Artificial & Computational Intelligence

AIMLCZG557

Contributors & Designers of document content : Cluster Course Faculty Team

M5 : Probabilistic Representation and Reasoning

Presented by
Faculty Name
BITS Email ID

BITS Pilani
Pilani Campus

Artificial and Computational Intelligence

Disclaimer and Acknowledgement

- Few content for these slides may have been obtained from prescribed books and various other source on the Internet
- I hereby acknowledge all the contributors for their material and inputs and gratefully acknowledge people others who made their course materials freely available online.
- .I have provided source information wherever necessary
- This is not a full fledged reading materials. Students are requested to refer to the textbook w.r.t detailed content of the presentation deck that is expected to be shared over e-learning portal - taxilla.
- I have added and modified the content to suit the requirements of the class dynamics & live session's lecture delivery flow for presentation

- **Slide Source / Preparation / Review:**
- From BITS Pilani WILP**:** Prof.Raja vadhana, Prof. Indumathi, Prof.Sangeetha
- From BITS Oncampus & External : Mr.Santosh GSK

BITS Pilani, Pilani Campus

Course Plan

M1 Introduction to AI

M2 Problem Solving Agent using Search

M3 Game Playing

M4 Knowledge Representation using Logics

M5 Probabilistic Representation and Reasoning

M6 Reasoning over time

M7 Ethics in AI

Module 5:

Probabilistic Representation and Reasoning

A. Inference using full joint distribution

B. Bayesian Networks

I. Knowledge Representation

II. Conditional Independence

III. Exact Inference

IV. Introduction to Approximate Inference

Reasoning

- Monotonic Reasoning
- Non- Monotonic Reasoning

Monotonic	Non-Monotonic
Consistent	Relaxed Consistency
Complete Knowledge	Incomplete Knowledge
Static	Dynamic
Discrete	Continuous & Learning Agent
Predicate Logic	Probabilistic Model

Reasoning

- Monotonic Reasoning

- Non- Monotonic Reasoning

  Dependency Directed Backtracking: when a statement is deleted as “ no more valid”, other related statements have to be backtracked and they should be either deleted or new proofs have to be found for them. This is called dependency directed backtracking (DDB)

Uncertainty

You can reach Bangalore Airport from MG Road within 90 mins if you go by route A.

- There is uncertainty in this information due to partial observability and non determinism
- Agents should handle such uncertainty

Previous approaches like Logic represent all possible world states

Such approaches can’t be used as multiple possible states need to be enumerated to handle the uncertainty in our information

Uncertainty

You can reach Bangalore Airport from MG Road within 90 mins if you go by route A.

Road Block	Festival Season	Weekend	Observation (20)	Prob
F	F	F	12	0.6
F	F	T	3	0.15
F	T	F	2	0.1
F	T	T	2	0.1
T	F	F	0	0
T	F	T	0	0
T	T	F	1	0.05
T	T	T	0	0
				=1

Probability Theory

Basics

Conditional Probability

Chain Rule

Independence

Conditional Independence

Belief Nets

Joint Probability distribution

Probability Basics – Refresher Self Study

Sample Space: Set of all possible outcomes.

- Ex: After tossing 2 coins, the set of all possible outcomes are
- {HH, HT, TH, TT}

Event: A subset of a sample space.

- An event of interest might be - {HH}

Probability Basics - Model

A fully specified **probability model** associates a numerical probability P(ω) with each possible world.

The **basic axioms**

1. Every possible world has a probability between 0 and 1
2. Sum of probabilities of possible worlds is 1 P(True) = 1
   P(False) = 0

3. P(a ∨ b) = P(a) + P(b) - P(a ∧ b)

E.g., P(HH) = 0.25; P(HT) = 0.25; P(TT) = 0.25, P(TH) = 0.25

$$0 \leq P(\omega) \leq 1 \text{ for every } \omega \text{ and } \sum_{\omega \in \Omega} P(\omega) = 1$$

Probability Basics – Exclusive / Exhaustive events

Mutually Exclusive Events:

- Two events (or propositions) are mutually exclusive or disjoint if they cannot both occur at the same time(be true).
- A clear example is the set of outcomes of a single coin toss, which can result in either heads or tails, but not both.

Exhaustive Events:

- A set of events is jointly or collectively exhaustive if at least one of the events must occur.
- E.g., when rolling a six-sided die, the events 1, 2, 3, 4, 5, and 6 are collectively exhaustive.

Probability Basics - Propositions

Probabilistic assertions (Propositions)

- Usually not a particular world event but about a set of them
- E.g., two dice when rolled, a proposition φ can be “the sum of dice is 11”

For any proposition φ,

$P(\varphi) = P(\text{sum} = 11) \quad = P((5, 6)) + P((6, 5))$

$= 1/36 + 1/36 = 1/18$

Probability Basics – Unconditional/Prior

Unconditional / Prior probabilities:

Propositions like P(sum = 11) or P(two dices rolling equals) are called Unconditional or Prior probabilities

They refer to degree of belief in absence of any other information

$$P(a \mid b) = \frac{P(a \wedge b)}{P(b)}$$

$$P(a \wedge b) = P(a \mid b) P(b)$$

Probability Basics - Conditional

However, most of the time we have some information, we call it **evidence**

E.g., we can interested in two dice rolling a double (i.e., 1,1 or 2,2, etc)
When one die has rolled 5 and the other die is still spinning
Here, we not interested in unconditional probability of rolling a double
Instead, the **conditional** or **posterior** probability for rolling a double given the first die has rolled a 5
$P(doubles \mid Die_1 = 5)$ where | is pronounced “given”

E.g., if you are going for a dentist for a checkup, P(cavity) = 0.2

- If you have a toothache, then P(cavity | toothache) = 0.6

Independence

If we have two random variables, TimeToBnlrAirport and HyderabadWeather

P(TimeToBnlrAirport, HyderabadWeather)

To determine their relation, use the product rule

= P(TimeToBnlrAirport | HyderabadWeather) / P(HyderabadWeather)

However, we would argue that HyderabadWeather and TimeToBnlrAirport doesn't have any relation and hence

P(TimeToBnlrAirport | HyderabadWeather) = P(TimeToBnlrAirport)

This is called Independence or Marginal Independence

Independence between propositions a and b can be written as

$$P(a \mid b) = P(a) \quad \text{or} \quad P(b \mid a) = P(b) \quad \text{or} \quad P(a \wedge b) = P(a)P(b)$$

Bayes Rule

Using the product rule for propositions a and b

$$P(a \wedge b) = P(a \mid b)P(b) \quad \text{and} \quad P(a \wedge b) = P(b \mid a)P(a)$$

Equating the right hand sides and dividing by P(a)

$$P(b \mid a) = \frac{P(a \mid b)P(b)}{P(a)}$$

This is called the Bayes Rule

Conditional Independence

2 random variables A and B are conditionally independent given C iff

$P(a, b \mid c) = P(a \mid c)\, P(b \mid c)$ for all values a, b, c

More intuitive (equivalent) conditional formulation

- A and B are conditionally independent given C iff

  $P(a \mid b, c) = P(a \mid c)$ OR $P(b \mid a, c) = P(b \mid c)$, for all values a, b, c
- Intuitive interpretation:

  $P(a \mid b, c) = P(a \mid c)$ tells us that learning about b, given that we already know c, provides no change in our probability for a, i.e., b contains no information about a beyond what c provides

  $P(R \mid F, P) = P(R \mid P)$

Joint Probability Distributions

Instead of distribution over single variable, we can model distribution over multiple variables, separated by comma

E.g., **P**(A, B) = **P**(A | B) . **P**(B)

P(A, B) is the probability distribution over combination of all values of A and B

E.g., if A = Weather and B = Cavity

$$
\begin{aligned}
P(W=sunny \wedge C=true) &= P(W=sunny|C=true)\, P(C=true)\\
P(W=rain \wedge C=true) &= P(W=rain|C=true)\, P(C=true)\\
P(W=cloudy \wedge C=true) &= P(W=cloudy|C=true)\, P(C=true)\\
P(W=snow \wedge C=true) &= P(W=snow|C=true)\, P(C=true)\\
P(W=sunny \wedge C=false) &= P(W=sunny|C=false)\, P(C=false)\\
P(W=rain \wedge C=false) &= P(W=rain|C=false)\, P(C=false)\\
P(W=cloudy \wedge C=false) &= P(W=cloudy|C=false)\, P(C=false)\\
P(W=snow \wedge C=false) &= P(W=snow|C=false)\, P(C=false)\,.
\end{aligned}
$$

Probabilistic Inference

Computation of posterior probabilities given observed evidence, i.e., full joint probability distribution

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

Query: P(cavity ∨ **toothache)**

0.108 + 0.012 + 0.072 + 0.008 + 0.016 + 0.064 = 0.28

Conditional Probability

Towards Chain Rule:

$$P(a \mid b) = P(a,b) / P(b)$$

$P(a, b) = P(a \mid b)\, P(b)$

$P(a, b, c) = P(a, x)$ where $x = b,c$

$$\begin{aligned} P(a,x) &= P(a \mid x) \,.\, P(x) \\ &= P(a \mid bc) \,.\, P(b, c) \\ &= P(a \mid bc) \,.\, P(b \mid c) \,.\, P(c) \end{aligned}$$

Hence : $P(a,b,c) = P(a \mid bc) \,.\, P(b \mid c) \,.\, P(c)$

Chain Rule : Generalization

$$P(X_1, X_2, \ldots X_K) = \prod P(X_i \mid X_{i-1}, \ldots X_1)$$

Where i = k to 1 (reverse)

Probability Theory

<ins>Independence</ins>

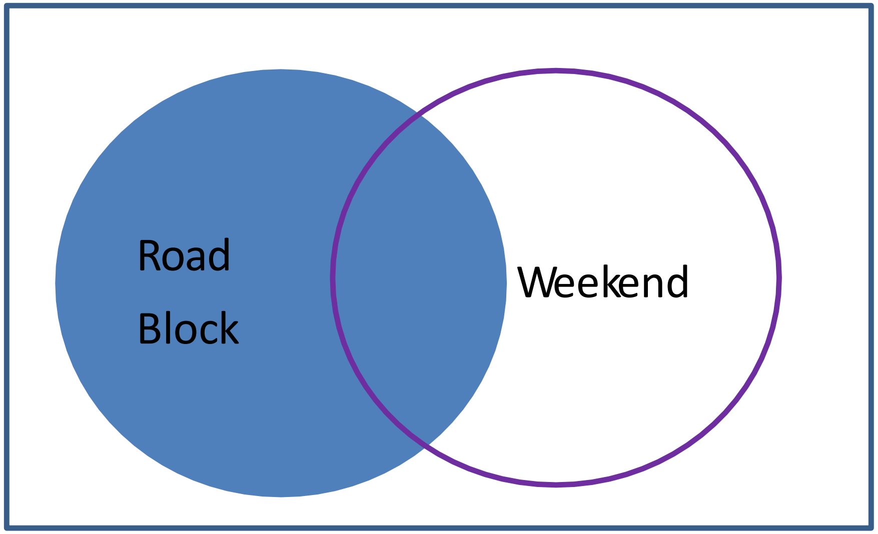
P(a | b) = P(a)

Implication:

P(a | b) = P(a,b) / P(b)

P(a) = P(a,b) / P(b)

P(a,b) = P(a) . P(b)



<ins>Conditional Independence</ins>

P(a | b c) = P(a | c)

Probability Theory

Conditional Independence

P(a | b c) = P(a | c)

Extension:

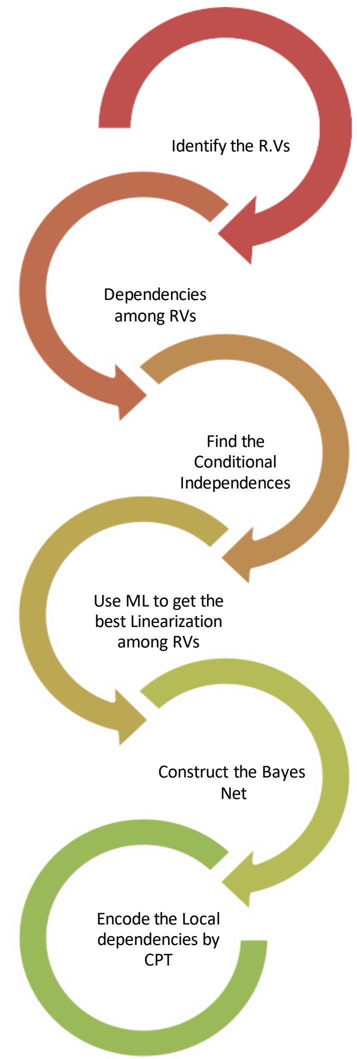
P(a b | c) = P(a | c) . P(b | c)

