Elective Course
Course Code: CS4103
Autumn 2025-26

Lecture #30

Artificial Intelligence for Data Science

Week-8:

Introduction to Probabilistic Reasoning [Part-II]

Reasoning Under Uncertainty and Probability Theory

Course Instructor:

Dr. Monidipa Das
Assistant Professor
Department of Computational and Data Sciences
Indian Institute of Science Education and Research Kolkata, India 741246

Bayesian Network

- Probabilistic graphical model
- Represents dependence/independence via *directed acyclic graph* (DAG)
- Structure of the graph ⬄ Conditional independence relations

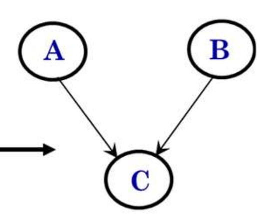
In general,

$$P(X1, X2, \ldots. XN) = \prod P(Xi \mid parents(Xi))$$

The full joint distribution   *The graph-structured approximation*

- Components of Bayesian network
  - *Qualitative: Graph structure*
  - *Quantitative: Numerical probabilities*

$$P(A,B,C) = P(C \mid A,B)P(A)P(B) \longleftrightarrow$$

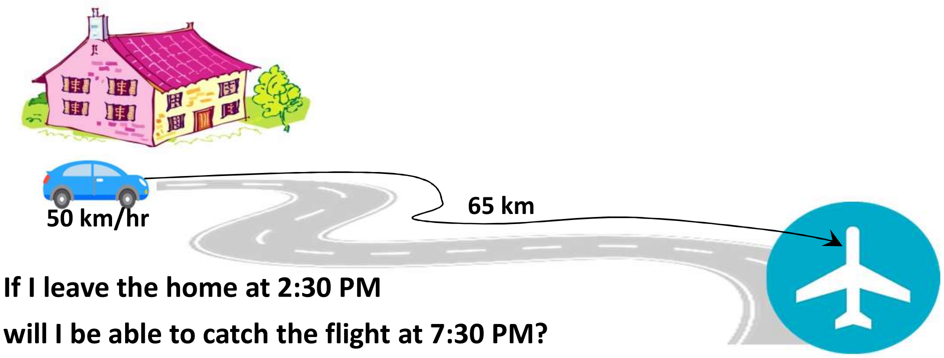


Example

3

Variables in the study:

1. Rush-hour (yes/no)
2. Bad-weather (yes/no)
3. Accident (yes/no)
4. Traffic-jam (heavy/light)
5. Miss-Flight (yes/no)



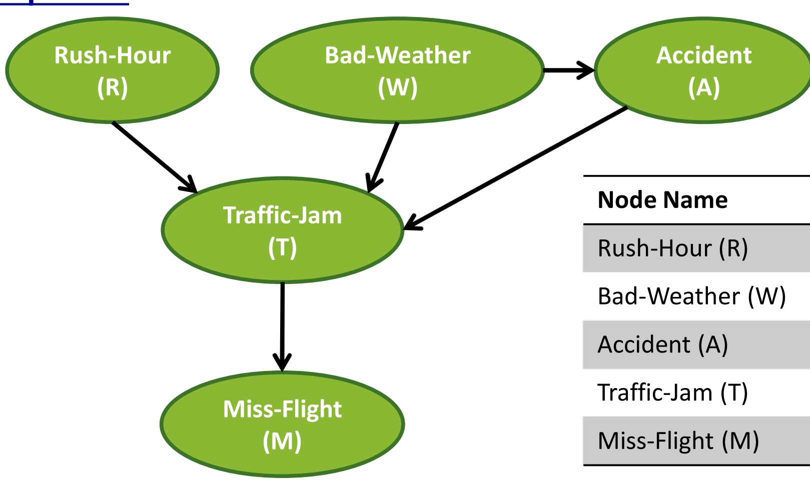
If I leave the home at 2:30 PM

will I be able to catch the flight at 7:30 PM?

