BBN

BBNs are networks that represent the probabilistic relationship between various classes

Defination → A BBN is a probabilistic graphical model which represents a set of variables and their conditional dependancies using a directed cyclic graph

They are built on probability distribution

They use probability for prediction

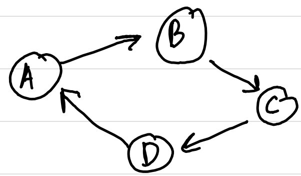
eg the probability of rain is given to the model as 0.6 then the model will predict it will rain.

BBNs address the issue of uncertainty in AI

BBNs are trained by Maximum likelyhood estimation, expectation Maximization, expert knowledge

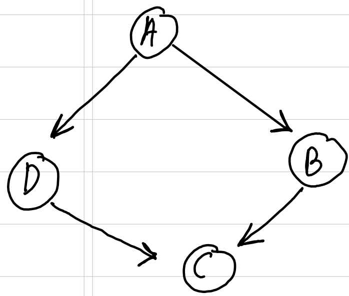
Useful for reasoning under uncertainty

Since it is a acyclic graph, the graph has no cycles



Not Possible

A BNN has Nodes → events
edges → probability of those edges



A B C D are random variables

A is parent of B

B is dependant on A

C is dependant on B
∴ C is dependant on A

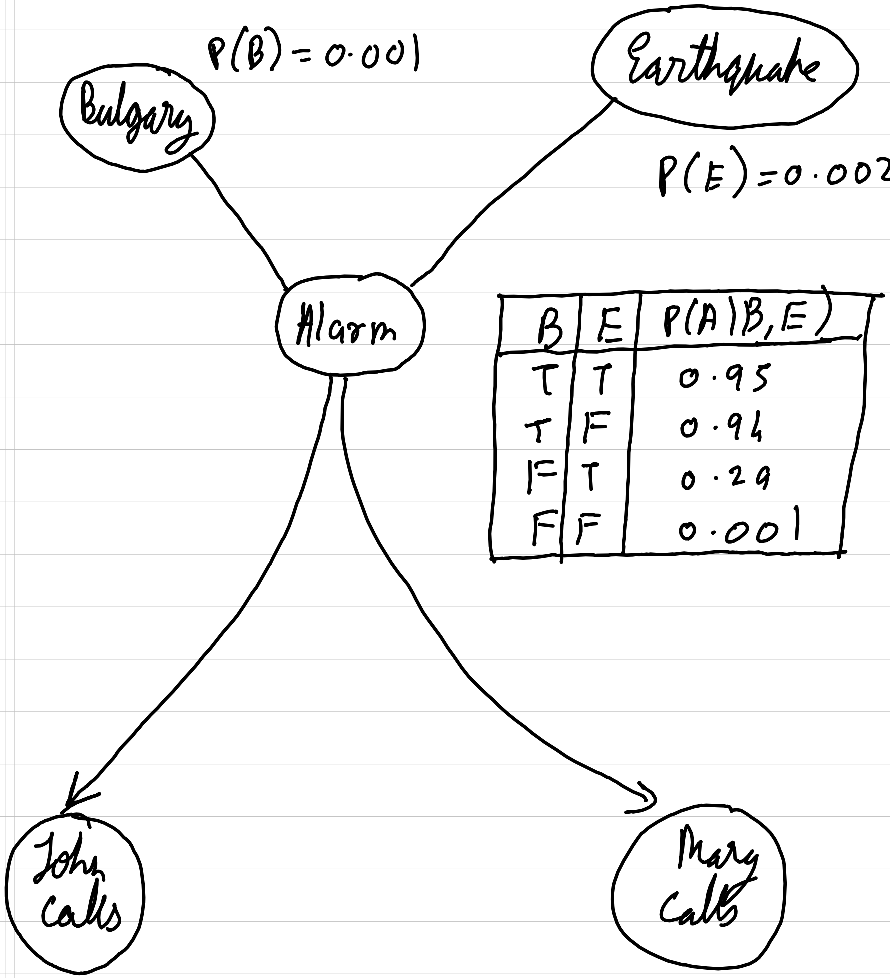
A is NOT dependant on B, C, D → Local Markov Property

BBN have Local Markov property

This means that a node is conditionally independant of its nondescendents given its parents

Basically you depend only on parents & not on children

Consider the following Network



B	E	P(A|B,E)
T	T	0.95
T	F	0.94
F	T	0.29
F	F	0.001

A	P(J|A)
T	0.96
F	0.05

A	P(M|A)
T	0.70
F	0.01

Probability of Bulgary $\rightarrow 0.001$

Probability of Alarm going off when bulgary occurs and earthquake occurs $\rightarrow 0.94$

Probability of John not calling when alarm does not occur

$\rightarrow 1 - 0.05 = 0.95$

Probability of John calling when Earthquake occurs?