Building a Bayesian Network

Example Bayesian Net #1

A simple world with four random variables

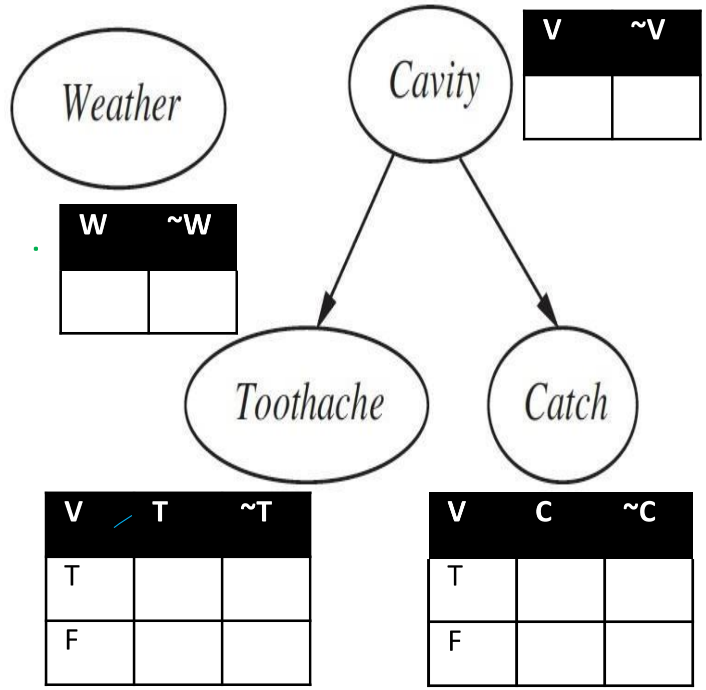
- Weather, Toothache, Cavity, Catch
- Weather is independent of other variables
- Toothache and Catch are conditionally independent given Cavity
- P(Toothache, Catch | Cavity) = P(Toothache | Cavity) . P(Catch | Cavity)
- Cavity is a direct cause of Toothache and Catch
- No direct relation between Toothache and Catch exists

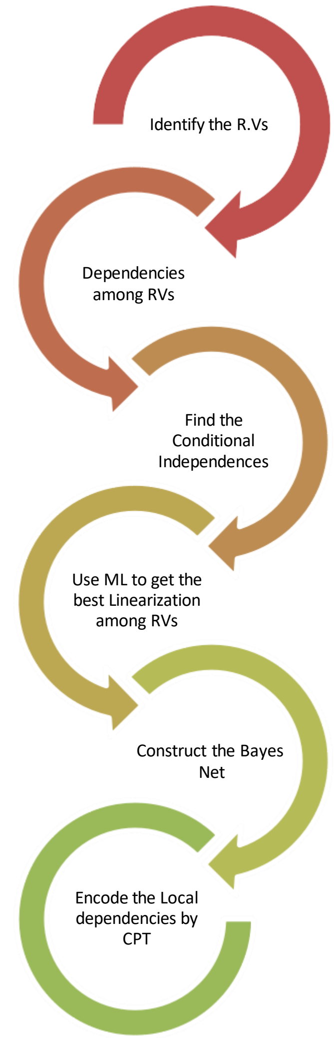


Example Bayesian Net #1

A simple world with four random variables

- Weather, Toothache, Cavity, Catch
- Weather is independent of other variables
- Toothache and Catch are conditionally independent given Cavity
- P(Toothache, Catch | Cavity) = P(Toothache | Cavity) . P(Catch | Cavity)
- Cavity is a direct cause of Toothache and Catch
- No direct relation between Toothache and Catch exists

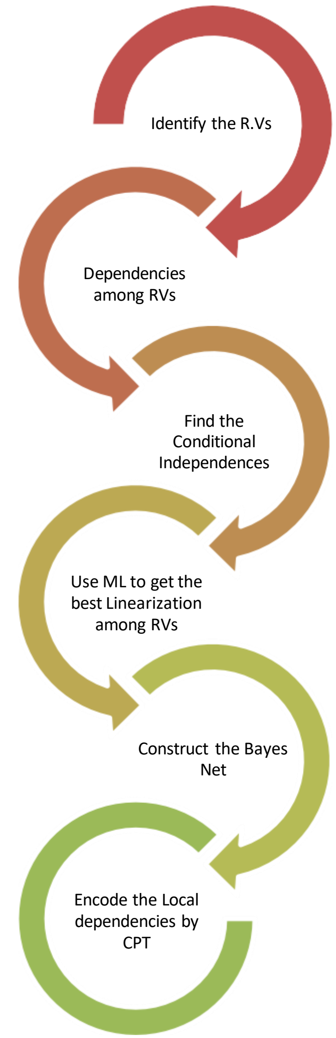




Example Bayesian Net #2

A Burglary Alarm System

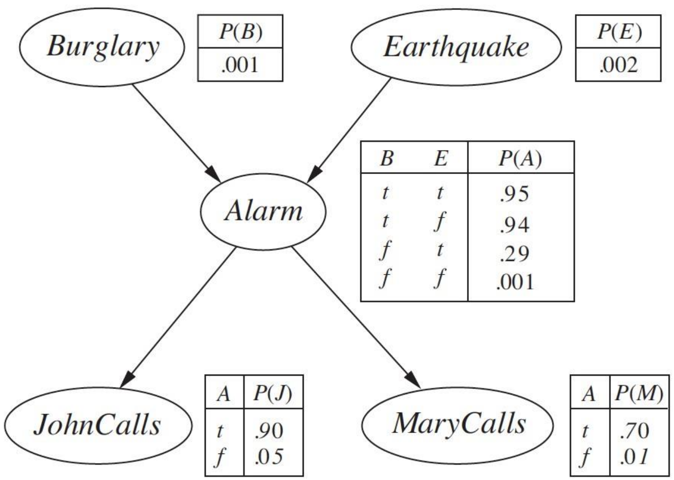
- Fairly reliable on detecting a burglary
- Also responds to earthquakes
- Two neighbors John and Mary are asked to call you at work when Burglary happens and they hear the Alarm
- John nearly always calls when he hears the alarm, however sometimes confuses the telephone ring with alarm and calls then too
- Mary like loud music and often misses the alarm altogether
- **Problem**: Given the information that who has / has not called we need to estimate the probability of a burglary

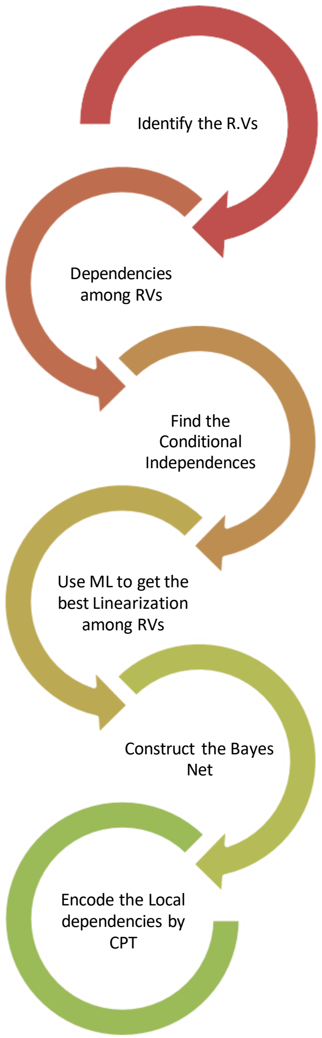


Example Bayesian Net #2

A Burglary Alarm System

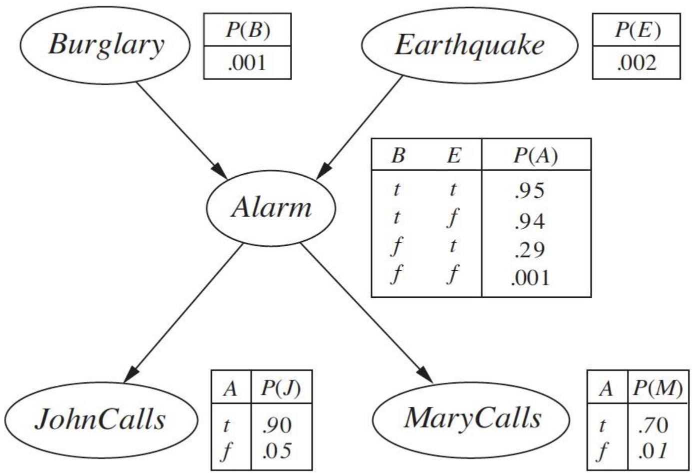
- Fairly reliable on detecting a burglary
- Also responds to earthquakes
- Two neighbors John and Mary ar asked to call you at work when Burglary happens and they hear the Alarm
- John nearly always calls when he hears the alarm, however sometimes confuses the telephone ring with alarm and calls then too
- Mary like loud music and often misses the alarm altogether
- **Problem**: Given the information that who has / has not called we need to estimate the probability of a burglary





Examples

1. Calculate the probability that alarm has sounded, but neither burglary nor earthquake happened, and both John and Mary called



$$
\begin{aligned}
P(j, m, a, \neg b, \neg e) &= P(j \mid a) P(m \mid a) P(a \mid \neg b \wedge \neg e) P(\neg b) P(\neg e) \\
&= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998 = 0.000628
\end{aligned}
$$

Example Bayesian Net #3

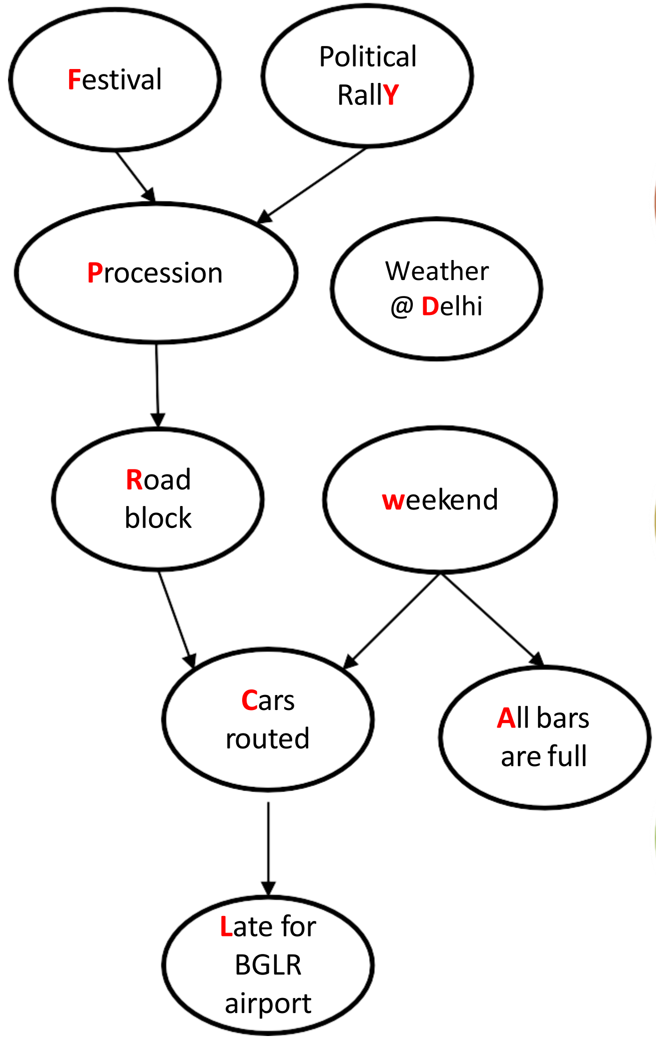
Traffic Prediction -Travel Estimation

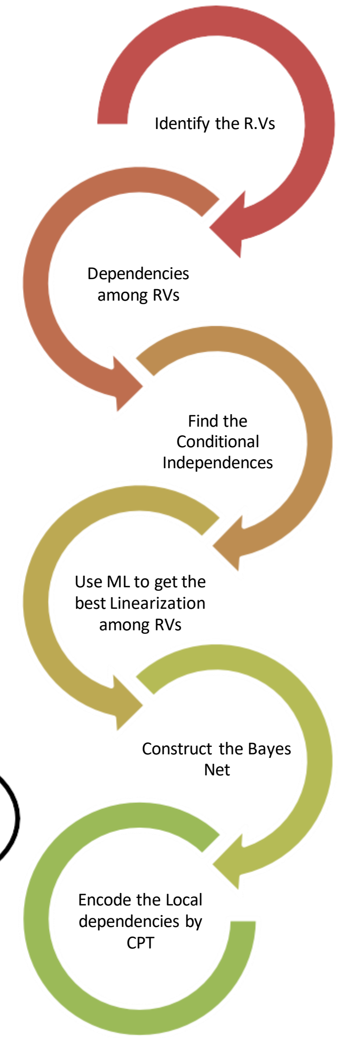
- AI system reminds traveler regarding start time
- Travel plan is to reach Delhi and the weather of Delhi may influence the accommodation plans
- Traveler always take car to reach airport
- Car may be rerouted either due to road block or weekday traffic during working hours which delays the arrival to airport
- Bars are always observed to be full on weekends
- Authorities block roads to safe the processions
- Processions observed during festive season or due to the political rally.
- **Problem**: Given the information that there is a political rally expected estimate the probability of late arrival

