Homework 5

FULYA KOCAMAN
CWID: 803023878

Due: **Check date on Canvas**. Prepare your answers as a **single PDF file**.
Group work: You may work in groups of 1-3. Include all group member names in the PDF file. Only one person in the group should submit to Canvas.

1. Consider the following dataset with weather conditions on ten winter days (Cloudy, Freezing) and if an accident occurred on that day.

Day	**Cloudy**	**Freezing**	**Accident?**
1	*False*	*False*	*No*
2	*True*	*False*	*No*
3	*False*	*True*	*No*
4	*True*	*True*	*Yes*
5	*True*	*True*	*Yes*
6	*True*	*False*	*No*
7	*False*	*True*	*No*
8	*False*	*True*	*Yes*
9	*True*	*True*	*Yes*
10	*True*	*True*	*Yes*

a) Apply the ID3 decision tree learning algorithm to build a complete Decision Tree to decide if an accident is likely to occur on a given winter day. Show the decision tree and calculations. [To make calculations easier, you can use Table 1.]

Entropy of the Accident:

p=5 (yes); n=5 (no); I(5,5) = 1

- Consider the attribute Cloudy: T=6; F=4;
  When Cloudy=True, 4 Accident=yes and 2 Accident=no; I(4,2)
  When Cloudy=False, 1 Accident=yes and 3 Accident=no; I(1,3)

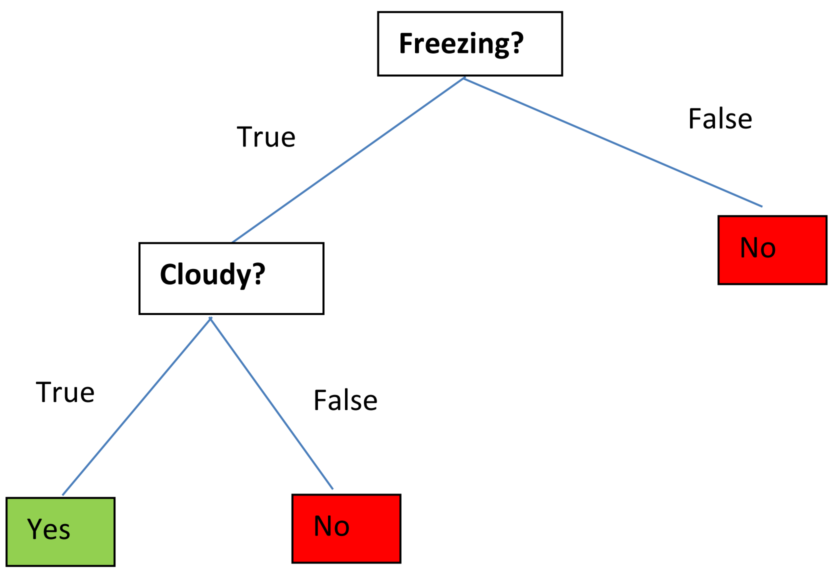
  Gain (Cloudy) = $1 - \left[\frac{6}{10} * \ I(4,2) + \frac{4}{10} * I(1,3)\right] = 1 - (0.6 * 0.918 + 0.4 * 0.811) = 0.1248$

- Consider the attribute Freezing: T=7; F=3;
  When Freezing =True, 5 Accident=yes and 2 Accident=no; I(5,2)
  When Freezing =False, 0 Accident=yes and 3 Accident=no; I(0,3)

  Gain (Freezing) = $1 - \left[\frac{7}{10} * I(5,2) + \frac{3}{10} * I(0,3)\right] = 1 - (0.7 * 0.863 + 0.3 * 0) = 0.3959$

Since Freezing attribute gives the highest information gain, we will choose the Freezing attribute as the root to start building the decision tree. With partition False of the Freezing attribute, we get all No to Accident, so the decision of the path from Freezing being False will be No accident.

