AI Project 2

Adrian Guzman afg30

November 2019

1

1.1 a

```
model = BayesianModel([('F', 'G'), ('B', 'G')])
cpd_b = TabularCPD(variable='B', variable_card=2, values=[[0.2, 0.8]])
cpd_f = TabularCPD(variable='F', variable_card=2, values=[[0.1, 0.9]])
cpd_g = TabularCPD(variable='G', variable_card=2,
                   values=[[0.9, 0.9, 0.8, 0.1],
                           [0.1, 0.1, 0.2,  0.9]],
                  evidence=['F', 'B'],
                  evidence_card=[2, 2])
model.add_cpds(cpd_f, cpd_b, cpd_g)
model.check_model()
```

Returns true verifying that the model is defined correctly.

1.2 b

From Bayes Rule

$$P(\neg F|\neg G) = \frac{P(\neg G|\neg F)P(\neg F)}{P(\neg G)} \tag{1}$$

From the Sum Rule

$$P(\neg G|\neg F) = \sum_B [P(\neg G|\neg F, B = b)P(B = b)]$$

$$P(\neg G) = \sum_F \sum_B P(\neg G|F = f, B = b)P(F = f)P(B = b)$$

Plugging back into (1)

$$P(\neg F|\neg G) = \frac{\sum_B [P(\neg G|\neg F, B = b)P(B = b)] * P(\neg F)}{\sum_F \sum_B P(\neg G|F = f, B = b)P(F = f)P(B = b)} \tag{2}$$

In equation (2) we are left with summations of P(G) given F and B, which are defined in the cpd for G, P(F), which is defined in the cpd for F, and P(B=b) which is defined in cpd for B. So we have expressed the $P(\neg F|\neg G)$ in terms of probabilities defined in the model.

1.3 c

```
print(infer.query(['F'], evidence={'G': 0}))
```

F	phi(F)
F(0)	0.2941
F(1)	0.7059

The equation queries the probabilities of F given that G is 0. From the above table, P(F=0) is 0.2941.

1.4 d

1) $P(F)$

```
print(infer.query(['F']))
```

F	phi(F)
F(0)	0.1000
F(1)	0.9000

2) $P(F|G=0)$

```
print(infer.query(['F'], evidence={'G': 0}))
```

F	phi(F)
F(0)	0.2941
F(1)	0.7059

3) $P(F|G=0, B=0)$

```
print(infer.query(['F'], evidence={'G': 0, 'B': 0}))
```

F	phi(F)
F(0)	0.1111
F(1)	0.8889

1.5 e

Only given that G=0, we must account for the probability that G=0 due to F=0 and/or due to B=0. There is a greater probability of G=0 due to F=0 than due to B=0. When we learn that B=0, for F=0 to be true, then G=0 would be due to F=0 *and* B=0, which has a lower probability than due to just F=0 as was the potential case before we learned that B=0.

ie. When we learn that the battery is dead, it increases the probability that the gauge reading empty is due to the dead battery, thus decreasing the probability that the gauge is reading empty due to the fuel tank being empty.

1.6 f

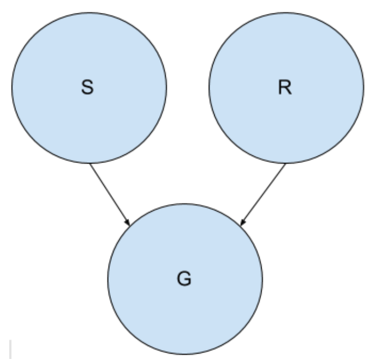
This can be modeled as $P(F|B = 0)$ which results in the following

F	phi(F)
F(0)	0.1000
F(1)	0.9000

Because the two variables are independent, observations about one of the variables does not affect the probability of the other variables outcomes. So our belief about the fuel tank does not change given the battery is dead.