Example Bayesian Net #3

Traffic Prediction -Travel Estimation

- AI system reminds traveler regarding start time
- Travel plan is to reach Delhi and the weather of Delhi may influence the accommodation plans
- Traveler always take car to reach airport
- Car may be rerouted either due to road block or weekday traffic during working hours which delays the arrival to airport
- Bars are always observed to be full on weekends
- Authorities block roads to safe the processions
- Processions observed during festive season or due to the political rally.
- **Problem**: Given the information that there is a political rally expected estimate the probability of late arrival

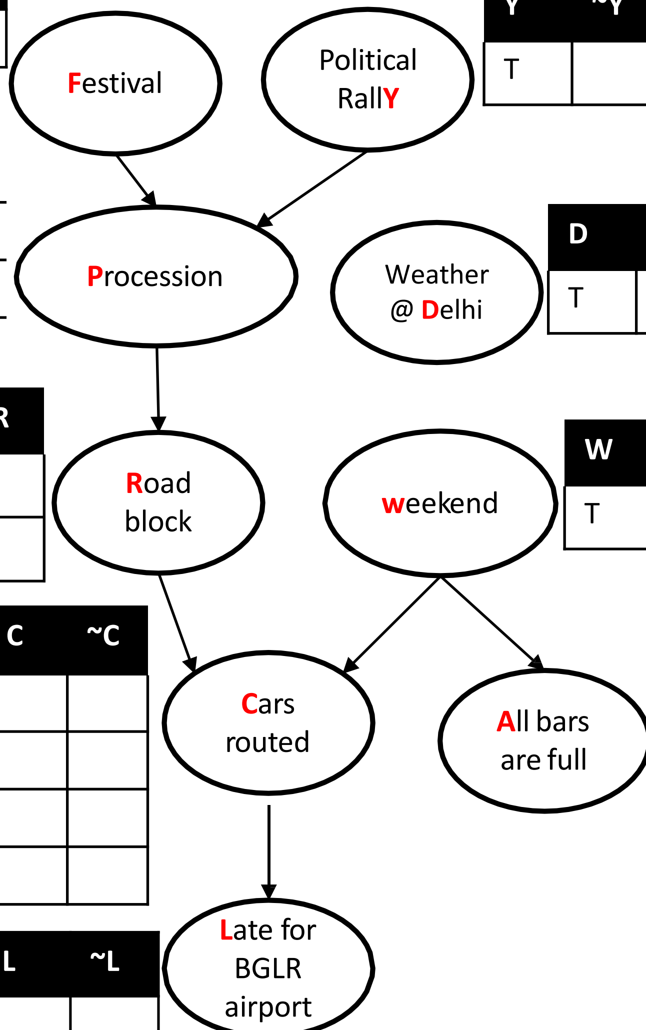




Example Bayesian Net #3

F	~F
T	

T	T		
T	F		
F	T		
F	F		



Y	~Y
T	

D	~D
T	

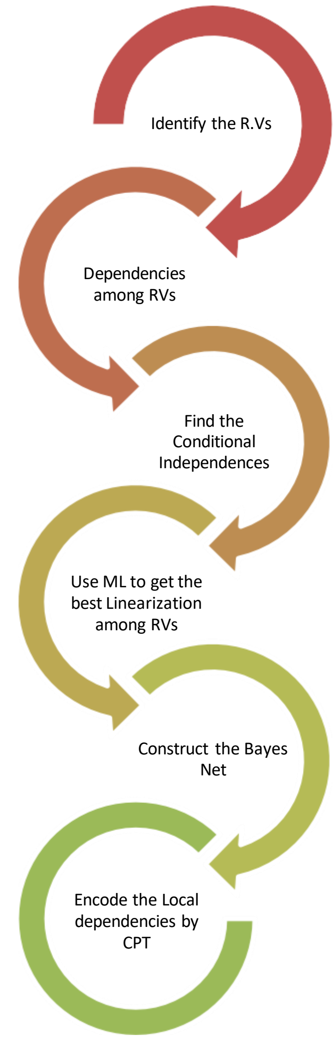
P	R	~R
T		
F		

W	~W
T	

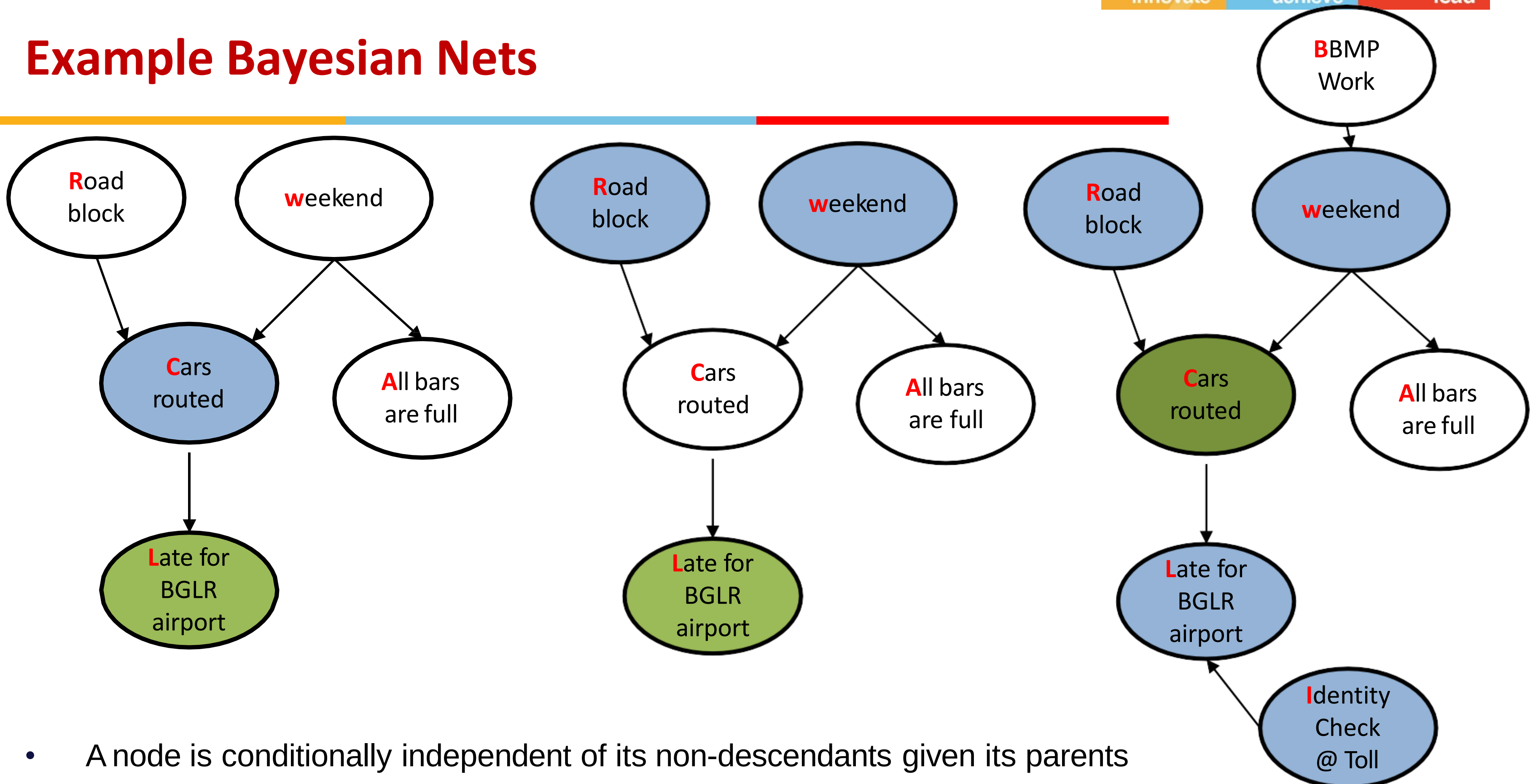
R	W	C	~C
T	T		
T	F		
F	T		
F	F		

W	A	~A
T		
F		

C	L	~L
T		
F		

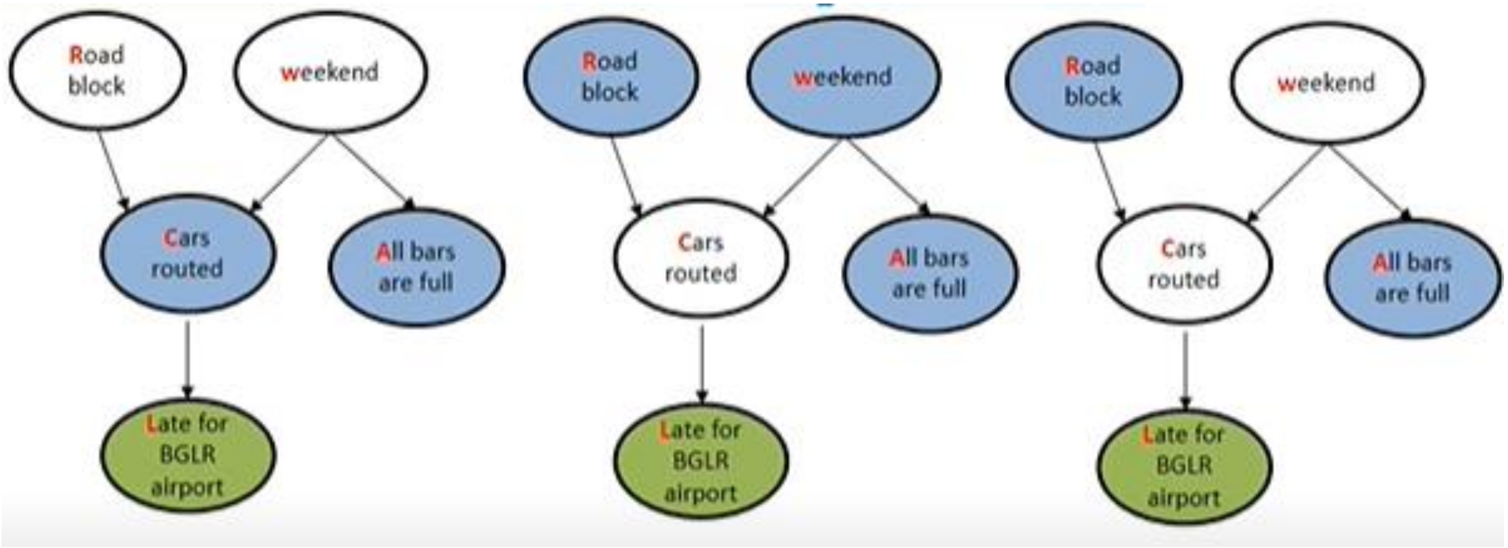


Example Bayesian Nets



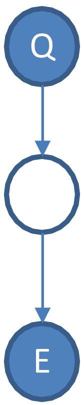
- A node is conditionally independent of its non-descendants given its parents
- A node is conditionally independent of all other nodes in the net , given its parents, children and children’s parents.