Next, we will further split partition True of the Freezing attribute into whether or not Cloudy.



b) For your decision tree, what is the accuracy on the given dataset? (Hint: predict every case and compare predictions with the given Accident column)

Day	Cloudy	Freezing	Predictions	Accident?
1	*False*	*False*	***No***	*No*
2	*True*	*False*	***No***	*No*
3	*False*	*True*	***No***	*No*
4	*True*	*True*	***Yes***	*Yes*
5	*True*	*True*	***Yes***	*Yes*
6	*True*	*False*	***No***	*No*
7	*False*	*True*	***No***	*No*
8	*False*	*True*	***No***	*Yes*
9	*True*	*True*	***Yes***	*Yes*
10	*True*	*True*	***Yes***	*Yes*

My predictions are calculated in the green Predictions column above. So, my decision tree predicts correctly 9 out of 10, so the accuracy on the given dataset is 0.9 or 90%.

c) Does there exist a decision tree (that only uses Cloudy and Freezing as input for its decision) that gives 100% accuracy on the given training data? Justify your answer (1-2 sentences).

No, there does not exist a decision tree that gives 100% accuracy on the given training data. Because we had two choices for the root; either Cloudy or Freezing. I already created a tree with Freezing being the root in part (b), which gave an accuracy of 0.9. The only other decision tree can be formed with Cloudy being the root. That such tree predicts correctly 8 out of 10 times which is less accurate than part (b). Therefore, there does not exist a decision tree that only uses Cloudy and Freezing as input for its decision, which gives 100% accuracy on the given training data

2. Using the same data in Problem 1:

(a) Construct a **naïve Bayes classifier** to predict if an accident will occur on a particular day that is/is not cloudy and freezing.

Frequency Tables for each attribute:

Cloudy	Yes	No	
True	4	2	6
False	1	3	4
	5	5	10

Freezing	Yes	No	
True	5	2	7
False	0	3	3
	5	5	10

Prior probabilities:

P(Accident=yes) = 5/10 = 0.5

P(Accident=no) = 5/10 = 0.5

Conditional probabilities:

P(Accident=yes | Cloudy=True, Freezing=True) = P(Accident=yes) x P(Cloudy=True | Accident=yes) x P(Freezing=True | Accident=yes) = 0.5 x 4/5 x 5/5 = 0.4

P(Accident=yes | Cloudy=True, Freezing=False) = P(Accident=yes) x P(Cloudy=True | Accident=yes) x P(Freezing=False | Accident=yes) = 0.5 x 4/5 x 0/5 = 0

P(Accident=yes | Cloudy=False Freezing=True) = P(Accident=yes) x P(Cloudy= False | Accident=yes) x P(Freezing=True | Accident=yes) = 0.5 x 1/5 x 5/5 = 0.1

P(Accident=yes | Cloudy=False, Freezing=False) = P(Accident=yes) x P(Cloudy= False | Accident=yes) x P(Freezing= False | Accident=yes) = 0.5 x 1/5 x 0/5 = 0

P(Accident=no | Cloudy=True, Freezing=True) = P(Accident=no) * P(Cloudy=True | Accident=no) x P(Freezing=True | Accident=no) = 0.5 x 2/5 x 2/5 = 0.08

P(Accident= no | Cloudy=True, Freezing=False) = P(Accident= no) x P(Cloudy=True | Accident= no) x P(Freezing=False | Accident= no) = 0.5 x 2/5 x 3/5 = 0.12

P(Accident= no | Cloudy=False Freezing=True) = P(Accident= no) x P(Cloudy= False | Accident= no) x P(Freezing=True | Accident= no) = 0.5 x 3/5 x 2/5 = 0.12

P(Accident= no | Cloudy=False, Freezing=False) = P(Accident= no) x P(Cloudy= False | Accident= no) x P(Freezing= False | Accident= no) = 0.5 x 3/5 x 3/5 = 0.18

(b) Using the above naïve Bayes classifier, determine if an accident is likely to occur on a day that is both Freezing and Cloudy. Show calculations.

