid D=T, S=F, P=L, X=pos)$ = 0.009665428

5. Let A, B be independent given C. Also, let A, C be independent. Which of the following is true? Why? (Only CSE 586 students)

    i) A,B are independent

    ii) B,C are independent

    iii) Both

    iv) None This is true because for A it is P(A|C) and B P(B|C) meaning the answer that results from them is going to be dependent on the value of C so i is false. ii is false since C is independent but as already mentioned B is dependent on C as such neither of the these are true.