James Hahn

CS1675 Machine Learning

Dr. Adriana Kovashka

**Discussion of results**

As requested, the probability of occurrence for five of the sentences have been calculated below:

|  |  |
| --- | --- |
| **Sentence** | **Probability of occurrence** |
| “john saw the cat.” | 0.000216177 |
| “john ate.” | 0.025678750 |
| “john saw mary.” | 0.000603579 |
| “mary saw john.” | 0.000601782 |
| “cat saw the john.” | 0.000046368 |

**Bayes net written exercises**

1. *Consider the example of the car fuel system shown in Figure 8.21, and suppose that instead of observing the state of the fuel gauge G directly, the gauge is seen by the driver D who reports to us the reading on the gauge. This report is either that the gauge shows full D = 1 or that it shows empty D = 0. Our driver is a bit unreliable, as expressed through the following probabilities*

*p(D = 1|G = 1) = 0.9*

*p(D = 0|G = 0) = 0.9.*

*Suppose that the driver tells us that the fuel gauge shows empty, in other words that we observe D = 0. Evaluate the probability that the tank is empty given only this observation. Similarly, evaluate the corresponding probability given also the observation that the battery is flat, and note that this second probability is lower. Discuss the intuition behind this result, and relate the result to Figure 8.54.*

P(D = 0 | F = 0) = P(D = 0 | G = 0)P(G = 0 | B = 0, F = 0)P(B = 0) +

P(D = 0 | G = 0)P(G = 0 | B = 1, F = 0)P(B = 1) +

P(D = 0 | G = 1)P(G = 1 | B = 0, F = 0)P(B = 0) +

P(D = 0 | G = 1)P(G = 1 | B = 1, F = 0)P(B = 1)

= [ 0.9 \* 0.9 \* 0.1 ]+[ 0.9 \* 0.8 \* 0.9 ]+[ 0.1 \* 0.1 \* 0.1 ]+[ 0.1 \* 0.2 \* 0.9 ]

= 0.081 + 0.648 + 0.001 + 0.018 = **0.748**

P(D = 0) = P(D = 0 | G = 0)P(G = 0 | B = 0, F = 0)P(B = 0)P(F = 0) +

P(D = 0 | G = 0)P(G = 0 | B = 0, F = 1)P(B = 0)P(F = 1) +

P(D = 0 | G = 0)P(G = 0 | B = 1, F = 0)P(B = 1)P(F = 0) +

P(D = 0 | G = 0)P(G = 0 | B = 1, F = 1)P(B = 1)P(F = 1) +

P(D = 0 | G = 1)P(G = 1 | B = 0, F = 0)P(B = 0)P(F = 0) +

P(D = 0 | G = 1)P(G = 1 | B = 0, F = 1)P(B = 0)P(F = 1) +

P(D = 0 | G = 1)P(G = 1 | B = 1, F = 0)P(B = 1)P(F = 0) +

P(D = 0 | G = 1)P(G = 1 | B = 1, F = 1)P(B = 1)P(F = 1)

= [ 0.9 \* 0.9 \* 0.1 \* 0.1 ] + [ 0.9 \* 0.8 \* 0.1 \* 0.9 ] + [ 0.9 \* 0.8 \* 0.9 \* 0.1 ] +

[ 0.9 \* 0.2 \* 0.9 \* 0.9 ] + [ 0.1 \* 0.1 \* 0.1 \* 0.1 ] + [ 0.1 \* 0.2 \* 0.1 \* 0.9 ] +

[ 0.1 \* 0.2 \* 0.9 \* 0.1 ] + [ 0.1 \* 0.8 \* 0.9 \* 0.9 ]

= 0.0081 + 0.0648 + 0.0648 + 0.1458 + 0.0001 + 0.0018 + 0.0018 + 0.0648

= **0.352**

P(F = 0 | D = 0) = P(D = 0 | F = 0)P(F = 0) / P(D = 0) = ( 0.748 \* 0.1 ) / 0.352 = **0.213**

P(D = 0 | F = 0, B = 0) = P(D = 0 | G = 0)P(G = 0 | B = 0, F = 0) +

P(D = 0 | G = 1)P(G = 1 | B = 0, F = 0)

= [ 0.9 \* 0.9 ] + [ 0.1 \* 0.1 ]

= **0.82**

P(D = 0 | F = 1, B = 0) = P(D = 0 | G = 0)P(G = 0 | B = 0, F = 1) +

P(D = 0 | G = 1)P(G = 1 | B = 0, F = 1)

