*Assignment on Bayesian Network*

*Subject: Artificial Intelligence*

Submitted by

Debonil Ghosh (M21AIE225)

## Q1. Define independent events. What is conditional independence? Explain the role of conditional independence in probabilistic inference through Bayesian network with a real-life example (other than the one taught in class). How is the notion of Markov Blanket useful in this context? 2+2+4+2

**Answer:**

**Independent events** are those events whose occurrence is not dependent on any other event.

In terms of probability, events A and B are said to be **independent** if and only if their joint probability equals to the product of their probabilities.

$$P(A \cap B) = P(A)P(B)$$

**Conditional Independence:**

Two events A1 and A2 are said to be conditionally independent given event B with $P(B) > 0$, if and only if

$$P(A1 \cap A2 | B) = P(A1 | B) \cdot P(A2 | B). \quad (3)$$

Otherwise, we say events A1 and A2 are conditionally dependent given B.

Example:

Let’s say A is the height of a child and B is the number of words that the child knows. It seems when A is high, B is high too.

There is a single piece of information that will make A and B completely independent. What would that be?

The child’s age.

The height and the # of words known by the kid are NOT independent, but they are conditionally independent if you provide the kid’s age

**Conditional Independence in Bayesian Network**

A Bayesian network represents a joint distribution using a graph. Specifically, it is a directed acyclic graph in which each edge is a conditional dependency, and each node is a distinctive random variable. It has many other names: belief network, decision network, causal network, Bayes model or probabilistic directed acyclic graphical model, etc

Example:

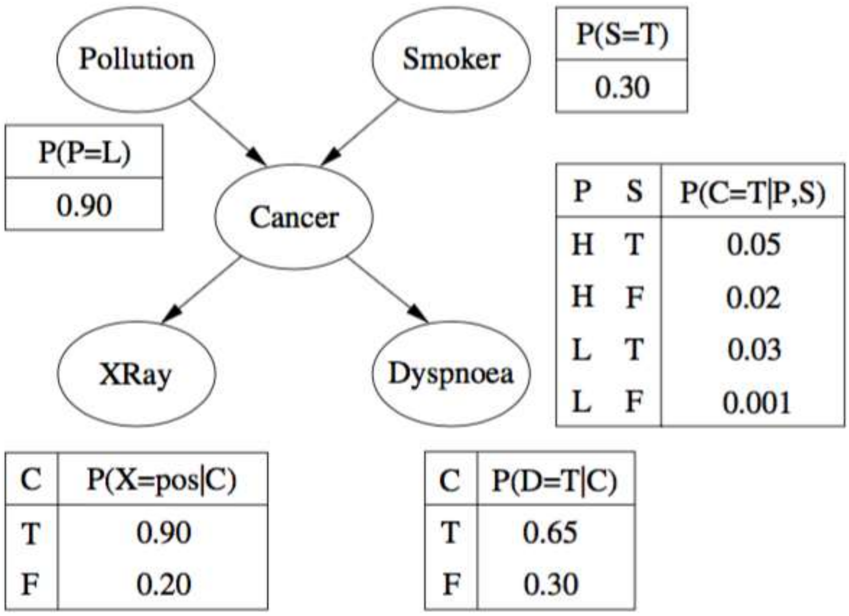


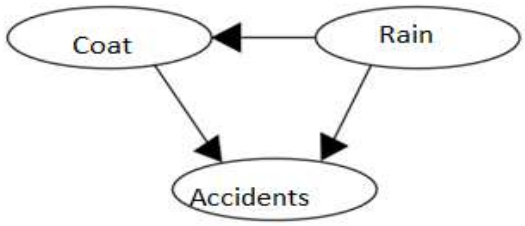

**FIGURE 2.1**
A BN for the lung cancer problem.

**Another Example:**

A study has shown a positive and significant correlation between the number of accidents and taxi drivers’ wearing coats. They found that coats might hinder the driver’s movements and cause accidents. A new law was ready to ban taxi drivers from wearing coats while driving.

