Homework 7

Morgan Ciliv

April 5, 2017

1 Exact Inference

Note: The important parts are typed. I've also included some of the math work done on paper for your reference.

For all my variable orderings: I chose the variable ordering because by using the adapted min-fill heuristic mentioned in class. After that, I determined it to be the variable ordering as it appeared to minimize the number of sums needed to evaluate the variable after several checks such that the summations were moved as far right as possible.

1.1 Part 1

Using the rule of conditional probability:
$P(S|C+, D+) = \frac{P(S,C+,D+)}{P(D+,C+)}$

Calculate the numerator first:
The numerator is equivalent to
$P(S+) \sum_B P(B|S+)P(D+|C+, B) \sum_L P(L|S+) \sum_T P(C+|T, L \sum_V P(T|V)P(V)$

Calculated the following factors:
$f_1(T) = [.0104, 0.9896]$
$f_2(L) = [1, .0104]$

Then, the numerator is .044836.

Calculate the denominator now:

The denominator is equivalent to (Note: f_2 taken from earlier)
$\sum_S P(S) \sum_B P(B|S)P(D+|C+, B) \sum_L L|Sf_2L$

Calculated the following factors:
We can use both of the previous factors, $f_1 and f_2$.

$f_3(S) = [.10936, .020296]$
Then, the denominator is .05255

Therefore, we divide the numerator by the denominator as such.
=.044836/.05255008 =0.8532356183
The answer to the first question is 0.8532 which is the same as the value obtained by using my Weka Bayes Net.

Below is my work done on paper for this question thus far:

Q. 1 P. 1

$$P(S+, C+, D+) = \sum P(S+)\,P(L|S+)\,P(C+|T,L)\,P(B|S+)\,P(D+|C+,B)\,P(T|V)\sum_{V} P(V)$$

$$\sum_{B} P(S+)\,P(B|S+)\,P(D+|C+,B)\sum_{L} P(L|S+)\sum_{T} P(C+|T,L)\underbrace{\sum_{V} P(T|V)\,P(V)}_{f_1(T)}$$

now factor

$$f_1(T) = \begin{cases} P(T+|V+)P(V+) + P(T+|V-)P(V-) \\ P(T-|V+)P(V+) + P(T-|V-)P(V-) \end{cases}$$

$$(\;) = \begin{cases} (.05)(.01) + (.01)(.99) \\ .95(.01) + .99(.99) \end{cases}$$

$$f_1(T) = \begin{cases} .0104 & T+ \\ .9896 & T- \end{cases}$$

$$f_2(L) = \begin{cases} P(C+|T+,L+)\,f_1(T+) + P(C+|T-,L+)\,f_2(T-) \\ P(C+|T+,L-)\,f_1(T+) + P(C+|T-,L-)\,f_2(T-) \end{cases}$$

$$= \begin{cases} 1(.0104) + 1(.9896) \\ 1(.0104) + 0 \end{cases}$$

$$f_2(L) = \begin{cases} 1 & L+ \\ 0.0104 & L- \end{cases}$$

$$P(S+)\sum_B P(B|S+)P(D+|C+,B)\sum_L P(L|S+)f_2(L)$$

$$= (0.5)\left[P(B+|S+)P(D+|C+,B+)\left[P(L+|S+)f_2(L+)+P(L-|S+)f_2(L-)\right]\ldots\right.$$

$$\sum_L P(L|S+)f_2(L)$$
$$= [P(L+|S+)f_2(L+)+P(L-|S+)f_2(L-)]$$
$$= 0.1(1)+0.9(0.0104)$$
$$= 0.1+0.00936$$
$$= 0.10936$$

$$= 0.5\sum_B P(B|S+)P(D+|C+,B)\,0.10936$$

$$= 0.5[P(B+|S+)P(D+|C+,B+)+$$
$$P(B-|S+)P(D+|C+,B-)]\,0.10936$$

$$= 0.5[0.6(0.9)+0.4(0.7)]\,0.10936$$
$$= 0.044836$$

Should be .8532 (Weka) I think

Numerator

Nevermind, this is the final answer.