From part (a), we know that

P(Accident=yes | Cloudy=True, Freezing=True) = 0.4

P(Accident=no | Cloudy=True, Freezing=True) = 0.08

We are more likely to have an accident on a day that is both Freezing and Cloudy because the probability of having an accident, 0.4, is greater than the probability of not having an accident, 0.08.

(c) What is the accuracy of your naïve Bayes classifier on the given dataset? (Hint: predict every case and compare predictions with the given Accident column)

Case 1: When Cloudy=True, Freezing=True => Yes from part (b)

Case 2: When Cloudy=True, Freezing=False => No since the probability of not having an accident, 0.12, is greater than the probability of having it, 0.

P(Accident=yes | Cloudy=True, Freezing=False) = P(Accident=yes) x P(Cloudy=True | Accident=yes) x P(Freezing=False | Accident=yes) = 0.5 x 4/5 x 0/5 = 0

P(Accident=no | Cloudy=True, Freezing= False) = P(Accident=no) x P(Cloudy=True | Accident=no) x P(Freezing= False | Accident=no) = 0.5 x 2/5 x 3/5 = 0.12

Case 3: When Cloudy=False, Freezing=True => No as in case 2

Case 4: When Cloudy=False, Freezing=True => No since the probability of not having an accident, 0.12, is greater than the probability of having it, 0.1.

P(Accident=yes | Cloudy=False, Freezing=True) = P(Accident=yes) x P(Cloudy= False | Accident=yes) x P(Freezing=True | Accident=yes) = 0.5 x 1/5 x 5/5 = 0.1

P(Accident=no | Cloudy=False, Freezing=True) = P(Accident=no) x P(Cloudy=False | Accident=no) x P(Freezing= True | Accident=no) = 0.5 x 3/5 x 2/5 = 0.12

Case 5: When Cloudy=True, Freezing=False => No as in case 4

Case 6: When Cloudy=False, Freezing=False => No since the probability of not having an accident, 0.18, is greater than the probability of having it, 0.

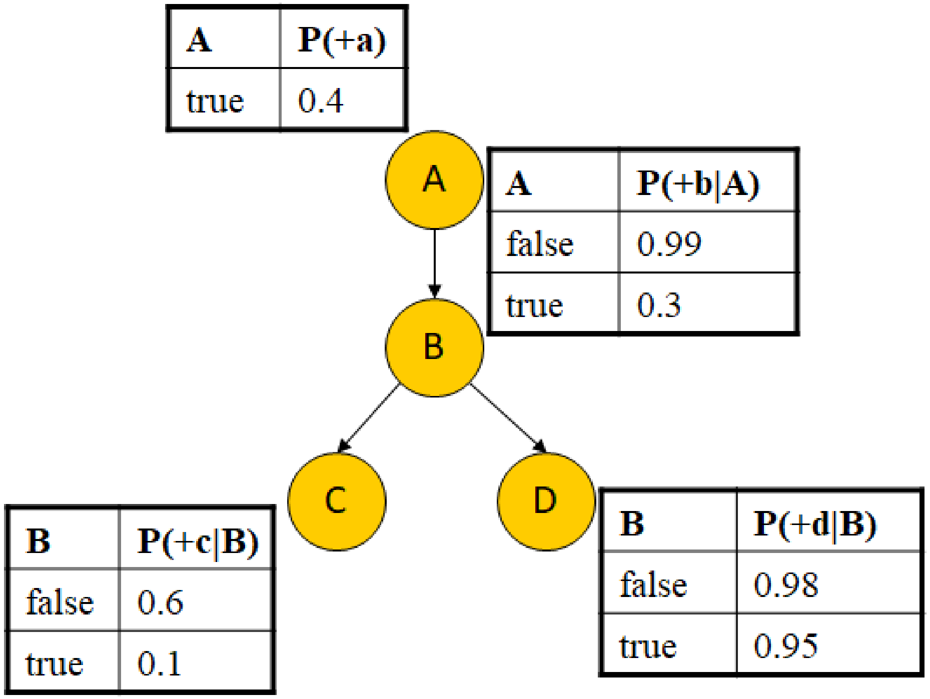
P(Accident=yes | Cloudy=False, Freezing=False) = P(Accident=yes) x P(Cloudy= False | Accident=yes) x P(Freezing= False | Accident=yes) = 0.5 x 1/5 x 0/5 = 0