Until another study pointed out that people wear coats when it rains

| Rain | Coat T | Coat F |
|---|---|---|
| F | 0.4 | 0.6 |
| T | 0.01 | 0.99 |




| Rain T | Rain F |
|---|---|
| 0.2 | 0.8 |

| Coat | Rain | T | F |
|---|---|---|---|
| F | F | 0.0 | 1.0 |
| F | T | 0.8 | 0.2 |
| T | F | 0.9 | 0.1 |
| T | T | 0.99 | 0.01 |

In For the Bayesian network to model a probability distribution, it relies on the important assumption: each variable is conditionally independent of its non-descendants, given its parents.

Conditional independence between variables can greatly reduce the number of parameters.

This reduces so much of the computation since we now only consider its parent and disregard everything else.

Let's take a look at the numbers.

Let's say you have n binary variables (= n nodes).

The unconstrained joint distribution requires O(2^n) probabilities.

For a Bayesian Network, with a maximum of k parents for any node, we need only O(n * 2^k) probabilities. (This can be carried out in linear time for certain numbers of classes.)

n = 30 binary variables, k = 4 maximum parents for nodes

• Unconstrained Joint Distribution: needs 2^30 (about 1 million) probabilities -> Intractable!

• Bayesian Network: needs only 480 probabilities

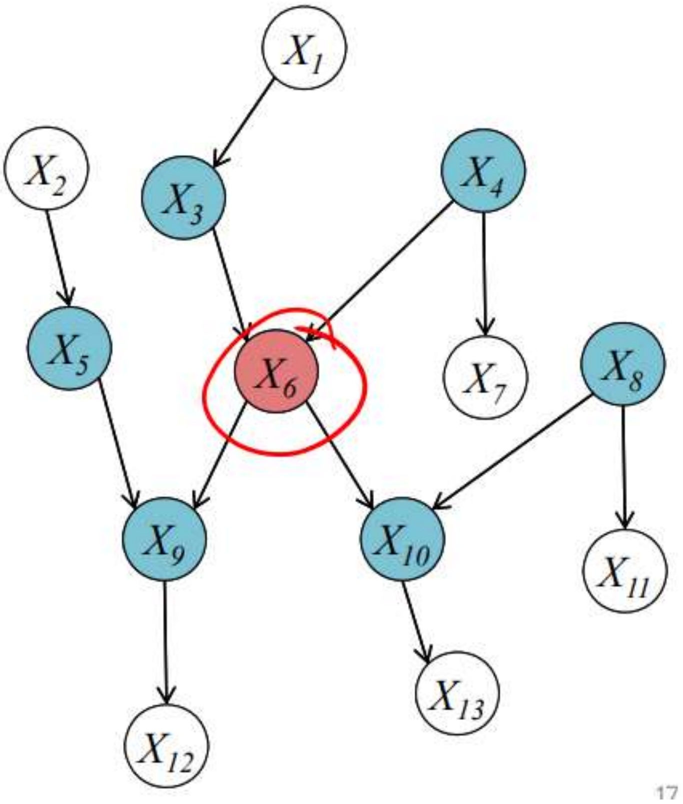
**Markov Blanket:** When one wants to infer a random variable with a set of variables, usually a subset is enough, and other variables are useless. Such a subset that contains all the useful information is called a Markov blanket. Markov blankets in a Bayesian network is used feature selection. The Markov blanket of a class attribute in a Bayesian network is a unique yet minimal feature subset for optimal feature selection if the probability distribution of a data set can be faithfully represented by this Bayesian network.