2

2.1 a



2.2 b

The cpd for R

R(0)	0.8
R(1)	0.2

The cpd for S

S(0)	0.9
S(1)	0.1

The cpd for G

R	R(0)	R(0)	R(1)	R(1)
S	S(0)	S(1)	S(0)	S(1)
G(0)	1	0.05	0	0
G(1)	0	0.95	1	1

I assume the sprinkler makes the grass wet, and that when it is left on, it breaks 5% of the time, meaning that it won't get the grass wet 5% of the time it is left on

2.3 c

```
model = BayesianModel([('S', 'G'), ('R', 'G')])
cpd_r = TabularCPD(variable='R', variable_card=2, values=[[0.8, 0.2]])
cpd_s = TabularCPD(variable='S', variable_card=2, values=[[0.9, 0.1]])
cpd_g = TabularCPD(variable='G', variable_card=2,
                    values=[[1, 0.05, 0, 0],
                            [0, 0.95, 1, 1]],
                   evidence=['R', 'S'],
                   evidence_card=[2, 2])

model.add_cpds(cpd_r, cpd_s, cpd_g)
model.check_model()
```

Returns true verifying that the model is valid.

2.4 d

For her first query

```
print(infer.query(['R'], evidence={'G': 1}))
```

R	phi(R)
R(0)	0.2754
R(1)	0.7246

The probability that the grass is wet due to rain last night is 72.46%.

For her second query

```
print(infer.query(['S'], evidence={'G': 1}))
```

S	phi(S)
S(0)	0.6522
S(1)	0.3478

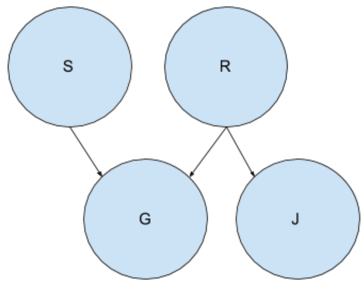
The probability that the grass is wet because she left the sprinkler on is 34.78%.

Combined queries

```
print(infer.query(['S','R'], evidence={'G': 1}))
```

R	S	phi(R,S)
R(0)	S(0)	0.0000
R(0)	S(1)	0.2754
R(1)	S(0)	0.6522
R(1)	S(1)	0.0725

2.5 e



The cpd for J

R	R(0)	R(1)
J(0)	0.85	0.0
J(1)	0.15	1.0

```
model = BayesianModel([('S', 'G'), ('R', 'G'), ('R', 'J')])
cpd_j = TabularCPD(variable='J', variable_card=2,
                   values=[[.85, 0],
                           [.15, 1]],
                  evidence=['R'],
                  evidence_card=[2])
```

These are the lines that I altered or added from the previous model definition. I add a new edge from R to J, and added the probability values for J.

2.6 f

$$P(S|G,J) = \frac{P(G,J,S)}{P(G,J)} \tag{3}$$

$$\begin{aligned} P(G,J,S) &= P(S|G,J)P(G,J) \\ &= P(S|G)P(G,J) \\ &= P(S|G)P(G)P(J) \\ &= \frac{P(S,G)}{P(G)}P(G)P(J) \\ &= P(S,G)P(J) \end{aligned}$$

$$= P(G|S)P(S)P(J)$$

Plug back in to (3)

$$P(S|G,J) = \frac{P(G|S)P(S)P(J)}{P(G,J)} \tag{4}$$

$$= \frac{P(G|S)P(S)P(J)}{P(G)P(J)}$$

$$= \frac{P(G|S)P(S)}{P(G)}$$

$$= \frac{\sum_R [P(G|S,R=r)P(R=r)] * P(S)}{\sum_R \sum_S P(G,R=r,S=s)P(R=r)P(S=s)}$$

Finally

$$P(S|G,J) = \frac{\sum_R [P(G|S,R=r)P(R=r)] * P(S)}{\sum_R \sum_S P(G,R=r,S=s)P(R=r)P(S=s)} \tag{5}$$

S(0)	0.8515
S(1)	0.1485

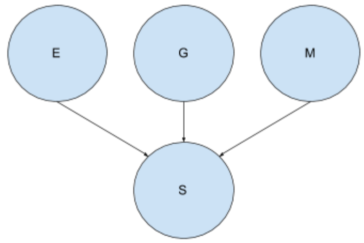
There is a 14.85% chance she left the sprinkler on after observing her and Jack's lawn are wet.

2.7 g

Because the rain affects both of their grasses, if both are wet then it is more likely that it rained last night. If it is more likely that it rained last night, then it is less likely that the grass being wet is due to the sprinkler being left on which would only affect Tracey's lawn.

3 Extra Credit

For my own Bayesian network, I will attempt to model the probability of University of Florida student, Mini Golf, starting her homework on time. This probability will be represented as S. S has three contributing factors: whether Mini enjoys the content of the class (E), The current letter grade Mini has in the class (G), and whether she has mini-golf practice the night that she should start her homework (M).