Example

4

Qualitative Component



Node Name	Type	Values
Rush-Hour (R)	Binary	{Yes (Y), No (N)}
Bad-Weather (W)	Binary	{Yes (Y), No (N)}
Accident (A)	Binary	{Yes (Y), No (N)}
Traffic-Jam (T)	Binary	{Heavy (H), Light (L)}
Miss-Flight (M)	Binary	{Yes (Y), No (N)}

Example

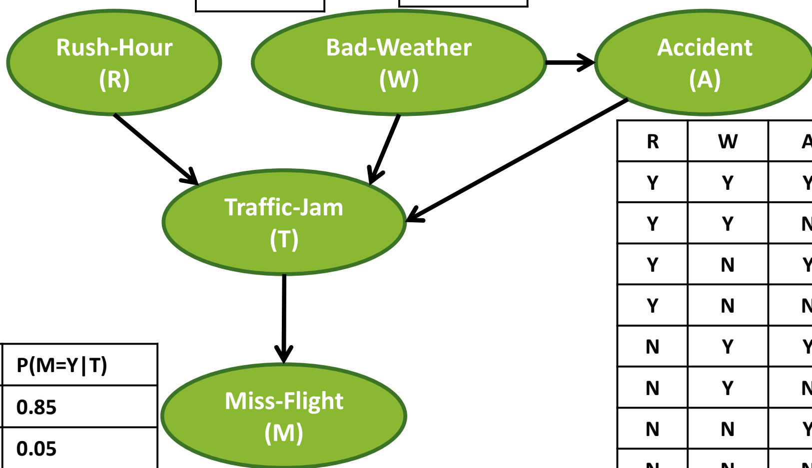
5

Quantitative Component

P(R=Y)
0.25

P(W=Y)
0.5

W	P(A=Y|W)
Y	0.75
N	0.10



R	W	A	P(T=H|R, W, A)
Y	Y	Y	0.99
Y	Y	N	0.92
Y	N	Y	0.96
Y	N	N	0.85
N	Y	Y	0.95
N	Y	N	0.70
N	N	Y	0.90
N	N	N	0.05

T	P(M=Y|T)
H	0.85
L	0.05

$$\boldsymbol{P(R,W,A,T,M) = P(R).P(W).P(A|W).P(T|R,W,A).P(M|T)}$$

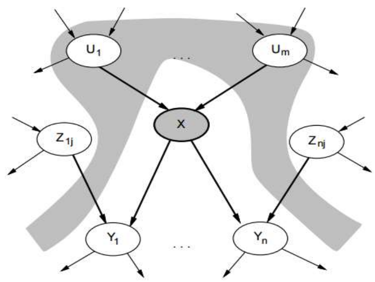
Advantage

6

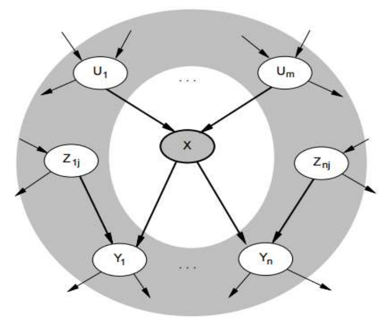
- If we have a Bayesian network, with a maximum of k parents for any node, then we need $O(n.2^k)$ probabilities
  - Bayesian network
    - n = 40, k = 3: need 320 probabilities
  - Full joint distribution
    - n = 40: need 10^{12} probabilities for full joint distribution

Conditional Independence Relations in Bayesian Network

1. Each node is conditionally independent of its non-descendants given its parents



2. Each node is conditionally independent of all others given its Markov blanket: (parents + children + children's parents)



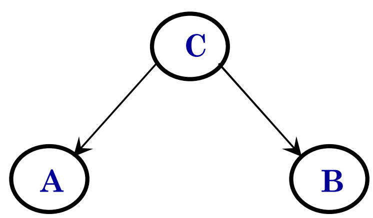
Conditional Independence

Conditionally independent effects:

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

$$\begin{aligned} P(A,B,C) &= P(A|B,C)P(B,C) \\ &= P(A|B,C)P(B|C)P(C) \\ &= P(A|C)P(B|C)P(C) \end{aligned}$$

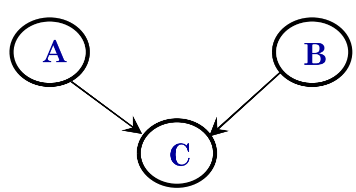
A and B are conditionally independent Given C



Independent Causes:

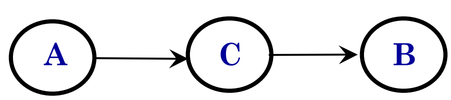
$$P(A,B,C) = P(C|A,B)P(A)P(B)$$

A and B are (marginally) independent
but become dependent once C is known



Markov dependence:

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



Inference in Bayesian Networks

9

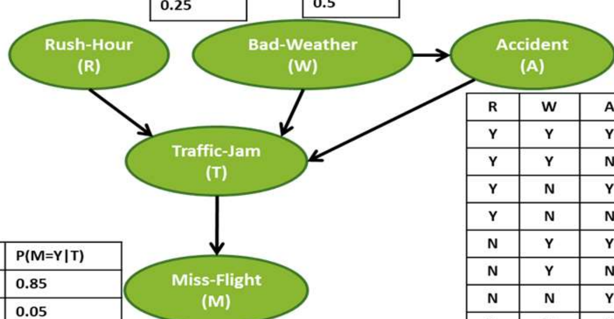
- Consider answering a query in a Bayesian Network
  - Q = set of query variables
  - e = evidence (set of instantiated variable-value pairs)
  - Inference = computation of conditional distribution $P(Q|e)$
- **Examples**
  - $P(MissFlight|BadWeather = Y)$
  - $P(MissFlight\ |RushHour = N)$
  - $P(TrafficJam, Accident|BadWeather = N, RushHour = Y)$

Inference in Bayesian Networks

10

P(R=Y)
0.25

P(W=Y)
0.5



W	P(A=Y|W)
Y	0.75
N	0.10

R	W	A	P(T=H|R, W, A)
Y	Y	Y	0.99
Y	Y	N	0.92
Y	N	Y	0.96
Y	N	N	0.85
N	Y	Y	0.95
N	Y	N	0.70
N	N	Y	0.90
N	N	N	0.05

T	P(M=Y|T)
H	0.85
L	0.05

Exact Method

$$P(MissFlight|BadWeather = Y) \equiv P(M|W = Y)$$

$$= P(M, W = Y)|P(W = Y) = \gamma . P(M, W = Y) = \gamma \sum_R \sum_A \sum_T P(R, W = Y, A, T, M)$$

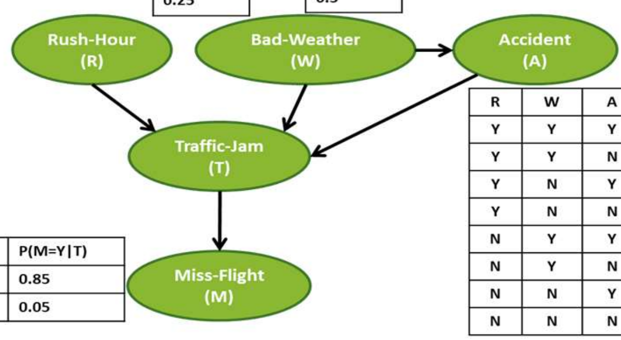
$$= \gamma \sum_R \sum_A \sum_T P(R). P(W = Y). P(A|W = Y). P(T|R, W = Y, A). P(M|T)$$

Inference in Bayesian Networks

11

P(R=Y)
0.25

P(W=Y)
0.5



W	P(A=Y|W)
Y	0.75
N	0.10

R	W	A	P(T=H|R, W, A)
Y	Y	Y	0.99
Y	Y	N	0.92
Y	N	Y	0.96
Y	N	N	0.85
N	Y	Y	0.95
N	Y	N	0.70
N	N	Y	0.90
N	N	N	0.05

T	P(M=Y|T)
H	0.85
L	0.05

$$P(MissFlight|BadWeather = Y, Accident = N) \equiv P(M|W = Y, A = N)$$

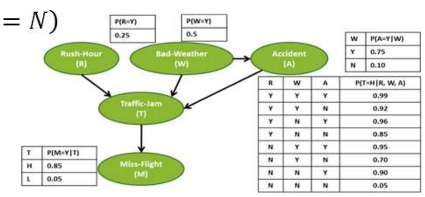
$$= P(M, W = Y, A = N)|P(W = Y, A = N) = \gamma . P(M, W = Y, A = N) = \gamma \sum_{T} \sum_{R} P(R, W = Y, A = N, T, M)$$

$$= \gamma(P(R = Y, W = Y, A = N, T = H, M) + P(R = N, W = Y, A = N, T = H, M) + P(R = Y, W = Y, A = N, T = L, M) + P(R = N, W = Y, A = N, T = L, M))$$

