CS540 – Section 2: Introduction to Artificial Intelligence

Homework #5

Solution

Problem 1: Conditional Independence and Joint Probability [15]

(a) Answer: 80

If we want to uniquely specify the joint probability P(A, B, C, D), in general we need all combinations of possible variable values. Then, we need 3^4 = 81 different values to specify every possible joint probability. Now notice that sum of all P(a, b, c, d) is 1 (a, b, c, d are values of A, B, C, D), which means if we know 80 of the joint probabilities, we can determine the remained one. Thus, we need only 80 different variables.

Generally, for n variables, we need at least **3^n – 1** parameters.

(b) Answer: 20

From the given Bayes Net, we need to specify 4 CPTs of A, B given A, C given B and D given C. Then, we can uniquely specify the joint probability, which is

$$P(A, B, C, D) = P(D|C)P(C|B)P(B|A)P(A)$$

- For the CPT of A, we only need to know probabilities of 2 values of A.
- For the CPT of B given A, each value of A, we only need to know probabilities of 2 values of B, which means 6 values
- Similarly, we need 6 values for each CPT of C and D
- Thus, we need 20 values to specify the joint probability. It's clear that it's much less than 80 values in question a.

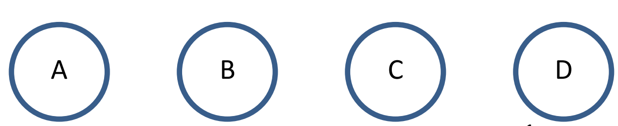
(c) Based on your assumption about what is known, we have different acceptable answers:

- Assuming nothing else is given (implicity): A and D are not independent. It's easy to design CPTs to see for some A = a and D = d, $P(a, d) \neq P(a)P(d)$.
- Assuming B or C is given: A and D are conditionally independent given B or C. You can give a proof for this equality P(A, D|C) = P(A|C)P(D|C) using Bayes rules.

(d) Answer: 4 (Note: this question is for binary variables, for 4 3-value variables, we need 8 parameters)

The optimal network happens when 4 variables are independent with each other. Thus, for each variable, we just need the probabilities of its 1 value to determine the CPT, which could uniquely specify the joint probability. Thus, the minimum number is 4.

(e) The plot is shown below:



Problem 2: Using a Bayesian Network [20]

(a) [1] What is the probability that there was a bird on the lawn?

P(B) = P(B|S)* P(S) + P(B|¬S)*P(¬S) = **0.304**

(b) [3] If there was a bird on the lawn, what is the probability that the sprinkler was on?

P(S|B) = P(B,S)/P(B) = P(B|S)*P(S)/P(B) = 0.01*0.4/0.304 = **0.0132**

(c) [3] If there was a bird on the lawn, what is the probability that the lawn was wet?

P(W|B)= P(W,B)/P(B) = 0.061/0.304 = **0.201**

P(W,B) = P(R,S,W,B) + P(R,¬S,W,B) + P(¬R,S,W,B) + P(¬R,¬S,W,B)
= P(R)*P(S)*P(W|R,S)*P(B|S) + P(R)*P(¬S)*P(W|R,¬S)*P(B|¬S)
+ P(¬R)*P(S)*P(W|¬R,S)*P(B|S) + P(¬R)*P(¬S)*P(W|¬R,¬S)*P(B|¬S)
= 0.1*0.4*1*0.01 + 0.1*0.6*1*0.5 + 0.9*0.4*1*0.01 + 0.9*0.6*0.1*0.5
= 0.061

(d) [3] If there was a bird on the lawn, what is the probability that the lawn was dry, and it did not rain, and the sprinkler was not on?

P(¬W,¬R,¬S|B) = P(¬W,¬R,¬S,B)/P(B) = 0.243/0.304 = **0.7993**

P(¬W,¬R,¬S,B) = P(¬R)*P(¬S)*P(¬W|¬R,¬S)*P(B|¬S) = 0.9*0.6*0.9*0.5 = 0.243

(e) [3] What is the probability that the lawn was wet?

P(W) = P(W|R,S)*P(R)*P(S) + P(W|R,¬S)*P(R)*P(¬S)
+ P(W|¬R,S)*P(¬R)*P(S) + P(W|¬R,¬S)*P(¬R)*P(¬S)
= 1*0.1*0.4 + 1*0.1*0.6 + 1*0.9*0.4 + 0.1*0.9*0.6
= **0.514**

(f) [3] If the lawn was wet and the sprinkler was on, what is the probability that it rained?

P(R|W,S) = P(R,W,S)/P(W,S) = 0.04/0.4 = **0.1**

P(W,S) = P(W,S,R) + P(W,S,¬R)
= P(W|R,S)*P(S)*P(R) + P(W|S,¬R)*P(S)*P(¬R)
= 1*0.1*04 + 1*0.9*0.4 = 0.4

(g) [4] No, it's not. A Naïve Bayes of W should have all directed connections from W to other variables. The directions in this network are inverse.