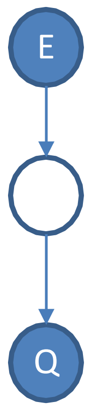
Example Bayesian Nets

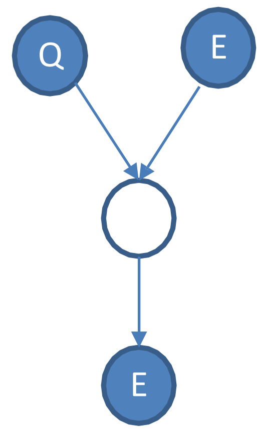


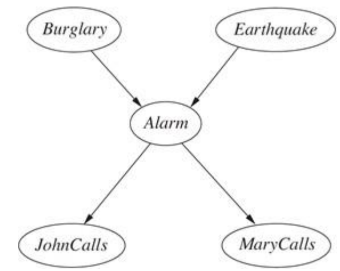
Belief Nets

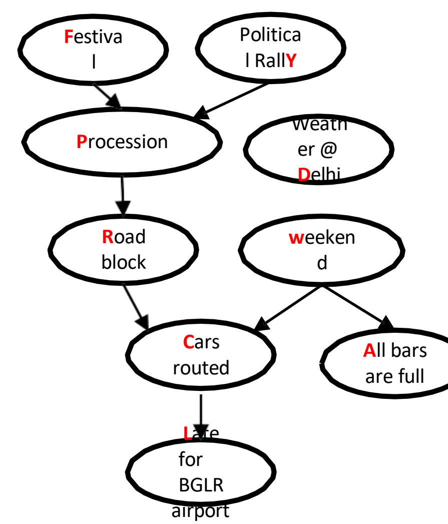
Diagnostic **Causal** **Inter-Casual** **Mixed Inferences**



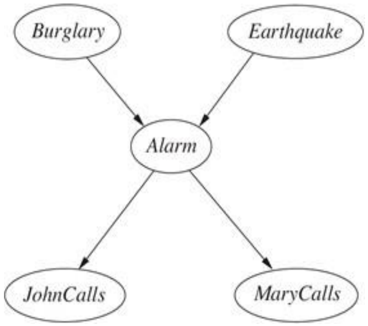




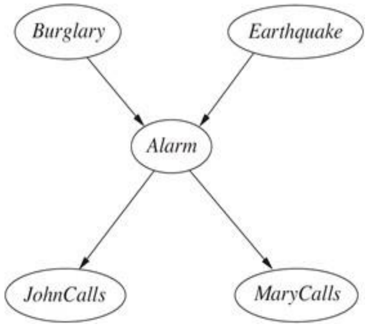




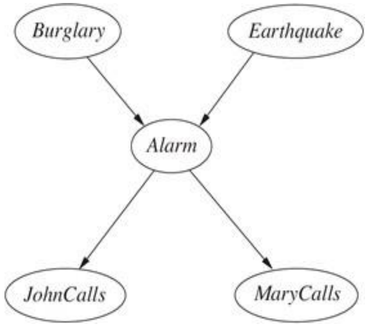
Belief Nets



Burglary
Earthquake
Alarm
JohnCalls
MaryCalls



Burglary
Earthquake
Alarm
JohnCalls
MaryCalls

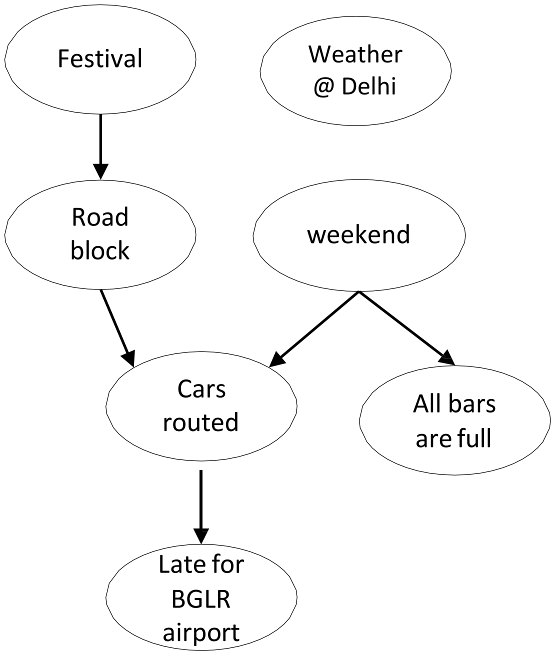


Burglary
Earthquake
Alarm
JohnCalls
MaryCalls

Inferences in Bayesian Nets

Enumeration

Examples

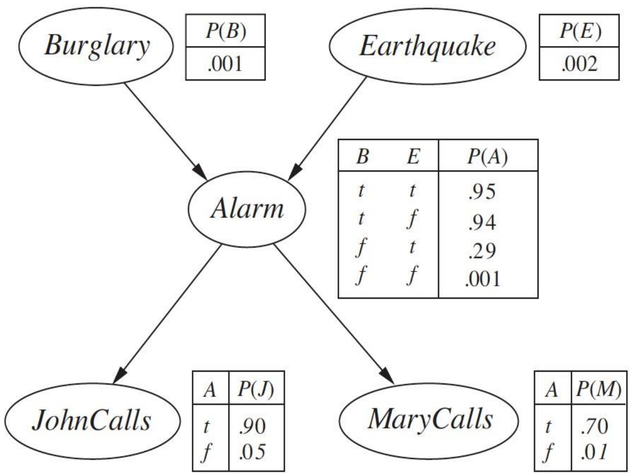


1. Calculate the probability that arrival at airport was delayed during a weekend but there was no road block or festival and car was not routed anywhere.

2. What is the probability that it is a festival season given cars where routed?

3. What is the probability that car arrived late at airport given it's a festival day?

2. What is the probability that Burglary happened given John & Mary called the police



$$P(B \mid J,M) = \frac{P(B, J, M)}{P(J, M)}$$

$$P(B \mid J,M) = \frac{\sum_{A, E} P(J, M, A, B, E)}{\sum_{A, B, E} P(J, M, A, B, E)}$$

$$P(B \mid JM) + P(\neg B \mid JM) = 1$$

$$\frac{P(BJM)}{P(JM)} + \frac{P(\neg BJM)}{P(JM)} = 1$$

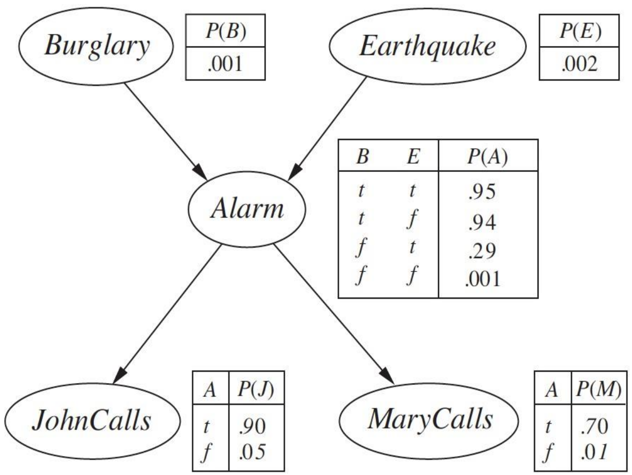
$$\frac{1}{P(JM)}\left[P(BJM) + P(\neg BJM)\right] = 1$$

let $\alpha = \frac{1}{P(JM)}$

$$\alpha = \frac{1}{P(BJM) + P(\neg BJM)} \rightarrow 1$$

innovate achieve lead

2. What is the probability that Burglary happened given John & Mary called the police



$$P(B \mid J,M) = \frac{P(B, J, M)}{P(J, M)}$$

$$P(B \mid J,M) = \frac{\sum_{A,E} P(J, M, A, B, E)}{\sum_{A,B,E} P(J, M, A, B, E)}$$

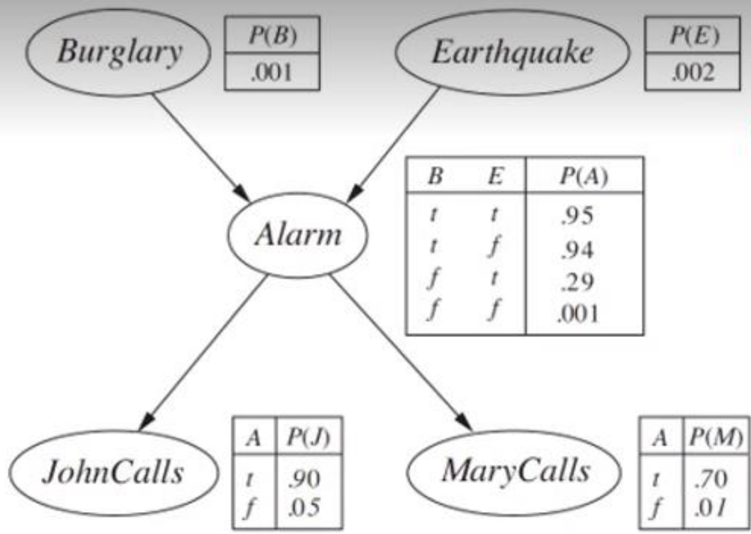
$$P(B \mid JM) + P(\neg B \mid JM) = 1$$

$$\frac{P(BJM)}{P(JM)} + \frac{P(\neg BJM)}{P(JM)} = 1$$

$$\frac{1}{P(JM)}\left[P(BJM) + P(\neg BJM)\right] = 1$$

let $\alpha = \frac{1}{P(JM)}$

$$\alpha = \frac{1}{P(BJM) + P(\neg BJM)} \rightarrow 1$$



P(B)
.001

P(E)
.002

B	E	P(A)
t	t	.95
t	f	.94
f	t	.29
f	f	.001

A	P(J)
t	.90
f	.05

A	P(M)
t	.70
f	.01

2. What is the probability that Burglary happened given John & Mary called the police

$$P(B \mid J,M) = \frac{P(B, J, M)}{P(J, M)}$$

$$P(B \mid J,M) = \frac{\sum_{A,E} P(J, M, A, B, E)}{\sum_{A,B,E} P(J, M, A, B, E)}$$

$$\boxed{P(B\, J\, M)} = \sum_{A,E} P(J, M, A, B, E)$$

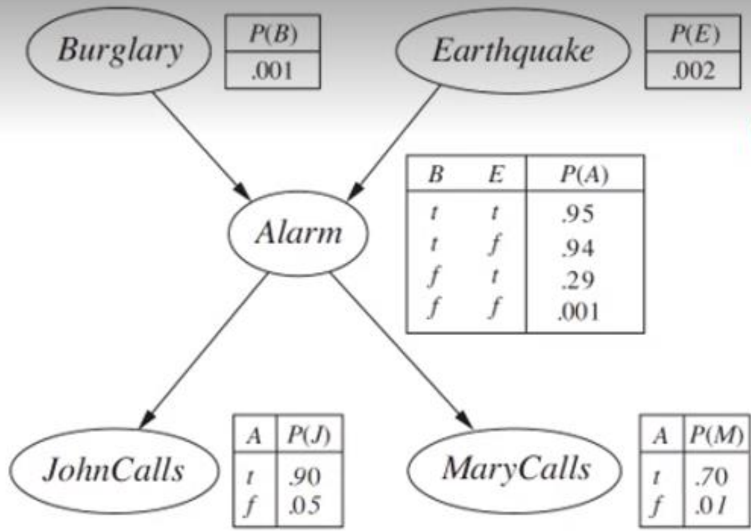
$$= \sum_{AE} P(J \mid MABE) \cdot P(M \mid ABE) \cdot P(A \mid BE) \cdot P(B \mid E) \cdot P(E)$$

$$= \sum_{AE} P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid BE) \cdot P(B) \cdot P(E)$$

$$= \sum_{A} \sum_{E} P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid BE) \cdot P(B) \cdot P(E)$$

$$= \sum_{A} \{P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid BE) \cdot P(B) \cdot P(E)\} + \{P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid B \neg E) \cdot P(B) \cdot P(\neg E)\}$$

$$= [P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid BE) \cdot P(B) \cdot P(E)] + [P(J \mid A) \cdot P(M \mid A) \cdot P(A \mid B \neg E) \cdot P(B) \cdot P(\neg E)]$$
$$+ [P(J \mid \neg A) \cdot P(M \mid \neg A) \cdot P(\neg A \mid BE) \cdot P(B) \cdot P(E)] + [P(J \mid \neg A) \cdot P(M \mid \neg A) \cdot P(\neg A \mid B \neg E) \cdot P(B) \cdot P(\neg E)]$$



$$\boxed{P(B|J,M)} = \sum_{A,E} P(J,M,A,B,E)$$

$$= \sum_{A E} P(J|M A B E)\cdot P(M|A B E)\cdot P(A|B E)\cdot P(B|E)\cdot P(E)$$

$$= \sum_{A E} P(J|A)\cdot P(M|A)\cdot P(A|BE)\cdot P(B)\cdot P(E)$$

$$= \sum_{A}\sum_{E} P(J|A)\cdot P(M|A)\cdot P(A|BE)\cdot P(B)\cdot P(E)$$

$$= \sum_{A} \{P(J|A)\cdot P(M|A)\cdot P(A|BE)\cdot P(B)\cdot P(E)\} + \{P(J|A)\cdot P(M|A)\cdot P(A|B\neg E)\cdot P(B)\cdot P(\neg E)\}$$

2. What is the probability that Burglary happened given John & Mary called the police

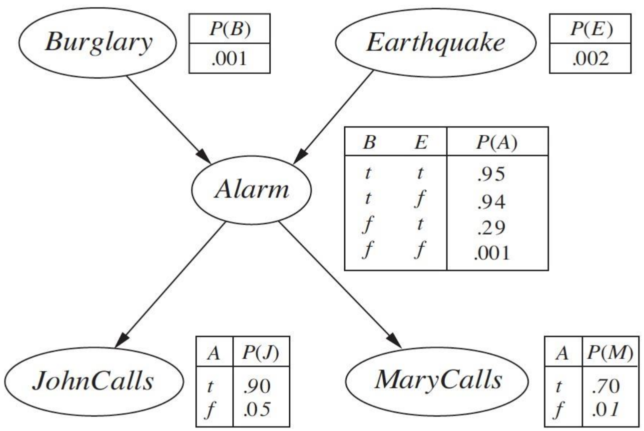
$$P(B \mid J,M) = \frac{P(B, J, M)}{P(J, M)}$$

$$P(B \mid J,M) = \frac{\sum_{A,E} P(J, M, A, B, E)}{\sum_{A,B,E} P(J, M, A, B, E)}$$

$$= [P(J|A)\cdot P(M|A)\cdot P(A|BE)\cdot P(B)\cdot P(E)] + [P(J|A)\cdot P(M|A)\cdot P(A|B\neg E)\cdot P(B)\cdot P(\neg E)]$$
$$+ [P(J|\neg A)\cdot P(M|\neg A)\cdot P(\neg A|BE)\cdot P(B)\cdot P(E)] + [P(J|\neg A)\cdot P(M|\neg A)\cdot P(\neg A|B\neg E)\cdot P(B)\cdot P(\neg E)]$$