P(Accident=no | Cloudy=False, Freezing=False) = P(Accident=no) v P(Cloudy=False | Accident=no) x P(Freezing= False | Accident=no) = 0.5 x 3/5 x 3/5 = 0.18

Day	Cloudy	Freezing	Predictions	Accident?
1	*False*	*False*	***No***	*No*
2	*True*	*False*	***No***	*No*
3	*False*	*True*	***No***	*No*
4	*True*	*True*	***Yes***	*Yes*
5	*True*	*True*	***Yes***	*Yes*
6	*True*	*False*	***No***	*No*
7	*False*	*True*	***No***	*No*
8	*False*	*True*	***No***	*Yes*
9	*True*	*True*	***Yes***	*Yes*
10	*True*	*True*	***Yes***	*Yes*

The table above shows the predictions column in green from Naïve Bayes classifier on this data, So, my Naïve Bayes classifier predicts correctly 9 out of 10, so the accuracy on the given dataset is 0.9 or 90%. As a conclusion, my decision tree in the question part (a) and Naïve Bayes classifier both have the same accuracy as 0.9 for this data.

3. Consider the following Bayesian Network (same network used for class work).



Calculate:

1. P(A=true, B=false | C=false, D=true)

   = P(A=true, B=false , C=false, D=true) / P(C=false, D=true)

   The numerator: P(A=true, B=false , C=false, D=true)

   = P(A = true) x P(B = false | A = true) x P(C = false | B = false) x P(D = true | B = false)

   = 0.4 x (1-0.3) x (1-0.6) x (0.98) = 0.4 x 0.7 x 0.4 x 0.98 = 0.10976

   Now, the denominator: P(C=false, D=true)

   = P (A=false, B=false, C=false, D=true) + P (A=false, B=true, C=false, D=true) + P (A=true, B=false, C=false, D=true) + P (A=true, B=true, C=false, D=true)

   P(A = false, B = false, C = false, D = true)

= P(A = false) x P(B = false | A = false) x P(C = false | B = false) x P(D = true | B = false)

= (1-0.4) x (1-0.99) x (1-0.6) x (0.98) = 0.6 x 0.01 x 0.4 x 0.98 = 0.002352

P(A = true, B = false, C = false, D = true) = 0.10976 from above

P(A = false, B = true, C = false, D = true)

= P(A = false) x P(B = true | A = false) x P(C = false | B = true) x P(D = true | B = true) =

(1-0.4) x 0.99 x (1-0.1) x 0.95 = 0.6 x 0.99 x 0.9 x 0.95 = 0.50787

P(A = true, B = true, C = false, D = true)

= P(A = true) x P(B = true | A = true) x P(C = false | B = true) x P(D = true | B = true) =

0.4 x 0.3 x (1-0.1) x 0.95 = 0.4 x 0.3 x 0.9 x 0.95 = 0.1026

Then, the denominator: P(C=false, D=true)

= 0.002352+0.10976+0.50787+0.1026 = 0.722582

Therefore, P(A=true, B=false | C=false, D=true) = 0.10976/0.722582 = **0.1519**

2. P(B=false | C=false, D=true)

   = P (B=false, C=false, D=true) / P (C=false, D=true)

   The numerator: P (B=false, C=false, D=true)

   = P (A=false, B=false, C=false, D=true) + P (A=true, B=false, C=false, D=true)

   P(A = false, B = false, C = false, D = true) = 0.002352 from part(1)