Inference in Bayesian Networks

12

$$P(MissFlight = Y|BadWeather = Y, Accident = N) \equiv P(M = Y|W = Y, A = N)$$



$$= \gamma(P(R = Y, W = Y, A = N, T = H, M = Y) + P(R = N, W = Y, A = N, T = H, M = Y) + P(R = Y, W = Y, A = N, T = L, M = Y) + P(R = N, W = Y, A = N, T = L, M = Y))$$

$$= \gamma(P(R = Y). P(W = Y). P(A = N|W = Y). P(T = H|R = Y, W = Y, A = N). P(M = Y|T = H) + P(R = N). P(W = Y). P(A = N|W = Y). P(T = H|R = N, W = Y, A = N). P(M = Y|T = H) + P(R = Y). P(W = Y). P(A = N|W = Y). P(T = L|R = Y, W = Y, A = N). P(M = Y|T = L) + P(R = N). P(W = Y). P(A = N|W = Y). P(T = L|R = N, W = Y, A = N). P(M = Y|T = L))$$

$$= \gamma((0.25 * 0.5 * 0.25 * 0.92 * 0.85) + (0.75 * 0.5 * 0.25 * 0.70 * 0.85) + (0.25 * 0.5 * 0.25 * 0.08 * 0.05) + (0.75 * 0.5 * 0.25 * 0.3 * 0.05))$$

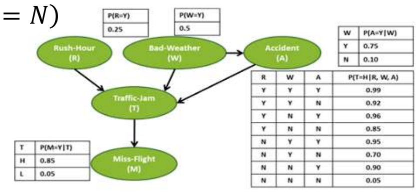
$$= \gamma(0.0244 + 0.0558 + 0.0001 + 0.0014)$$

$$= 0.082\gamma$$

Inference in Bayesian Networks

13

$P(MissFlight = N|BadWeather = Y, Accident = N) \equiv P(M = N|W = Y, A = N)$



$= \gamma(P(R = Y, W = Y, A = N, T = H, M = N) + P(R = N, W = Y, A = N, T = H, M = N) + P(R = Y, W = Y, A = N, T = L, M = N) + P(R = N, W = Y, A = N, T = L, M = N))$

$= \gamma(P(R = Y). P(W = Y). P(A = N|W = Y). P(T = H|R = Y, W = Y, A = N). P(M = N|T = H) + P(R = N). P(W = Y). P(A = N|W = Y). P(T = H|R = N, W = Y, A = N). P(M = N|T = H) + P(R = Y). P(W = Y). P(A = N|W = Y). P(T = L|R = Y, W = Y, A = N). P(M = N|T = L) + P(R = N). P(W = Y). P(A = N|W = Y). P(T = L|R = N, W = Y, A = N). P(M = N|T = L))$

$= \gamma((0.25 * 0.5 * 0.25 * 0.92 * 0.15) + (0.75 * 0.5 * 0.25 * 0.70 * 0.15) + (0.25 * 0.5 * 0.25 * 0.08 * 0.95) + (0.75 * 0.5 * 0.25 * 0.3 * 0.95))$

$= \gamma(0.004 + 0.0098 + 0.0024 + 0.0267)$

$= 0.043\gamma$

Inference in Bayesian Networks

14

$$P(MissFlight = Y|BadWeather = Y, Accident = N) + P(MissFlight = N|BadWeather = Y, Accident = N) = 1$$

$$\Rightarrow P(M = Y|W = Y, A = N) + P(M = N|W = Y, A = N) = 1$$
$$\Rightarrow 0.082\gamma + 0.043\gamma = 1$$
$$\Rightarrow 0.125\gamma = 1$$
$$\Rightarrow \gamma = 8$$

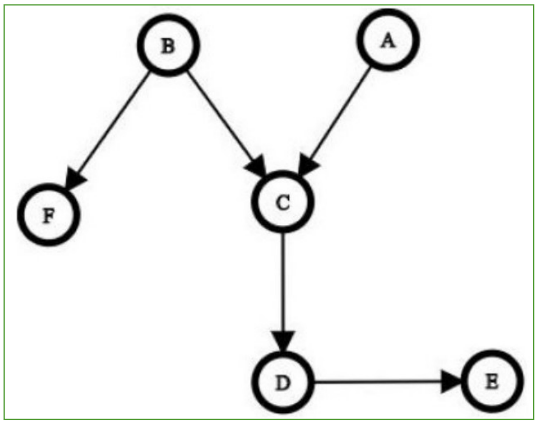
$\therefore P(MissFlight = Y|BadWeather = Y, Accident = N) = 0.082 * 8 = 0.656$

$\therefore P(MissFlight = N|BadWeather = Y, Accident = N) = 0.043 * 8 = 0.344$

Do it now

15

Consider the following Bayesian Network. Given evidence about C which of the following pair of variables are conditionally independent.