Examples

3. What is the probability that John calls given earthquake occurred?



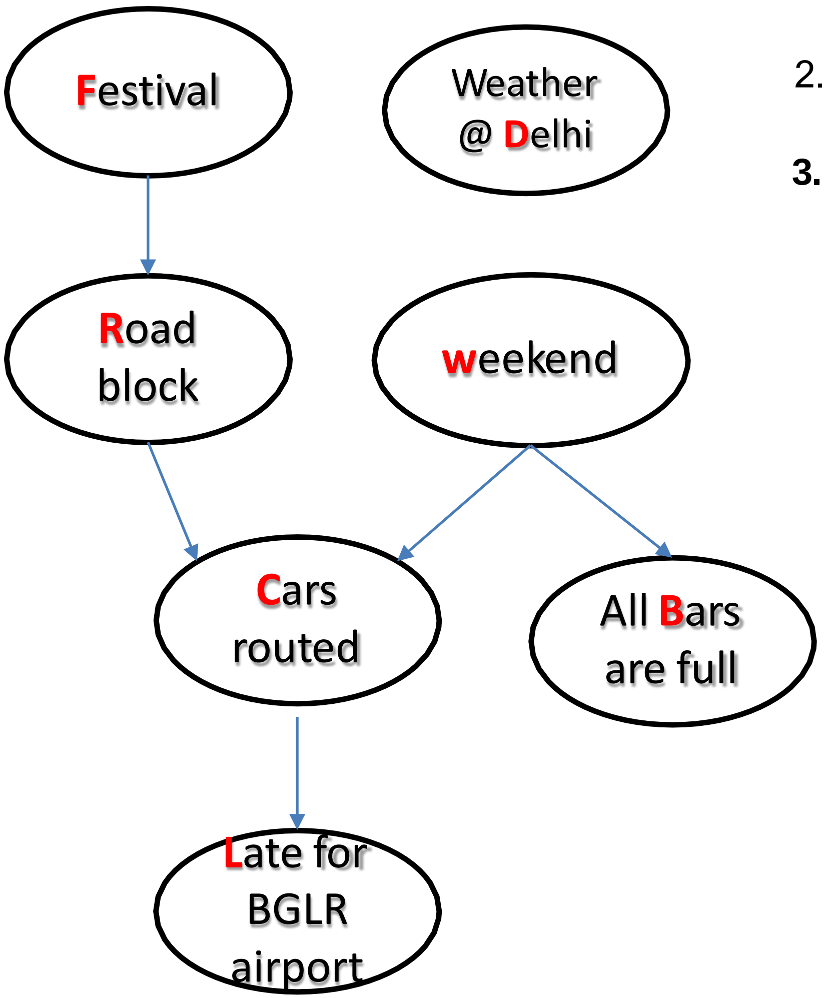
$$P(J \mid E) = \frac{P(J, E)}{P(E)}$$

$$P(J \mid E) = \frac{\sum_{M, A, B} P(J, M, A, B, E)}{\sum_{J, M, A, B} P(J, M, A, B, E)}$$

Inferences in Bayesian Nets

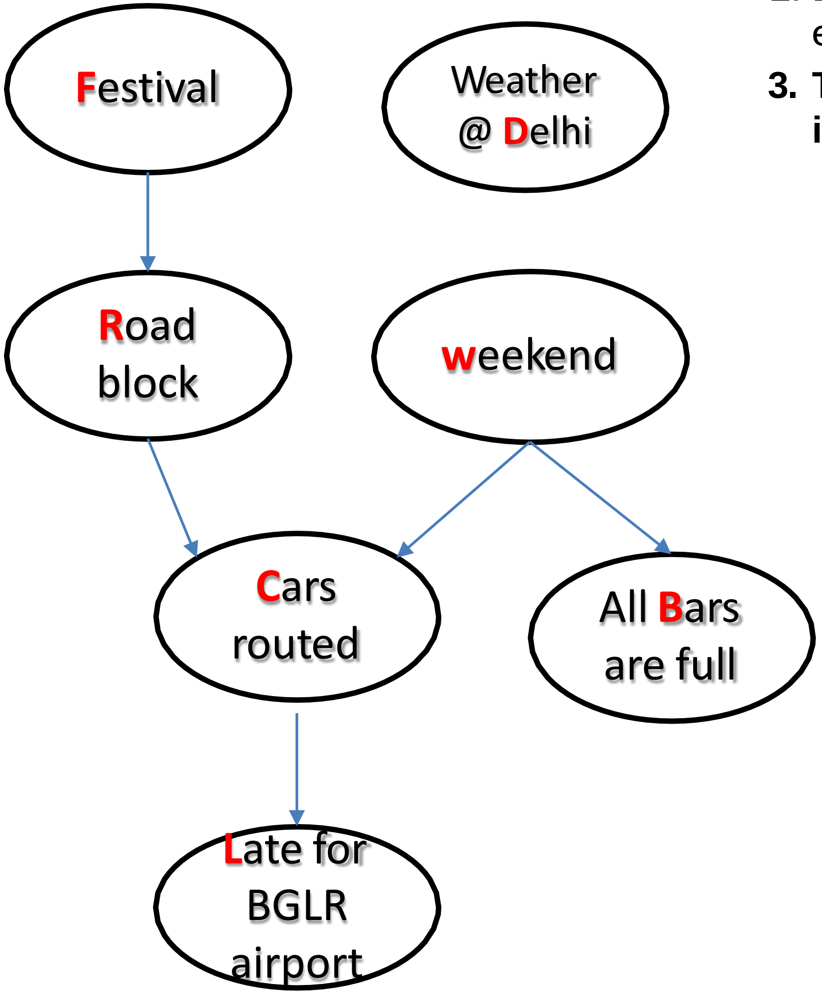
Variable Elimination
Reduce Guaranteed Independent nodes

D-Connectedness Vs D-Separation



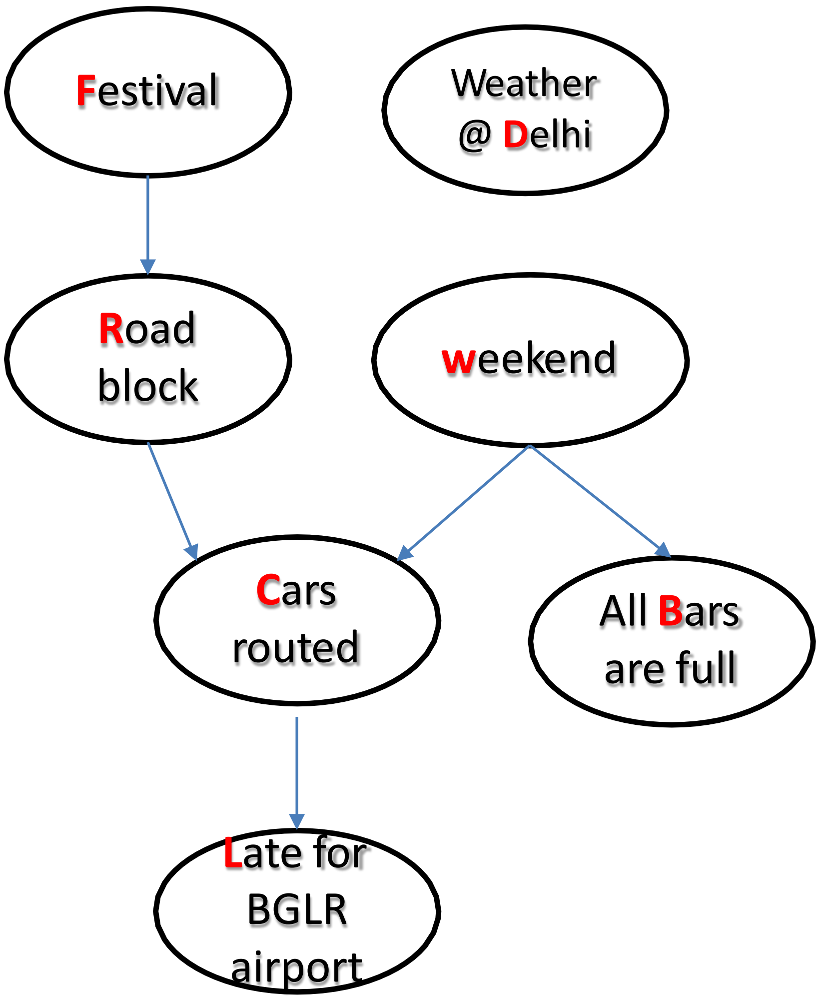
1. Each variable is conditionally independent of its non-descendants, given its parents
2. Eliminate the hidden variables that is neither a query nor an evidence
3. **Two variables are d-separated if they are conditionally independent given evidences**

D-Connectedness Vs D-Separation



1. Each variable is conditionally independent of its non-descendants, given its parents
2. Eliminate the hidden variables that is neither a query nor an evidence
3. **Two variables are d-separated if they are conditionally independent given evidences**

Try it & Test



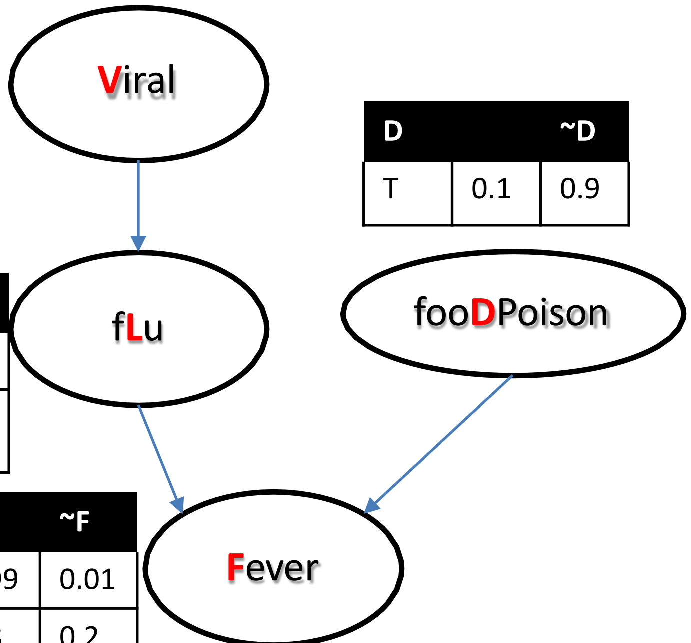
X	Y	Evidence Z	d-sep?
F	W	C	No
L	W	R	No
R	**L**	**C**	**Yes**
B	R	C	No

- P(R | L , C) = P(R| L)

R & L are d-separated ie., conditionally independent given C

D-Separation in Inference

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

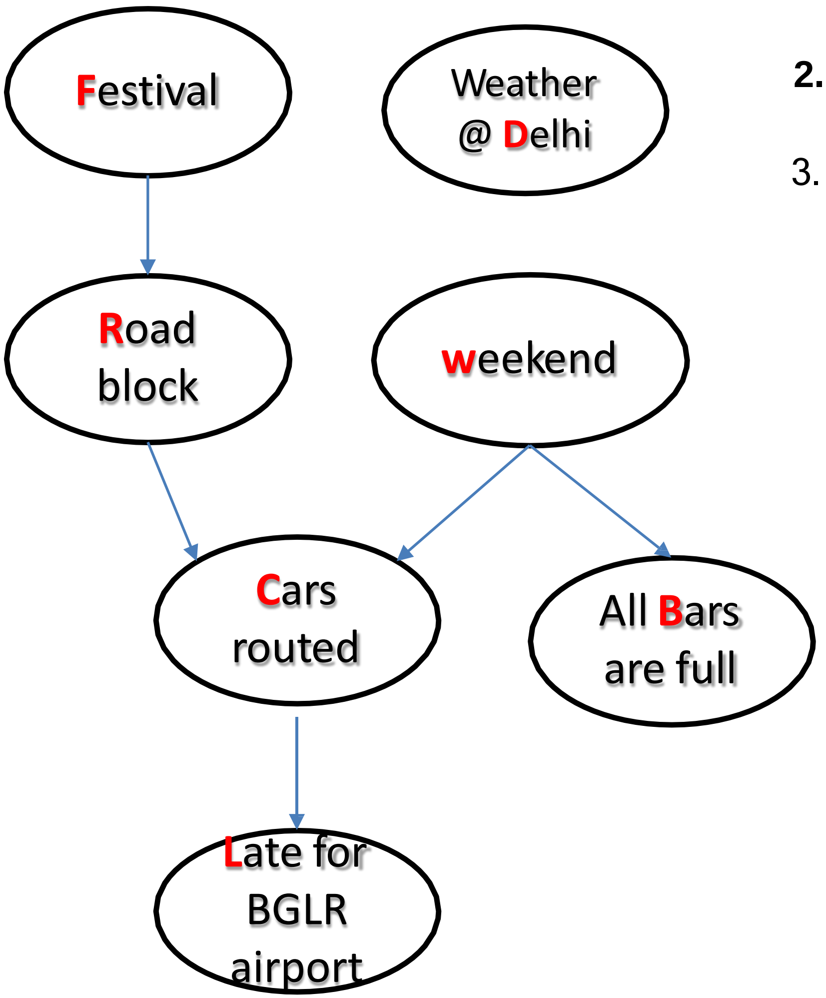
V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