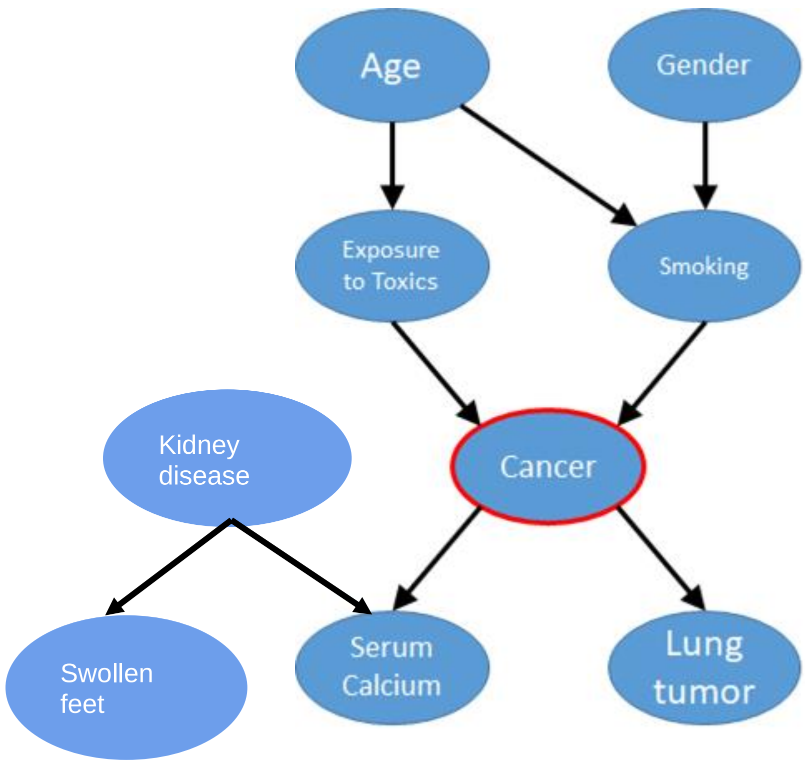
   P(A = true, B = false, C = false, D = true) = 0.10976 from part(1)

Then, the numerator: P(B=false, C=false, D=true) = 0.002352 + 0.10976 = 0.112112

Now, the denominator: P(C=false, D=true) = 0.722582 from part(1)