**Graphical Definition & Example:**

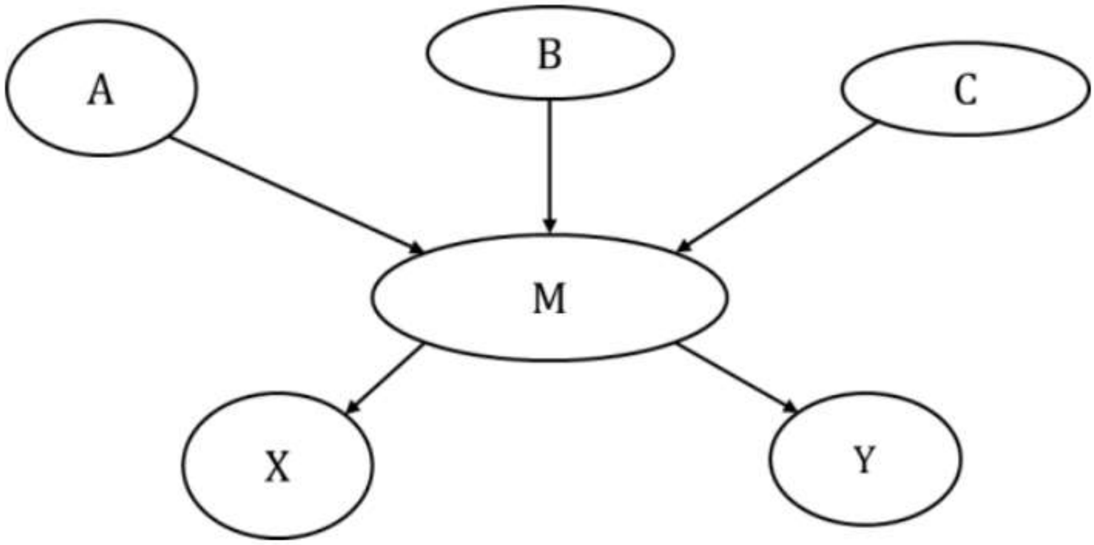
**Def:** The Markov Blanket of a node is the set containing the node's parents, children, and

co-parents

**Example: The Markov Blanket of X6 is {X3, X4, X5, X8, X9, X10}**



## Q2. Consider the following Bayesian network




The possible values of random variables are

| Random Variables | Values |
|---|---|
| $A$ | $a1, a2, a3$ |
| $B$ | $b1, b2$ |
| $C$ | $c1, c2$ |
| $M$ | $m1, m2, m3$ |
| $X$ | $x1, x2$ |
| $Y$ | $y1, y2, y3$ |

Write a pseudocode and a Python (strongly preferred)/ Java/ C++ implementation to find out P(A|x2, y3). You may assume the conditional probability tables satisfying the necessary constraints and then use those tables for the subsequent computations. Your code should output P(a1|x2, y3), P(a2|x2, y3), and P(a3|x2, y3). You need to submit the following.

a. The pseudocode

b. Your code and a file describing how to run the code

c. Manual computation (showing the steps) for P(a2|x2, y3)

10+10+10

**Answer:**

**a. The pseudocode**

```
f1(m):
    return P( x2 | m) * P(y3 | m)

f2(a, b, c):
    result = 0
    for each outcome m of Random Variable M:
        result += P(m|a,b,c) * f1(m)
    return result

f3(a):
    result = 0
    for each outcome b of Random Variable B:
        for each outcome c of Random Variable C:
            result += P(b) * P(c) * f2(a, b, c)
    return result

f4(a):
    return P(a) * f3(a)

pAX2Y3 = 0
```

```
for each outcome a of Random Variable A:
    pAX2Y3 += f4(a)

alpha = 1 / pAX2Y3

print alpha

for each outcome a of Random Variable A:
    print  'P(a{a+1}|x2, y3) = ' + (alpha * f4(a))
```