X	Y	Evidence Z	d-sep?
V	F	L	Yes
V	D	L	Yes

- P(V | F , L)
- P(V | D , L)

Variable Elimination



1. Each variable is conditionally independent of its non-descendants, given its parents
2. **Eliminate the hidden variables that is neither a query nor evidence**
3. Two variables are d-separated if they are conditionally independent given evidences

- $P(B) = \sum_{L, B, W, R, F} P(L, C, B, W, R, F)$
  $= \sum_{L} \sum_{B} P(L|C) \cdot P(B \mid W) \cdot \sum_{W} P(C \mid W, R) \cdot \sum_{R} P(R \mid F) \cdot \sum_{F} P(F)$
  $= \mathbf{P(B \mid W)}$

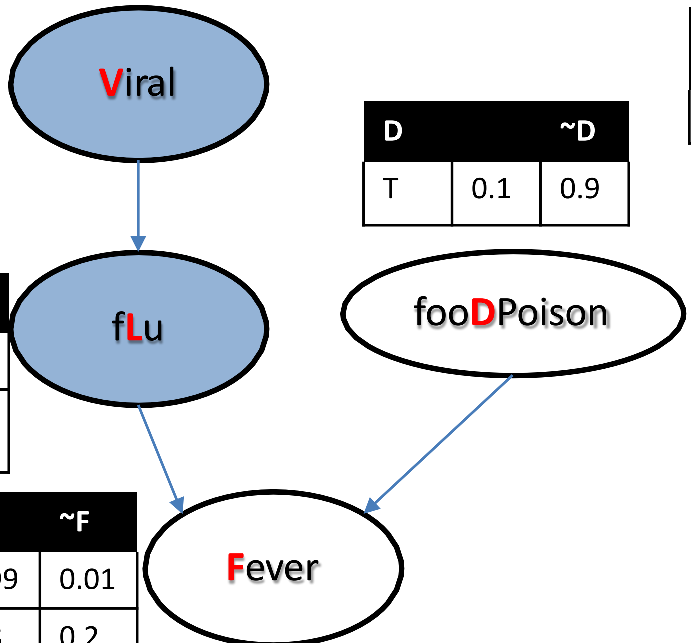
All other variables are hidden w.r.t to B as (L, C, R, F) are neither evidence nor query nor (L, C, R, F) ∊ Ancestors(W, B)

This is variable elimination example targeting irrelevant nodes

Inference

Variable Elimination: V

V	~V	
T	0.6	0.4



D	~D	
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	~V	
T	0.6	0.4

V	L	~L
T	0.8	0.2
F	0.01	0.99

V	L	
T	T	0.48
T	F	0.12
F	T	0.004
F	F	0.396

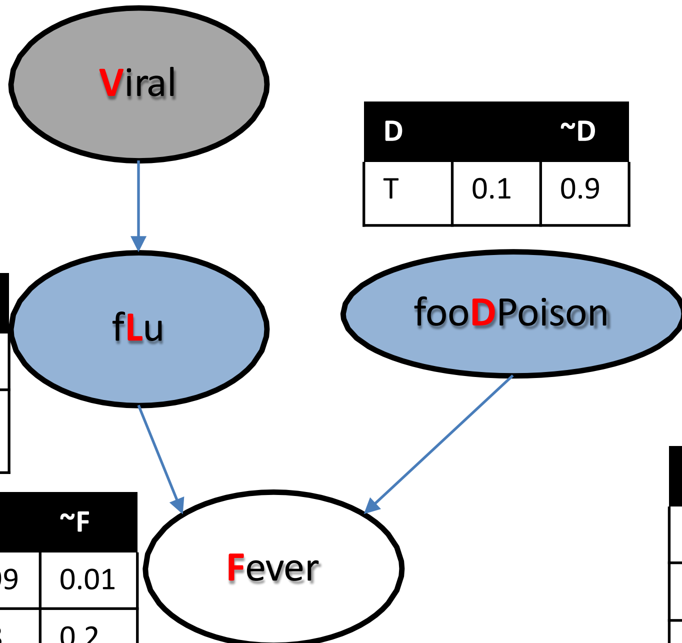
L	
T	0.484
F	0.516

P(V)
P(L|V)
P(D)
P(F|L,D)

Inference

Variable Elimination: L,D

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

P(L)
P(D)
P(F|L,D)

L	
T	0.484
F	0.516

D		~D
T	0.1	0.9

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

D	L	F	
T	T	T	
T	F	T	
F	T	T	
F	F	T	
T	T	F	
T	F	F	
F	T	F	
F	F	F	

L	F	
T	T	0.08928
F	T	0.37174
T	F	0.0108
F	F	0.52818

F	
T	0.46102
F	0.53898

Approximate Inferences in Bayesian Nets
Introduction

Prior Sampling

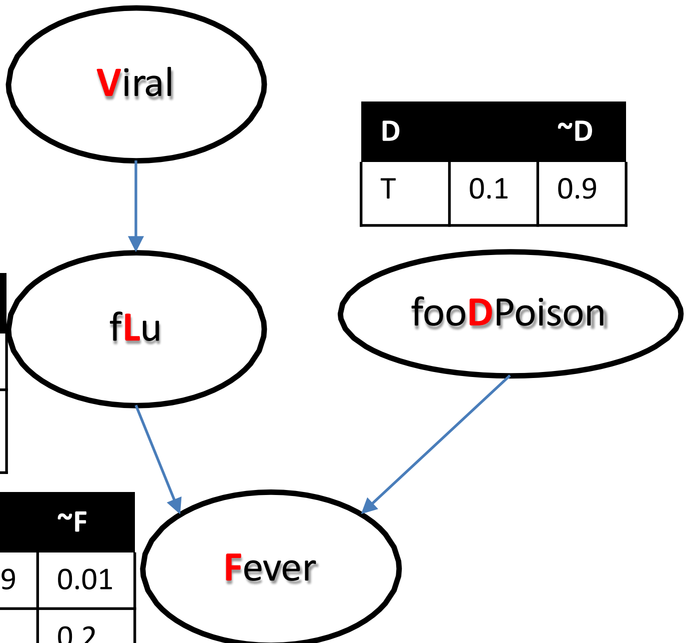
Sample Generation by Randomization

V		~V
T	0.6	0.4

D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95



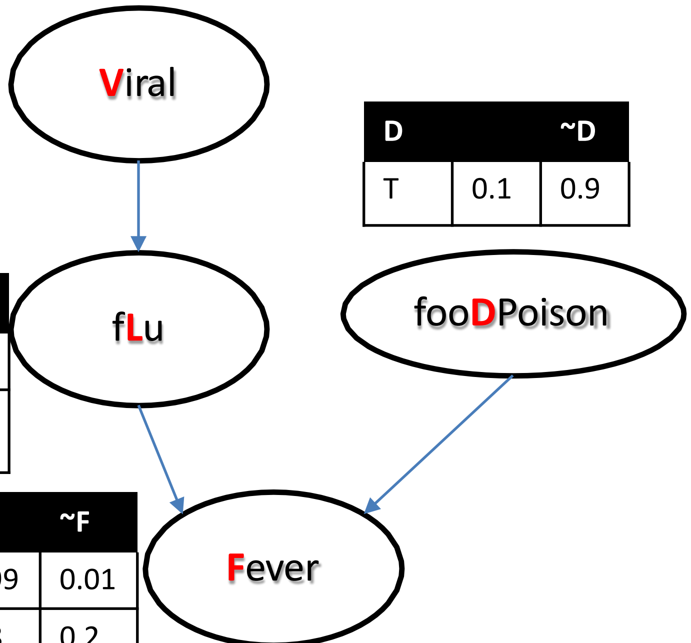
V	L	D	F

0.3, 0.2, 0.6, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0 . 9 , 0 . 5 5 ………. . .

Prior Sampling

Sample Generation by Randomization

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

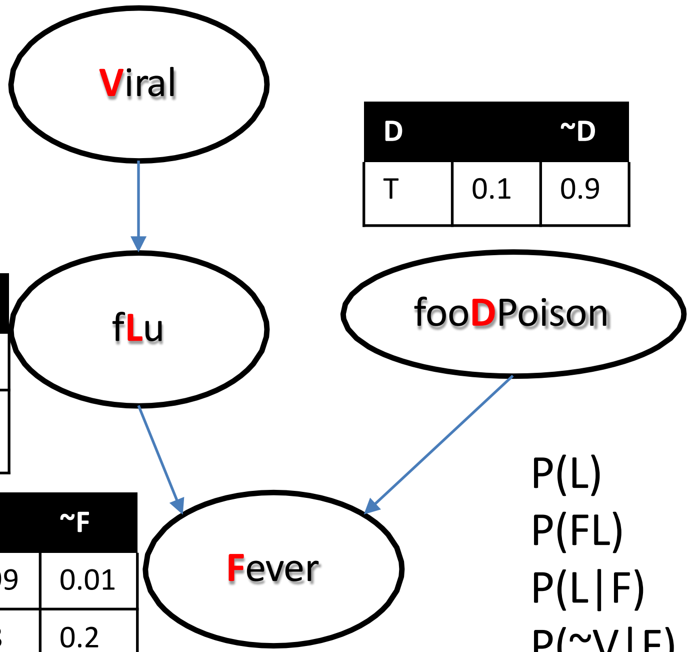
L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	L	D	F
T	T	F	T
F	F	F	F
T	F	F	T
F	T	F	T
..			
.....			

0.3, 0.2, 0.6, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0 . 9 , 0 . 5 5 ………. . .

Prior Sampling

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

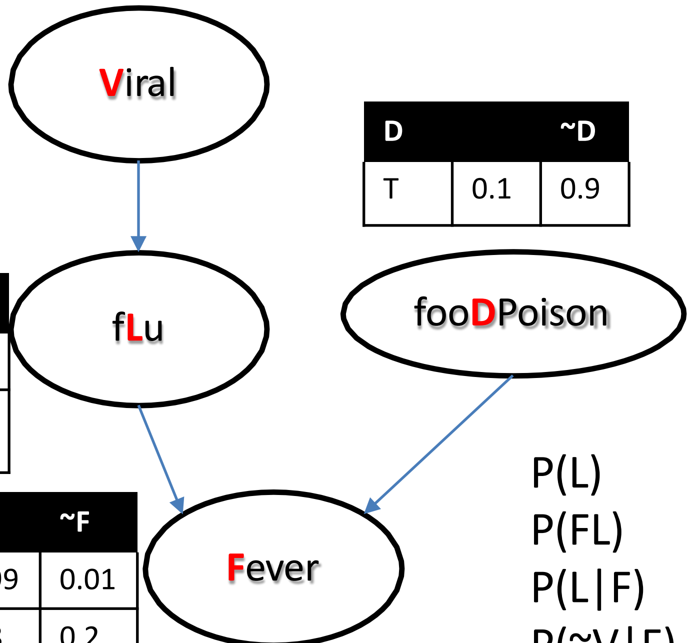
P(L) = 3/8
P(FL) = 3/8
P(L|F) = 3/5
P(~V|F) = 2/5
P(L|V~F) = 0
P(F|D) = ?????

V	L	D	F
T	T	F	T
F	F	F	F
T	F	F	T
F	T	F	T
T	T	F	T
T	F	F	F
F	F	F	T
T	F	F	F

Rejection Sampling

Sample Generation by Randomization

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	L	D	F
…..			