Therefore, P(B=false | C=false, D=true) = 0.112112/0.722582 = **0.1552**

4. Consider the Bayesian network shown below.

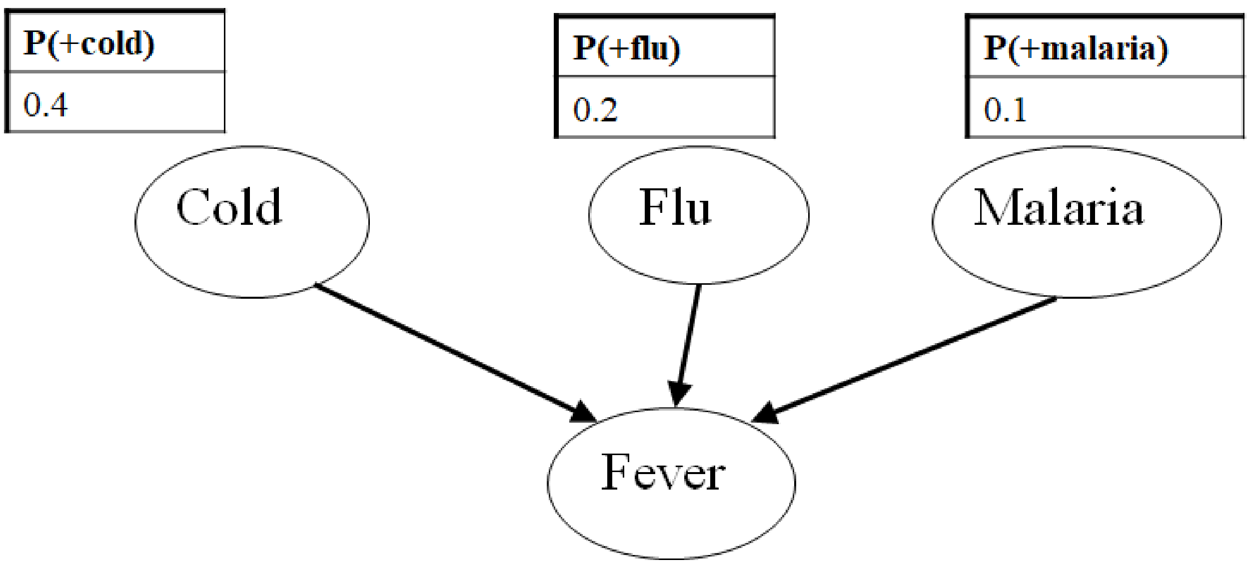


Which of the following statements about conditional independence is true?

1. Cancer is independent of Age = False
2. Cancer is independent of Serum calcium = False
3. Cancer is independent of Swollen feet = False
4. Cancer is independent of Age given Exposure to Toxics = False
5. Cancer is independent of Serum calcium given Exposure to Toxics = False
6. Cancer is independent of Swollen feet given Exposure to Toxics = False
7. Cancer is independent of Age given Exposure to Toxics and Smoking = True

8. Cancer is independent of Serum calcium given Exposure to Toxics and Smoking = False
9. Cancer is independent of Swollen feet given Exposure to Toxics and Smoking = True

5. Consider the Bayesian network shown below where the Conditional Probability table for the Fever variable is represented as a Noisy-OR (same network as in class work).



Inhibition probabilities for Fever:

- P(~Fever | cold, ~flu, ~malaria) = 0.6
- P(~Fever | ~cold, flu, ~malaria) = 0.2
- P(~Fever | ~cold, ~flu, malaria) = 0.1

Complete the Conditional Probability table at Fever:

Cold	**Flu**	**Malaria**	**P(+Fever|Cold,Flu,Malaria**
False	False	False	0 (1-1.0)
False	False	True	0.9 (1-0.1)
False	True	False	0.8 (1-0.2)
False	True	True	0.98 (1-(0.2 x 0.1))
True	False	False	0.4 (1-0.6)
True	False	True	0.94 (1-(0.6 x 0.1))
True	True	False	0.88 (1-(0.6 x 0.2))

True	True	True	0.988 (1-(0.6 x 0.2 x 0.1))

P(Cold=False, Flu=False, Malaria=False, Fever=False) = P(Cold=False) x P(Flu=False) x P(Malaria=False) x P(Fever=False | Cold=False, Flu=False, Malaria=False)

= 0.6 x 0.8 x 0.9 x 1.0 = 0.432

P(Cold=False, Flu=False, Malaria=False, Fever=True) = P(Cold=False) x P(Flu=False) x P(Malaria=False) x P(Fever=True | Cold=False, Flu=False, Malaria=False)

= 0.6 x 0.8 x 0.9 x 0 = 0

P(Cold=False, Flu=False, Malaria=True, Fever=False) = P(Cold=False) x P(Flu=False) x P(Malaria= True) x P(Fever=False | Cold=False, Flu=False, Malaria= True)

= 0.6 x 0.8 x 0.1 x 0.1= 0.0048

P(Cold=False, Flu=False, Malaria=True, Fever= True) = P(Cold=False) x P(Flu=False) x P(Malaria= True) x P(Fever= True | Cold=False, Flu=False, Malaria= True)

= 0.6 x 0.8 x 0.1 x 0.9= 0.0432

P(Cold=False, Flu=True, Malaria=False, Fever=False) = P(Cold=False) x P(Flu= True) x P(Malaria=False) x P(Fever=False | Cold=False, Flu= True, Malaria=False)