**b. Your code and a file describing how to run the code:**

**Python Code for the given problem:**

```python
# Events and outcomes

A = {'a1': 0, 'a2': 1, 'a3': 2}
B = {'b1': 0, 'b2': 1}
C = {'c1': 0, 'c2': 1}
M = {'m1': 0, 'm2': 1, 'm3': 2}
X = {'x1': 0, 'x2': 1}
Y = {'y1': 0, 'y2': 1, 'y3': 2}


# Conditional Probabilities as per given Bayesian Network


pA = [0.4, 0.35, 0.25]
pB = [0.6, 0.4]
pC = [0.3, 0.7]

pM = [
    # m1, m2, m3
    [0.8, 0.15, 0.05],  # a1,b1,c1
    [0.75, 0.1, 0.15],  # a2,b1,c1
    [0.7, 0.25, 0.05],  # a3,b1,c1
    [0.6, 0.02, 0.2],  # a1,b2,c1
    [0.55, 0.25, 0.02],  # a2,b2,c1
    [0.4, 0.5, 0.01],  # a3,b2,c1
    [0.4, 0.1, 0.5],  # a1,b1,c2
    [0.1, 0.5, 0.4],  # a2,b1,c2
    [0.1, 0.2, 0.7],  # a3,b1,c2
```

```
    [0.1, 0.4, 0.5],  # a1,b2,c2
    [0.01, 0.01, 0.98],  # a2,b2,c2
    [0.01, 0.14, 0.85],  # a3,b2,c2
]

pX = [
    # x1, x2
    [0.25, 0.75],  # m1
    [0.90, 0.10],  # m2
    [0.28, 0.72],  # m3
]

pY = [
    # y1, y2, y3
    [0.8, 0.15, 0.05],  # m1
    [0.15, 0.25, 0.6],  # m2
    [0.02, 0.18, 0.8],  # m3
]


# Now need to find  P(a1|x2, y3), P(a2|x2, y3), and P(a3|x2, y3)

# P( A| x2, y3 )= α(P(A) Sum-B P(B) Sum-C P(C) Sum-M P(M|A,B,C) P( x2 | M)
P(y3 | M))


# P( x2 | M) P(y3 | M)


def f1(m):
    return pX[m][X['x2']]*pY[m][Y['y3']]


# P(M|A,B,C) f1(M)

def f2(a, b, c):
    res = 0
    for m in M.values():
        res += pM[a+b*3+c*6][m] * f1(m)
    return res

# Sum-B P(B) Sum-C P(C) f2(a, b, c)


def f3(a):
    res = 0
    for b in B.values():
        for c in C.values():
            res += pB[b] * pC[c] * f2(a, b, c)
```

```
    return res


def f4(a):
    return pA[a] * f3(a)


pAX2Y3 = 0

for a in A.values():
    pAX2Y3 += f4(a)


alpha = 1 / pAX2Y3

print(f'Alpha = {alpha}\n')


for a in A.values():
    print(f'P(a{a+1}|x2, y3) = {alpha * f4(a)}\n')
```

---

**How to run above code:**

1. Above code may be copy-pasted in a text file with name *bayesian-network.py.*
2. Python 3 need to be installed if not there already.
3. Run command " *py bayesian-network.py* "

**Output of the above code:**

Alpha = 3.494137562204166

$P(a1|x2, y3)$ = 0.34530744575731426

$P(a2|x2, y3)$ = 0.3627536920760874

$P(a3|x2, y3)$ = 0.29193886216659837

**c. Manual computation (showing the steps) for P(a2|x2, y3):**

$$P(\, A|\, x2, y3\,) = \alpha(P(A) * \sum_B P(B) * \sum_C P(C) * \sum_M P(M|A,B,C) * P(\, x2 \mid M) * P(y3 \mid M))$$

1. **Calculating $\mathbf{P(\, x2 \mid M) * P(y3 \mid M)) = f1(M)}$ from given condition probability table:**

$P(X|M)$ = [

  # x1, x2

  [0.25, 0.75],  # m1

  [0.90, 0.10],  # m2

  [0.28, 0.72],  # m3

]

$P(Y|M)$ = [

  # y1, y2, y3

  [0.8, 0.15, 0.05],  # m1

  [0.15, 0.25, 0.6],  # m2

  [0.02, 0.18, 0.8],  # m3

]