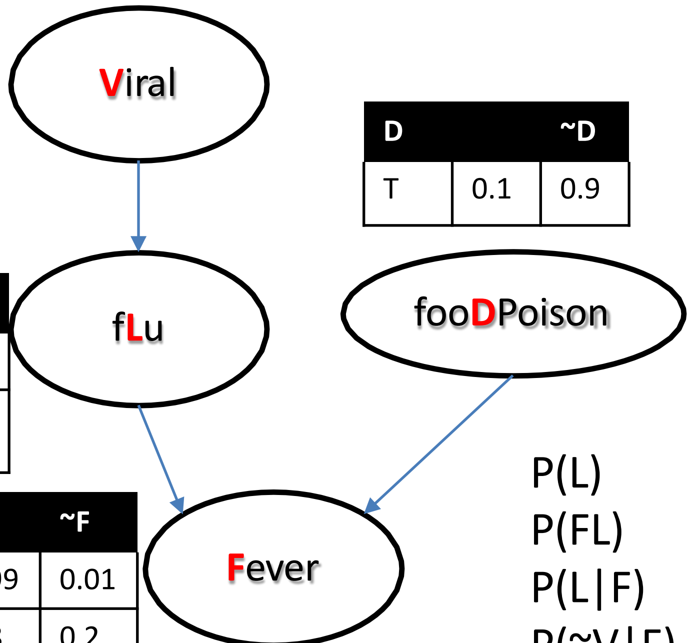
P(L) = 3/8
P(FL) = 3/8
P(L|F) = 3/5
P(~V|F) = 2/5
P(L|V~F) = 0
P(F|D) = ?????

0.3, 0.2, 0.6, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0.9,0.555,0.38………..

Rejection Sampling

Sample Generation by Randomization

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	L	D	F
T	T	F	
T	T	F	
T	T	F	
F	F	T	F
T	F	T	T
…..			

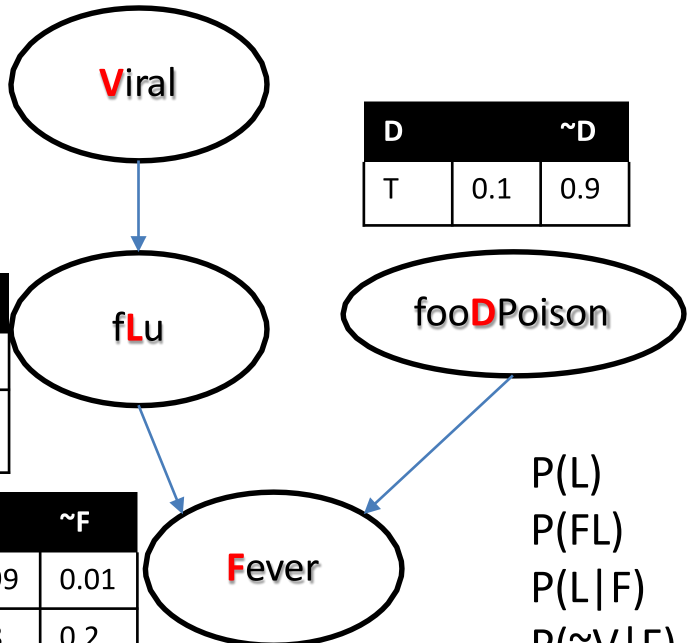
P(L) = 3/8
P(FL) = 3/8
P(L|F) = 3/5
P(~V|F) = 2/5
P(L|V~F) = 0
P(F|D) = ?????

0.3, 0.2, 0.6, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0.9,0.555,0.38………..

Rejection Sampling

Inference

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

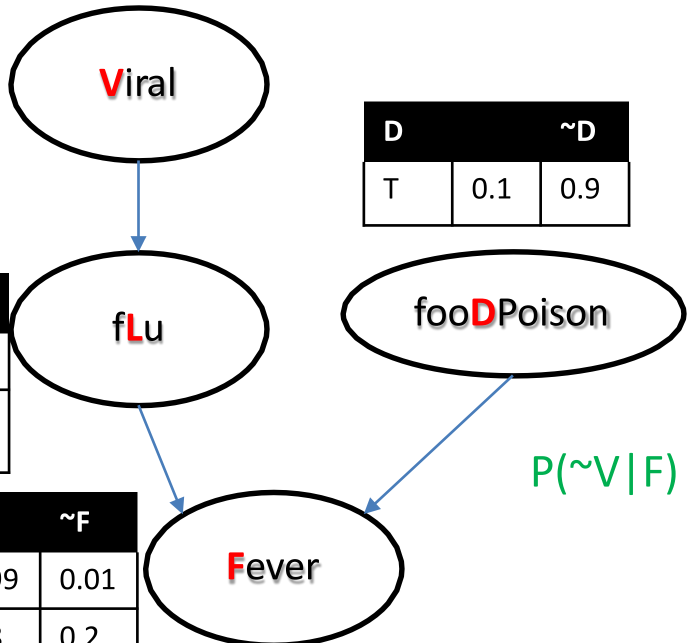
P(L) = 3/8
P(FL) = 3/8
P(L|F) = 3/5
P(~V|F) = 2/5
P(L|V~F) = 0
P(F|D) = 5/8

V	L	D	F
T	T	T	T
F	F	T	F
T	F	T	T
F	T	T	T
T	T	T	T
T	F	T	F
F	F	T	T
T	F	T	F

Rejection Sampling

Sample Generation by Randomization

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

P(~V|F) = ?

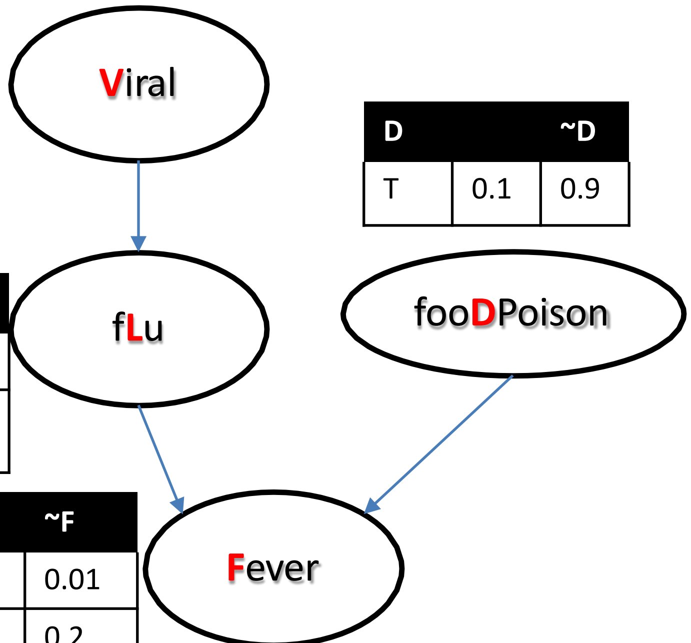
V	L	D	F
T	T	F	T
F	F	F	
T	F	F	T
F	T	F	T
…..			

0.3, 0.2, 0.6, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0.9 , 0.555 ,………..

Likelihood Weighing

Sample Generation by Randomization

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	L	D	F	wgt
F		T		
F		T		
F		T		
F		T		
F		T		
F		T		
F		T		

P(~F|D,~V) = 0.04 / 7*0.04

0.3, 0.2, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0.99, , ……..

Likelihood Weighing

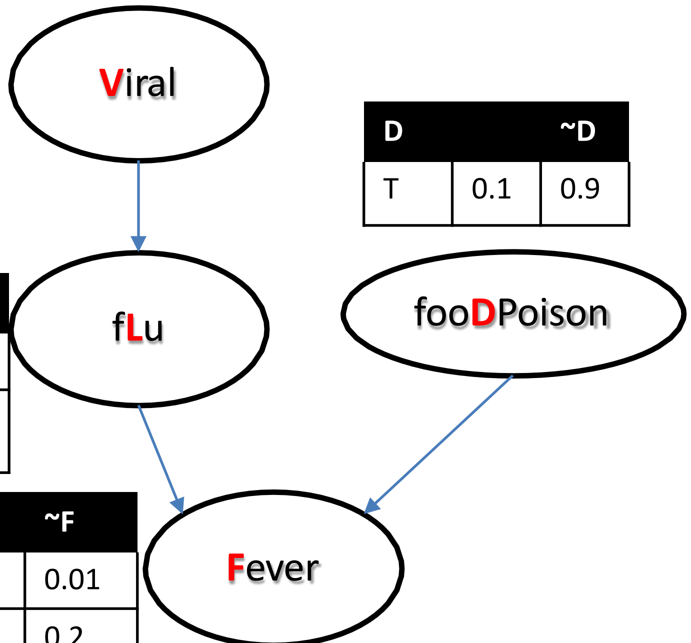
Sample Generation by Randomization

V		~V
T	0.6	0.4

D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95



V	L	D	F	wgt
F	F	T	T	0.4*1* 0.1 *1=
F	F	T	T	
F	F	T	T	
F	F	T	T	
F	F	T	T	
F	T	T	T	
F	T	T	F	

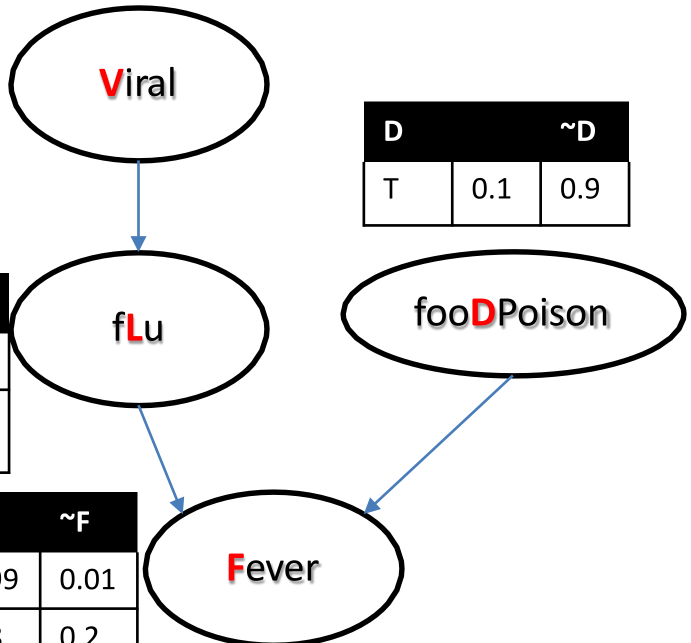
P(~F|D,~V) = 0.04 / 7*0.04

0.3, 0.2, 0.58, 0.73, 0.87, 0.15, 0.6, 0.57, 0.85, 0.12, 0.004, 0.93 , 0.0002, 0.99, ,……..

Likelihood Weighing

Inference

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

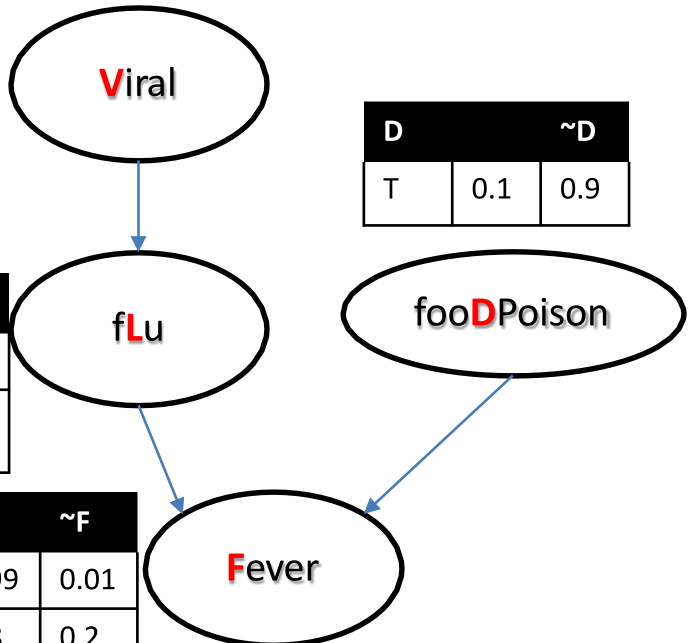
V	L	D	F	wgt
F	F	T	F	0.4*1* 0.1 *1=
F	T	T	T	0.4*1* 0.1 *1=
F	F	T	T	0.4*1* 0.1 *1=
F	F	T	F	0.4*1* 0.1 *1=

$P(F|D,\sim V) \qquad = 0.04+0.04 \quad / \; 4*0.04$

Likelihood Weighing

Inference

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

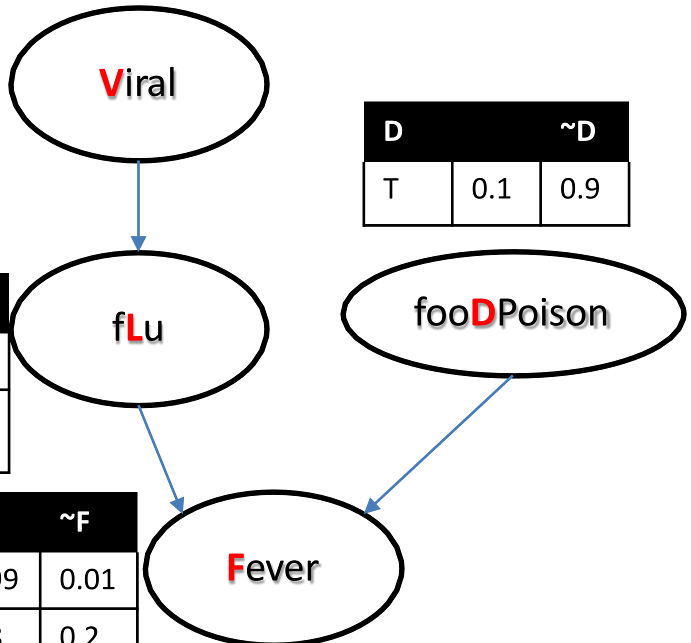
V	L	D	F	wgt
F	F	T	F	
F	F	T	T	
F	F	T	T	
T	F	T	F	