= 0.6 x 0.2 x 0.9 x 0.2 = 0.0216

P(Cold=False, Flu=True, Malaria=False, Fever= True) = P(Cold=False) x P(Flu= True) x P(Malaria=False) x P(Fever= True | Cold=False, Flu= True, Malaria=False)

= 0.6 x 0.2 x 0.9 x 0.8 = 0.0864

P(Cold=False, Flu=True, Malaria= True, Fever=False) = P(Cold=False) x P(Flu= True) x P(Malaria= True) x P(Fever=False | Cold=False, Flu= True, Malaria= True)

= 0.6 x 0.2 x 0.1 x 0.02 = 0.00024

P(Cold=False, Flu=True, Malaria= True, Fever= True) = P(Cold=False) x P(Flu= True) x P(Malaria= True) x P(Fever= True | Cold=False, Flu= True, Malaria= True)

= 0.6 x 0.2 x 0.1 x 0.98 = 0.01176

P(Cold=True, Flu=False, Malaria=False, Fever=False) = P(Cold= True) x P(Flu=False) x P(Malaria=False) x P(Fever= False | Cold= True, Flu=False, Malaria=False)

= 0.4 x 0.8 x 0.9 x 0.6 = 0.1728

P(Cold=True, Flu=False, Malaria=False, Fever= True) = P(Cold= True) x P(Flu=False) x P(Malaria=False) x P(Fever= True | Cold= True, Flu=False, Malaria=False)

= 0.4 x 0.8 x 0.9 x 0.4 = 0.1152

P(Cold=True, Flu=False, Malaria= True, Fever=False) = P(Cold= True) x P(Flu=False) x P(Malaria= True) x P(Fever= False | Cold= True, Flu=False, Malaria= True)

= 0.4 x 0.8 x 0.1 x 0.06 = 0.00192

P(Cold=True, Flu=False, Malaria= True, Fever= True) = P(Cold= True) x P(Flu=False) x P(Malaria= True) x P(Fever= True | Cold= True, Flu=False, Malaria= True)

= 0.4 x 0.8 x 0.1 x 0.94 = 0.03008

P(Cold=True, Flu= True, Malaria=False, Fever=False) = P(Cold= True) x P(Flu= True) x P(Malaria=False) x P(Fever= False | Cold= True, Flu= True, Malaria=False)

= 0.4 x 0.2 x 0.9 x 0.12 = 0.00864

P(Cold=True, Flu= True, Malaria=False, Fever= True) = P(Cold= True) x P(Flu= True) x P(Malaria=False) x P(Fever= True | Cold= True, Flu= True, Malaria=False)

= 0.4 x 0.2 x 0.9 x 0.88 = 0.06336

P(Cold=True, Flu= True, Malaria= True, Fever=False) = P(Cold= True) x P(Flu= True) x P(Malaria= True) x P(Fever= False | Cold= True, Flu= True, Malaria= True)

= 0.4 x 0.2 x 0.1 x 0.012 = 0.000096

P(Cold=True, Flu= True, Malaria= True, Fever= True) = P(Cold= True) x P(Flu= True) x P(Malaria= True) x P(Fever= True | Cold= True, Flu= True, Malaria= True)

= 0.4 x 0.2 x 0.1 x 0.988 = 0.007904

Complete the full Joint Probability table represented by the Bayesian network:

Cold	**Flu**	**Malaria**	**Fever**	**P(Cold,Flu,Malaria,Fever)**
False	False	False	False	0.432
False	False	False	True	0
False	False	True	False	0.0048
False	False	True	True	0.0432
False	True	False	False	0.0216
False	True	False	True	0.0864

False	True	True	False	0.00024
False	True	True	True	0.01176
True	False	False	False	0.1728
True	False	False	True	0.1152
True	False	True	False	0.00192
True	False	True	True	0.03008
True	True	False	False	0.00864
True	True	False	True	0.06336
True	True	True	False	0.000096
True	True	True	True	0.007904