= [ 0.9 \* 0.8 ] + [ 0.1 \* 0.2 ]

= **0.74**

P(F = 0 | D = 0, B = 0) = [ P(D = 0 | F = 0, B = 0)P(F = 0) ] / [ P(D = 0 | F = 0, B = 0)P(F = 0)

+ P(D = 0 | F = 1, B = 0)P(F = 1) ]

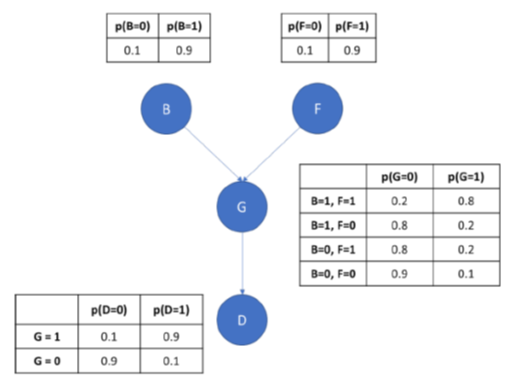
= [ 0.82 \* 0.1 ] / [ 0.82 \* 0.1 + 0.74 \* 0.9 ]

= 0.082 / 0.748

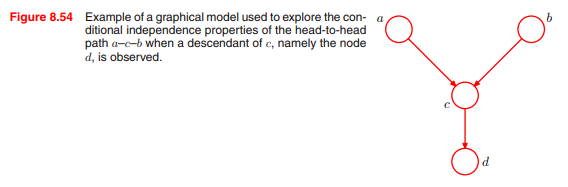
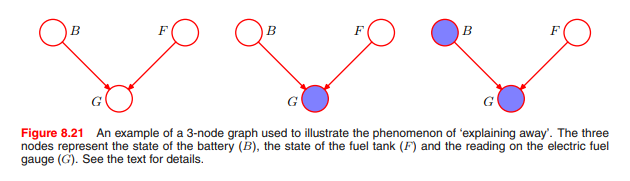
= **0.110**

The **intuition** is that if the battery is flat, the electrical gadgets, namely the electric fuel gauge (G) will not have enough power to supply correct measurements. Therefore, since the battery is flat, the electric fuel gauge will be faulty, so the odds of the tank actually being empty, when the humans thinks/says it’s empty, are lower.

The **relation** between the result and Figure 8.54 is represented in Figure 1 below. Clearly, they are orchestrated in similar fashion. The old graphical representation can be seen in Figure 8.21, but the new variable D is conditional on G, so it is a child node. Therefore, when D is added to Figure 8.21, it is converted into a structure similar to Figure 8.54.



**Figure 1.** This represents the new graphical representation of the problem with the newly added variable, D.



1. *In this exercise, we'll do some cross-domain recommendation, where we assume that there is a correlation between a user's taste in music and film. We'll only consider one music genre, namely jazz (which we'll denote by J), and four films, "Waking Life" (denoted by W), "Borat" (denoted by B), "Cinema Paradiso" (denoted by C) and "Requiem for a Dream" (denoted by R). We'll assume that conditioned on whether the user likes jazz, the movie likes/dislikes are independent. The prior probability of liking jazz is 30%. We've defined the following (combined) conditional probability table, where "=1" means "likes". We are conditioning on the first column.*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| J=1 | W=1 | B=1 | C=1 | R=1 |
| T | 80 | 20 | 70 | 50 |
| F | 30 | 50 | 30 | 40 |

P(J)P(W | J)P(~B | J)P(~C | J)P(R | J) = 0.0288

P(~J)P(W | ~J)P(~B | ~J)P(~C | ~J)P(R | ~J) = 0.0294

P(J)P(W | J)P(B | J)P(C | J)P(R | J) = 0.0168

P(~J)P(W | ~J)P(B | ~J)P(C | ~J)P(R | ~J) = 0.0126

1. P(J | W, ~B, ~C, R) = P(J)P(W | J)P(~B | J)P(~C | J)P(R | J) / [ P(J)P(W | J)P(~B | J)P(~C | J)P(R | J) + P(~J)P(W | ~J)P(~B | ~J)P(~C | ~J)P(R | ~J) ] = 0.0288 / [ 0.0288 + 0.0294 ] = **0.4948**
2. P(J | W, B, C, R) = P(J)P(W | J)P(B | J)P(C | J)P(R | J) = 0.0168 / [ 0.0168 + 0.0126 ] = **0.5714**