$P(F|D,\sim L) \quad = 0.099+0.099 \quad / \, (3*0.099 + 0.02)$

Likelihood Weighing

Inference

V		~V
T	0.6	0.4



D		~D
T	0.1	0.9

V	L	~L
T	0.8	0.2
F	0.01	0.99

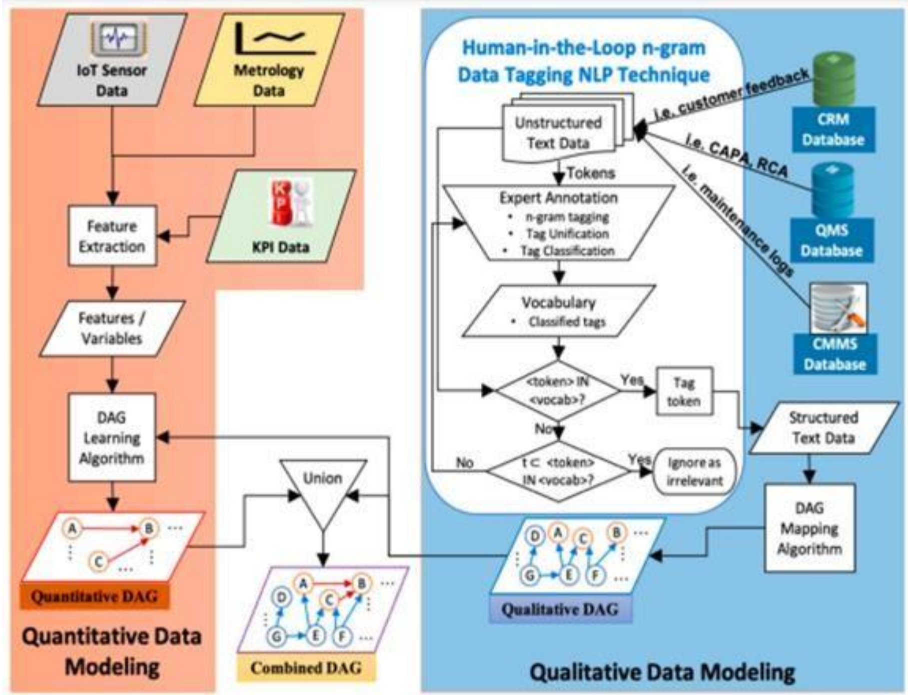
L	D	F	~F
T	T	0.99	0.01
T	F	0.8	0.2
F	T	0.8	0.2
F	F	0.05	0.95

V	L	D	F	wgt
F	F	T	F	1*0.99* 0.1 *1=
F	F	T	T	1*0.99* 0.1 *1=
F	F	T	T	1*0.99* 0.1 *1=
T	F	T	F	1*0.2* 0.1 *1=

$$P(F|D,\sim L) = 0.099+0.099 \quad / \left(3*0.099 + 0.02\right)$$

Bayesian Network

Fault Diagnostic System

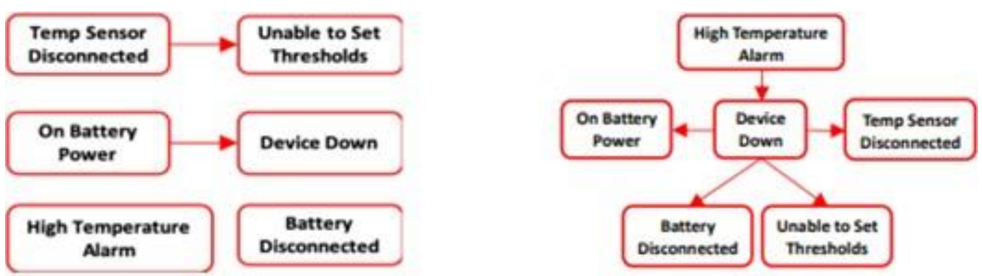


Source Credit : [Sensors 2021 : Fusion-Learning of Bayesian Network Models for Fault Diagnostics]

Bayesian Network

Fault Diagnostic System

Raw Data	Short Description			Resolution Notes		
	On battery power			Power outage due to transformer fire		
Classified Tags	Symptom		Cause(s)		Link	
	on_battery_power		power_outage, transformer_fire		due_to	
BN Mapping	Child Variable	Child State	Parent Variable	Parent State	Ancestor Variable	Ancestor State
	on_battery_power	yes	power_outage	yes	transformer	Fire



(a) (b)

Source Credit : [Sensors 2021 : Fusion-Learning of Bayesian Network Models for Fault Diagnostics]

Bayesian Network

Fault Diagnostic System

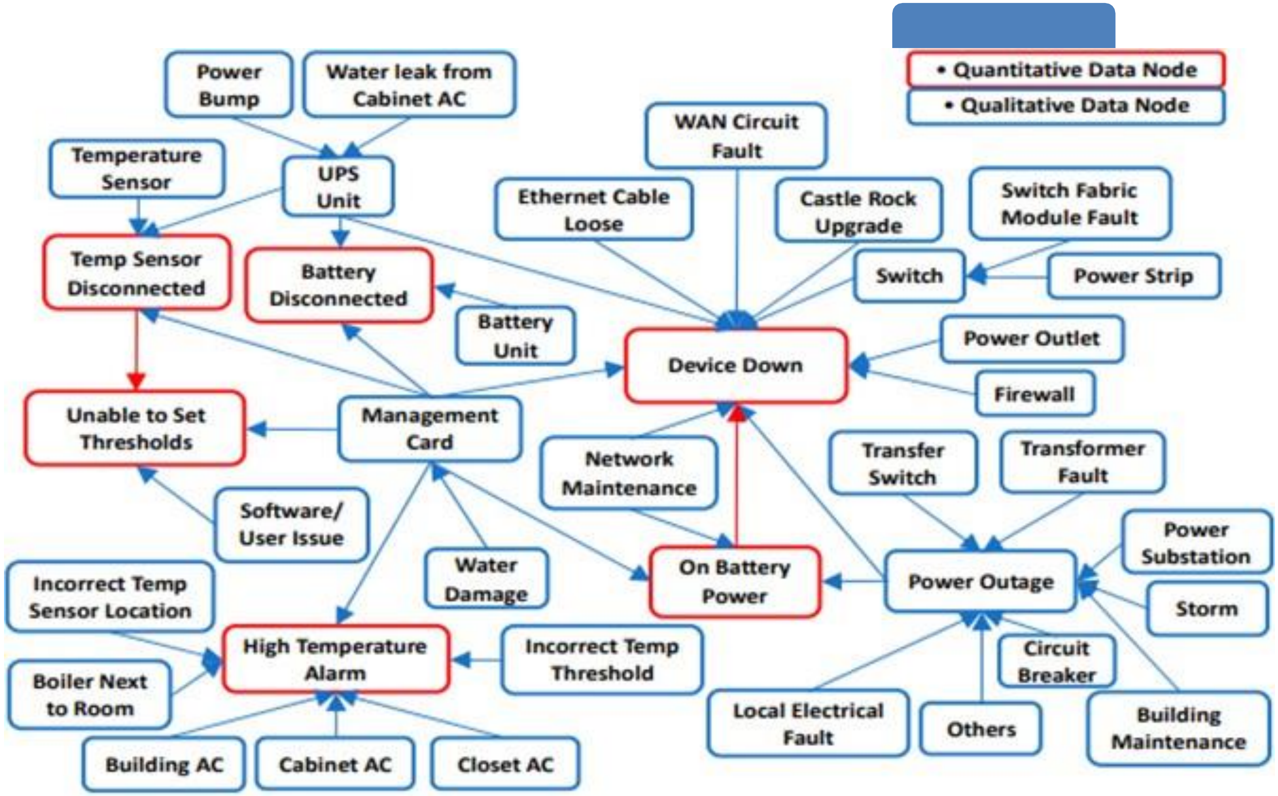
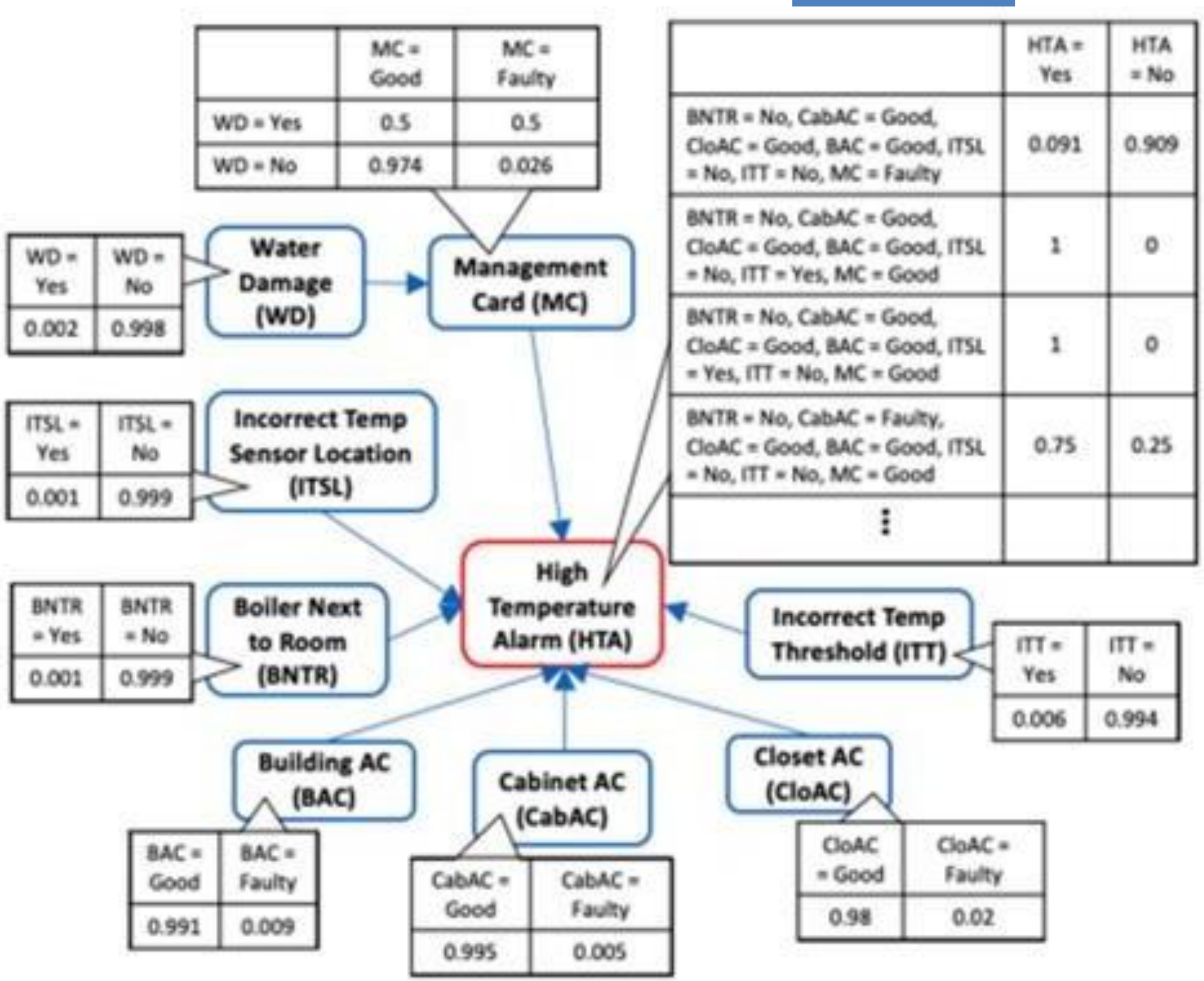


Figure 8. Fused Bayesian Network structure for top six occurring UPS messages.

Source Credit : [Sensors 2021 : Fusion-Learning of Bayesian Network Models for Fault Diagnostics]

Bayesian Network

Fault Diagnostic System



Source Credit : [Sensors 2021 : Fusion-Learning of Bayesian Network Models for Fault Diagnostics]

Required Reading: AIMA - Chapter # 4.1, #4.2, #5.1, #9 Refer to the handout

Next Session Plan:

- Hidden Markov Models

Thank You for all your Attention

Note : Some of the slides are adopted from AIMA TB materials