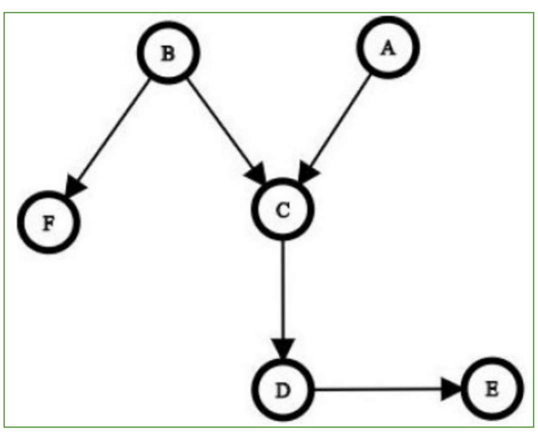


- A and B
- A and D
- D and E
- A and F

Do it now

16

Consider the following Bayesian Network.
Which of the following variables belong to Markov Blanket of F.

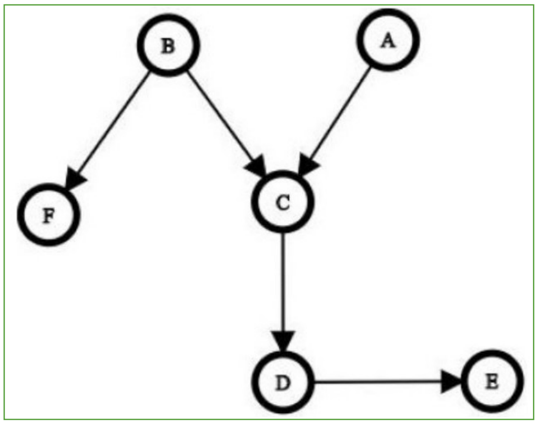


- A
- B
- C
- D

Do it now

17

How many minimum parameters are required to represent the following Bayesian Network, if each variable is Boolean?



Solving Homework

18

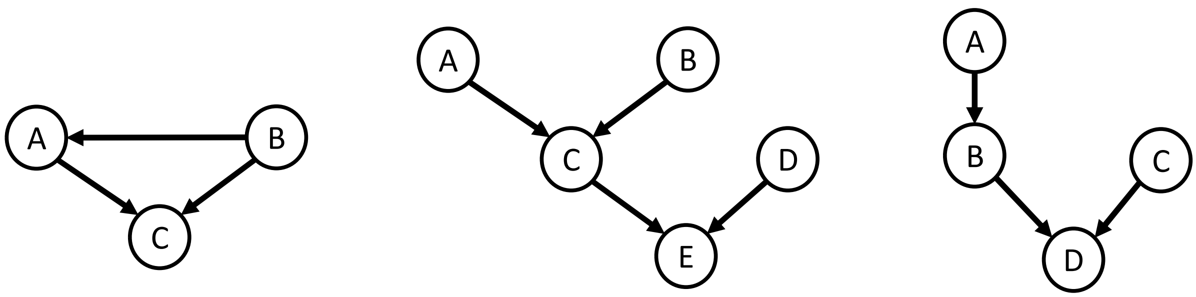
- **Given the full joint distribution shown below, calculate the following:**
  - **a. P(toothache).**
  - **b. P(Cavity).**
  - **c. P(Cavity | toothache ∨ catch).**

	$toothache$		$\neg toothache$	
	$catch$	$\neg catch$	$catch$	$\neg catch$
$cavity$	0.108	0.012	0.072	0.008
$\neg cavity$	0.016	0.064	0.144	0.576

Solved during lecture session

Solving Homework

Given the following networks, write the graph structured approximation of the full joint distributions



(a) (b) (c)

Solved during lecture session

Questions?

Slide Content taken from Prof. Cesare Tinelli, Prof. Stuart Russell, and Prof. Jim Martin