$= P(J \mid A=T)$ John calls when alarm goes on
$\times P(A \mid E=T)$ & alarm does go on when E

$+ P(J \mid A=F)$ John calls even if alarm doesnt go on
$\times P(A=F \mid E=T)$ & alarm doesnt go on when E

$$P(J \mid A) P(A \mid E) + P(J \mid \sim A) P(\sim A \mid E)$$

Probability of Bulgary occuring given Earthquake occurs

$\rightarrow$ E & B are independant events

$P(B \mid E) = P(B)$

Joint probability distribution

If we have variables $x_1\ x_2\ x_3 \ldots x_n$ then the probabilities of different combination of $x_1\ x_2\ x_3 \ldots x_n$ are known as joint probability distribution

$$P[x_1\ x_2\ x_3 \ldots x_n] =$$

$$P[x_1 \mid x_2\ x_3 \ldots x_n] \times P[x_2\ x_3 \ldots x_n]$$

ie probability that events A B C D all occur is

$$P[ABCD] = P[C \mid ABD] \times P[ABD]$$

$$= P[C \mid ABD]\ P[D \mid AB] \times P[AB]$$

$$= P[C \mid ABD] \times P[D \mid A] \times P[B \mid A] \times P[A]$$

ie. $$P(x_1 \ldots x_n) = \prod_{i=1}^{n} P(x_i \mid x_{i-1} \ldots x_n)$$

$$= \prod_{i=1}^{n} P(x_i \mid \text{Parents } x_i)$$

Probability that John calls, mary calls, alarm goes off & both burglary & earthquake occur?

$$P(J, M, A, B, E) = P(J \mid MABE)\, P(MABE)$$

$$= P(J \mid A)\, P(M \mid ABE)\, P(ABE)$$

Since A is only parent of J

$$= P(J \mid A)\, P(M \mid A)\, P(A \mid BE)\, P(B \mid E)\, P(E)$$

$P(B \mid E) = P(B)$
Independant

$$= P(J \mid A)\, P(M \mid A)\, P(A \mid BE)\, P(B)\, P(E)$$

ie: Probability of John calls when alarm goes off
Probability of Mary calls when alarm goes off
Alarm goes of when B & E
Bulgary occurs
Earthquake occurs

(Q1) what is the probability that alarm has sounded but neither a bulgar nor earthquake has occured & both john & merry call ?

→ $P(J, M, A, \sim B, \sim E)$

$= P(J|A)\,P(M|A)\,P(A|\sim E \sim B)\,P(\sim B)\,P(\sim E)$

$= (0.90)(0.70)(0.001)(1-0.001)(1-0.002)$

$= 0.0006281$

Note →

$P(A,B)$ also represented as $P(A \wedge B) / P(A \cap B)$

It actually is the intersection of probabilities of A & B

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

This is derived from Bayes theorem, hence the name Bayesian belief Network

$$P(A|B) = \frac{P(A \cap B)}{P(B)}$$

Q2 What is the probability that John calls?

$$P(J) = P(J, A) + P(J, \sim A)$$

$$= P(J|A)\,P(A) + P(J|\sim A)\,P(\sim A)$$

$$\begin{aligned} P(A) = \; & P(A|EB)\,P(E)\,P(B) \\ & + P(A|\sim E\,B)\,P(\sim E)\,P(B) \\ & + P(A|E\sim B)\,P(E)\,P(\sim B) \\ & + P(A|\sim E\sim B)\,P(\sim E)\,P(\sim B) \end{aligned}$$

$$= 0.002516$$

$$P(\sim A) = 1 - P(A) = 1 - 0.00256$$
$$= 0.9974$$

$$P(J) = 0.9 \times (0.002516) + 0.05 \times (0.9974)$$
$$= 0.05213$$

Applications of Baysian Networks

1. Spam filtering
2. Disease diagnosis (Blood tests)
3. Turbo codes → error correcting codes for WiFi
4. Risk assessment → complex insurance
5. Fraud detection
6. POS tagging in NLP

Advantages →

1. Probabilistic reasoning (Instead of classification like SVM, BBN gives probability that john calls and not say john will call)

2. Transparency (More interpretable)

3. Modularity (Modular design where different parts of the network can be developed independantly and then combined)

4. Handle missing data

Disadvantages

1. Independance assumption may not hold true (what if bulgar doesnt go when earthquake occurs? Those events may not be independant)

2. Knowledge - BBN needs to be trained on significant amount of data or by experts This is subjective and difficult process