$P(C+, D+)$ ← Denominator

$$= \sum_{SBLTV} \overset{P(S)}{\hat{P}}(B|S) P(D+|C+,B) \overset{P(L|S+)}{\hat{P}}(C+|T,L) \ddot{P}(T|V) P(V)$$

$$= \sum_S P(S) \sum_B P(B|S) P(D+|C+,B) \sum_L P(L|S)$$

$$\underbrace{\sum_T P(C+|T,L) \sum_V P(T|V) P(V)}_{\text{This section can be copied from previous}} \Big\} f_2(L)$$

∴ This with $f_2(L)$ at the end ∴ evaluate...

$$\sum_L P(L|S) f_2(L) = f_3(S)$$

$$= \begin{cases} P(L+|S+) f_2(L+) + P(L-|S+) f_2(L-) \\ P(L+|S-) f_2(L+) + P(L-|S-) f_2(L-) \end{cases}$$

$$= \begin{cases} 0.10(1) + 0.9(0.0104) \\ 0.01(1) + 0.99(0.104) \end{cases}$$

$$f_3(S) = \begin{cases} 0.10936 \\ 0.020296 \end{cases}, \quad f_4(S) \ldots$$

$$= \begin{cases} P(B+|S+) P(D+|C+,B+) f_3(S+) + (\text{'' w/ } B-\text{''}) \\ P(B+|S-) P(D+|C+,B+) f_3(S-) + (\text{'' w/ } B-\text{''}) \end{cases}$$

$$= \begin{cases} .6(.9)(.10936) + 0.4(0.7)(.10936) \\ .3(.9)(.020296) + .7(.7)(.020296) \end{cases}$$

$$= \begin{cases} 0.0448376 \\ 0.0071248 \end{cases}$$

$$\rightarrow 0.0448376 + 0.0071248$$

$$= 0.0525508$$ ∴ divide in typed version

1.2 Part 2

Using the rule of conditional probability:
$P(D|S+,C+) = \frac{P(S,C+,D+)}{P(S+,C+)}$

The numerator is the same as above which equals .044836

$P(S+)\sum_L P(L|S+)\sum_T P(C+|T,L)\sum_V P(T|V)P(V)$

Everything after $P(S+)$ in this equation is the same as part of the numerator of part 1, more specifically everything from the summation over L onwards, which is equivalent to $\sum_{P(L|S+)f_2(L)}$. Therefore, the equation for the denominator for this second part of this problem is:

$P(S+)\sum_{P(L|S+)f_2(L)}$ =.5(.10936)
=.05468

Now, we divide plug the values into the conditional probability equation above.
=.044836/.05468
=.8200, which is equal to what was obtained in my Bayes net in Weka.

2 Looking Back

1. Describe how you would formulate MAP Inference as a local search problem.

First, I think that local search would be a good match for MAP as MAPs running time can be exponential. I would formulate inference as a local search problem by setting a limit on the depth, such that local search constrains the problem such that local search allows Map to look at certain degree of nodes away from a particular node. More specifically, local search would use the probabilities of each branch to calculate and use the nodes down the path with the highest probability for MAP. We would be increases the likelihood that the branch found by local search is the answer. Nodes obviously represent variables.

2. Describe how hill climbing and simulated annealing would (roughly) go about solving MAP Inference

Hill climbing is a type of local search, so it would work similarly as the above. Given an initial start node, hill climbing would incrementally choose a new node that works better as a variable assignment.

Simulated annealing would identify neighboring nodes and based on the node, use probability to determine which node should be reached next. It does this each iteration. It would keep repeating this process until a set of states is met that meets the problem or till all computations have been calculated.

3. State which of the two you believe would do a better job, and why.

Hill climbing could get stuck in a local maxima quite easily. I also think that simulated annealing would be a good fit for this particular problem since we are finding a set of variable assignments. I also think simulated annealing would be a good fit because this model is probabilistic. Based of of these ideas, I believe simulated annealing would do a better job.

4. Describe an experiment you might run to see which does better, and justify your experimental design.

I would have 1000 varying and different-sized sets of different Bayes nets in which to run the experiments on. For each of the graphs, I would run both hill climbing and simulated annealing for 5 evenly spaced different amounts of time which would be $1/3, 2/3, 3/3, 4/3, 5/3$ of an approximated average running time of the experiments. Once this is all complete I would take the use statistics to determine if there is a significant different in the choosing of the variable assignments. I could simply identify which had the best variable assignments on average.

I chose the experiment above to test whether hill climbing or simulated annealing would be better because I would have a good sample size so that I have a better estimate of the entire population of possible graphs for the problem. I also, test each algorithm fairly so that there isn't any intentional bias in the problem. I also test the algorithms with different amounts of time so that a particular algorithm isn't chosen to work better for an exact time. Again, we would be getting a better estimate by doing this process with a range of values in terms of both space and time for the problem.