3.1 Cpd's

Cpd for E

E(0)	0.4
E(1)	0.6

60% chance Mini enjoys her class, 40% chance she does not.

Cpd for G

G(0)	0.3
G(1)	0.6
G(2)	0.1

30% chance Mini has an A in her class, 60% chance she has a B, 10% chance she has a C.

Cpd for M

M(0)	0.7
M(1)	0.3

30% chance Mini has mini-golf practice, 70% chance she does not.

Cpd for S

E	E(0)	E(0)	E(0)	E(0)	E(0)	E(0)	E(1)	E(1)	E(1)	E(1)	E(1)	E(1)
G	G(0)	G(0)	G(1)	G(1)	G(2)	G(2)	G(0)	G(0)	G(1)	G(1)	G(2)	G(2)
M	M(0)	M(1)	M(0)	M(1)	M(0)	M(1)	M(0)	M(1)	M(0)	M(1)	M(0)	M(1)
S(0)	0.85	0.95	0.75	0.85	0.4	0.5	0.65	0.85	0.5	0.7	0.2	0.4
S(1)	0.15	0.05	0.25	0.15	0.6	0.5	0.35	0.15	0.5	0.3	0.8	0.6

3.2 Implementing model with toolbox

```
model = BayesianModel([('E', 'S'), ('G', 'S'), ('M', 'S')])
cpd_e = TabularCPD(variable='E', variable_card=2, values=[[0.4, 0.6]])
cpd_g = TabularCPD(variable='G', variable_card=3, values=[[0.3, 0.6, 0.1]])
cpd_m = TabularCPD(variable='M', variable_card=2, values=[[0.7, 0.3]])
cpd_s = TabularCPD(variable='S', variable_card=2,
```

```
            values=[[0.85, 0.95, 0.75, 0.85, 0.4, 0.5, 0.65, 0.85, 0.5, 0
                   [0.15, 0.05, 0.25, 0.15, 0.6, 0.5, 0.35, 0.15, 0.5, 0.
           evidence=['E', 'G', 'M'],
           evidence_card=[2, 3, 2])

model.add_cpds(cpd_e, cpd_g, cpd_m, cpd_s)
model.check_model()
```

Returns true so it is a valid model.

3.3 Inferences

Mini has a project for her Introduction to AI class due in 3 days, she estimates that to finish it comfortably, she should begin tonight. She enjoys her class very much, her grades have not been published yet, but she has mini-golf practice tonight! What is the probability that she starts her homework tonight?

```
print(infer.query(['S'], evidence={'E': 1, 'M': 1}))
```

S	phi(S)
S(0)	0.7150
S(1)	0.2850

Only 28.5%! It looks like she is pretty unlikely to start her homework on time! If Mini emails her coach and asks for practice to be rescheduled to tomorrow, then what is the probability that she will start her homework tonight?

```
print(infer.query(['S'], evidence={'E': 1, 'M': 0}))
```

S	phi(S)
S(0)	0.5150
S(1)	0.4850

48.5%! Still lower than we would expect. The coach has not responded yet so she is currently unsure of whether she will have mini-golf practice tonight. But her professor has just released grades and it looks like Mini currently has a C! Looks like she better get to work on her project! What is the probability that she begins her homework tonight now that she knows her grade?

```
print(infer.query(['S'], evidence={'E': 1, 'G': 2}))
```

S	phi(S)
S(0)	0.2600
S(1)	0.7400

74%! Looks like Mini is motivated to keep her grades up!

Mini's coach has just responded to her email, and refused her request to reschedule practice. So Mini decides to forgo practice and finish her homework instead. But she does not email her coach, so he is left wondering why she has missed practice. He decides to compute the probabilities for reasons she could have missed practice.

G	E	phi(G)
G(0)	E(0)	0.0417
G(0)	E(1)	0.1513
G(1)	E(0)	0.1530
G(1)	E(1)	0.4591
G(2)	E(0)	0.0661
G(2)	E(1)	0.1287

It looks like it is most likely that Mini started her homework instead of going to practice because she enjoys her class and has a B, and wants to improve her grade. The coach can't fault her for her decision so he decides to only make her run 4 laps of suicides instead of the usual 8 as punishment.