$\mathrm{f1(M)} = P(\, x2 \mid M) * P(y3 \mid M) =$

M=m1 | 0.75 * 0.05 = 0.0375

M=m2 | 0.1 * 0.6 = 0.06

M=m3 | 0.72 * 0.8 = 0.576

2. **Calculating $\sum_M \boldsymbol{P(M|A,B,C)} * \mathbf{f1(M) = f2(A,B,C)}$ with help of above table:**

$\boldsymbol{P}(\boldsymbol{M}|\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) =$

```
# m1, m2, m3
[0.8, 0.15, 0.05],  # a1,b1,c1
[0.75, 0.1, 0.15],  # a2,b1,c1
[0.7, 0.25, 0.05],  # a3,b1,c1
[0.6, 0.02, 0.2],  # a1,b2,c1
[0.55, 0.25, 0.02],  # a2,b2,c1
[0.4, 0.5, 0.01],  # a3,b2,c1
[0.4, 0.1, 0.5],  # a1,b1,c2
[0.1, 0.5, 0.4],  # a2,b1,c2
[0.1, 0.2, 0.7],  # a3,b1,c2
[0.1, 0.4, 0.5],  # a1,b2,c2
[0.01, 0.01, 0.98],  # a2,b2,c2
[0.01, 0.14, 0.85],  # a3,b2,c2
```

$$\mathbf{f2}(\mathbf{A},\mathbf{B},\mathbf{C}) = \sum_{\boldsymbol{M}} \boldsymbol{P}(\boldsymbol{M}|\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}) * \mathbf{f1}(\mathbf{M}) =$$

∑_M 〖P(M|a1, b1, c1) * f1(M)〗 = ( 0.8 * 0.0375 ) + ( 0.15 * 0.06 ) + ( 0.05 * 0.576 ) = 0.0678

∑_M 〖P(M|a1, b1, c2) * f1(M)〗 = ( 0.4 * 0.0375 ) + ( 0.1 * 0.06 ) + ( 0.5 * 0.576 ) = 0.309

∑_M 〖P(M|a1, b2, c1) * f1(M)〗 = ( 0.6 * 0.0375 ) + ( 0.02 * 0.06 ) + ( 0.2 * 0.576 ) = 0.1389

∑_M 〖P(M|a1, b2, c2) * f1(M)〗 = ( 0.1 * 0.0375 ) + ( 0.4 * 0.06 ) + ( 0.5 * 0.576 ) = 0.31575

∑_M 〖P(M|a2, b1, c1) * f1(M)〗 = ( 0.75 * 0.0375 ) + ( 0.1 * 0.06 ) + ( 0.15 * 0.576 ) = 0.120525

∑_M 〖P(M|a2, b1, c2) * f1(M)〗 = ( 0.1 * 0.0375 ) + ( 0.5 * 0.06 ) + ( 0.4 * 0.576 ) = 0.26415

∑_M 〖P(M|a2, b2, c1) * f1(M)〗 = ( 0.55 * 0.0375 ) + ( 0.25 * 0.06 ) + ( 0.02 * 0.576 ) = 0.047145

∑_M 〖P(M|a2, b2, c2) * f1(M)〗 = ( 0.01 * 0.0375 ) + ( 0.01 * 0.06 ) + ( 0.98 * 0.576 ) = 0.565455

∑_M 〖P(M|a3, b1, c1) * f1(M)〗 = ( 0.7 * 0.0375 ) + ( 0.25 * 0.06 ) + ( 0.05 * 0.576 ) = 0.07005

∑_M 〖P(M|a3, b1, c2) * f1(M)〗 = ( 0.1 * 0.0375 ) + ( 0.2 * 0.06 ) + ( 0.7 * 0.576 ) = 0.41895

∑_M 〖 P(M|a3, b2, c1) * f1(M) 〗 = ( 0.4 * 0.0375 ) + ( 0.5 * 0.06 ) + ( 0.01 * 0.576 ) = 0.05076

∑_M 〖 P(M|a3, b2, c2) * f1(M) 〗 = ( 0.01 * 0.0375 ) + ( 0.14 * 0.06 ) + ( 0.85 * 0.576 ) = 0.498375

3. **Calculating $\sum_B \boldsymbol{P(B)} \sum_C \boldsymbol{P(C)} * \mathbf{f2(A, B, C)}$ with help of above table:**

$$f3(A) = \sum_B P(B) * \sum_C P(C) * f2(A, B, C) =$$

$\boldsymbol{P(B)}$= [0.6, 0.4]

$\boldsymbol{P(C)}$= [0.3, 0.7]

$$\begin{aligned} f3(\boldsymbol{a1}) &= P(b1) * P(c1) * f2(a1, b1, c1) + P(b1) * P(c2) * f2(a1, b1, c2) + P(b2) * P(c1) \\ &\quad * f2(a1, b2, c1) + P(b2) * P(c2) * f2(a1, b2, c2) \\ &= (0.6 * 0.3 * 0.0678) + (0.6 * 0.7 * 0.309) + (0.4 * 0.3 * 0.1389) + (0.4 \\ &\quad * 0.7 * 0.31575) \\ &= 0.24706199999999998 \end{aligned}$$

$$\begin{aligned} f3(\boldsymbol{a2}) &= P(b1) * P(c1) * f2(a2, b1, c1) + P(b1) * P(c2) * f2(a2, b1, c2) + P(b2) * P(c1) \\ &\quad * f2(a2, b2, c1) + P(b2) * P(c2) * f2(a2, b2, c2) \\ &= (0.6 * 0.3 * 0.120525) + (0.6 * 0.7 * 0.26415) + (0.4 * 0.3 \\ &\quad * 0.047145000000000006) + (0.4 * 0.7 * 0.5654549999999999) \\ &= 0.29662229999999995 \end{aligned}$$

$$\begin{aligned} f3(\boldsymbol{a3}) &= P(b1) * P(c1) * f2(a3, b1, c1) + P(b1) * P(c2) * f2(a3, b1, c2) + P(b2) * P(c1) \\ &\quad * f2(a3, b2, c1) + P(b2) * P(c2) * f2(a3, b2, c2) \\ &= (0.6 * 0.3 * 0.07005) + (0.6 * 0.7 * 0.41894999999999993) + (0.4 * 0.3 \\ &\quad * 0.05076) + (0.4 * 0.7 * 0.4983749999999999) \\ &= 0.33420419999999995 \end{aligned}$$

4. **Calculating $\boldsymbol{P(A)} * \boldsymbol{f3(A)}$ with help of above table:**

$\boldsymbol{P(A)}$ = [0.4, 0.35, 0.25]

$$f4(A) = P(A) * f3(A) =$$

$$f4(\boldsymbol{a1}) = P(a1) * f3(a1)$$

$$= 0.4 * 0.24706199999999998$$

$$= 0.09882479999999999$$

$f4(\boldsymbol{a2}) = P(a2) * f3(a2)$

$= 0.35 * 0.29662229999999995$

$= 0.10381780499999997$

$f4(\boldsymbol{a3}) = P(a3) * f3(a3)$

$= 0.25 * 0.33420419999999995$

$= 0.08355104999999999$

5. **Calculating $\alpha$ with help of above data:**

$$\alpha * \big(f4(a1) + f4(a2) + f4(a3)\big) = 1$$

$$\alpha * (0.09882479999999999 + 0.103817805 + 0.08355105) = 1$$

$$\alpha = \frac{1}{0.28619365} = 3.494137562204166$$

6. **Calculating $\mathbf{P(a2|x2, y3)}$ with help of above data:**

$$P(a2|x2, y3) = \alpha * f4(a2)$$
$$= 3.494137562204166 * 0.1038178045$$
$$= 0.3627536920760874$$