Constraint Satisfaction Problems

(CSP)

In search algorithms → what variables make up a state

→ determine the actions concern about the change of states; how a state can be changed, how the agent can go one state to another

→ never go inside a state, and try to manipulate the values of the variable agent

→ just perform action and change their state until it can find the goal state

Scanned with CamScanner

In CPS, we actually look up the variable that make up a state

→ try to manipulate the values of the variable and try to assign values to the variable in such a manner that we can actually construct the goal state

Formulation:

have to sepecify 3 things to formulating a csp

→ variables that make up a state

→ domains of the variable

→ set of constraints.

domain: allowable set of values that the variable can have/take

Constraint: the ~~ex~~ rules, or regulation or condition that you have to satisfy for a state to become a goal state

→ ~~always~~ when agent constrainting the goal state he always has to check against the set of constraint that none of being violated

set of constraints
ভঙ্গ করা যাবে না

→ if any constraint is violated then we can't call a state goal state.

Scanned with CamScanner

CSP Example: (Map Coloring)

Australia region (7 regions)
3 colors (Red, green, blue)

Problem: The objective of the agent is to color the map in such way that no two adjacent territory/regions can have the same color.

Soln: Variable: WA, NT, Q, NSW, V, SA, T

এমন entity যেগুলোতে different value দেওয়া যাবে

Domain: D = {red, green, blue}

Constraints: adjacent regions must have different colors

We can't write just like a line

2 ways to list/represent the constraints

$\underbrace{\text{Implicit}}_{\text{using equation}}$: WA ≠ NT

WA ≠ SA

NT ≠ SA

SA ≠ Q

⋮

→ Constraint list

(check while solving)

Or

$\underbrace{\text{Explicit}}_{\text{using value set}}$: (WA, NT) ∈ { (red, green),

(red, blue)

(blue, green)

(blue, red)

(green, blue)

(green, read) }

(WA, SA) ∈ . . . .

(NT, SA) ∈ . . . .

Scanned with CamScanner

Solutions are assingments satisfying all constraints

{WA = red, NT = green, Q = red, NSW = green, V = red, SA = blue, T = green}

Assignment

- → complete (all variables value assign)
- and → consistent (No constraint is violated)

→ it will be solution of this problem

Example: N-Queens:

— N by N checkerboard

have to place N queens in such a way that no queen can attrack ~~any one~~ each other.

Formulation 1:

- Variables: X_{ij}
- Domains: $\{0, 1\}$ — 0: no queen, 1: queen
- Constraints:

$X_{11} \neq X_{12}$ (Not express)

$\quad 0 \qquad\quad 0$

constraint:

Row: $\forall i, j, k \; (X_{ij}, X_{ik}) \in \{(0,0), (0,1), (1,0)\}$

$$\boxed{\begin{array}{l} i = 1, j = 1, k = 3 \\ \\ X_{11} \qquad\qquad X_{13} \end{array}}$$

Scanned with CamScanner

column $\rightarrow \forall i, j, k \ (x_{ij}, x_{kj}) = \{(0,0), (0,1), (1,0)\}$

Diagnol $\begin{cases} \forall i, j, k \ (x_{ij}, x_{i+k, j+k}) \\ \forall i, k, k \ (x_{ij}, x_{i+k, j-k}) \end{cases}$

Empty board কেউ solution বলতে পারে

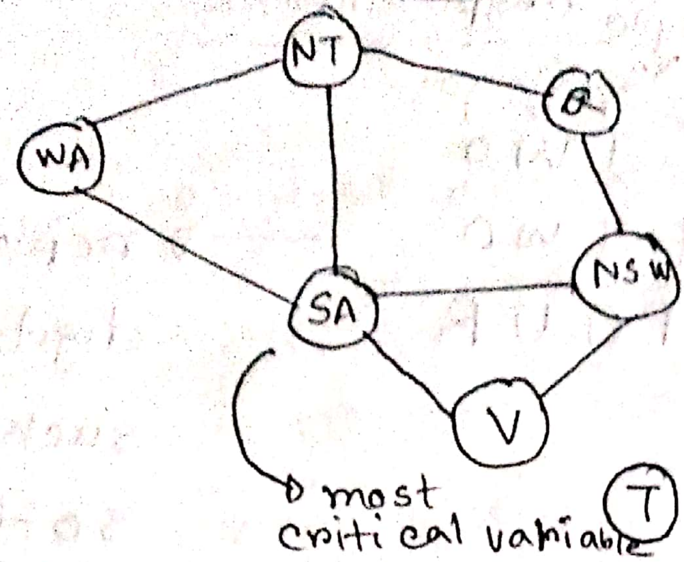
1 Q " " " "

2 Q " " " "

3 Q " " " "

$$\boxed{\sum_{i,j} X_{ij} = N}$$

Constraint Graph (Additional)



variable → nodes

constraint → arc

[যে যে variable এর মধ্যে আছে তাদের মধ্যে arc draw করে দিতে হবে]

Example: Cryptarithmetic

$$\begin{array}{r} X_3\ X_2\ X_1 \\ T\ W\ O \\ +\ T\ W\ O \\ \hline F\ O\ U\ R \end{array}$$

→ replace with digit in such manner so that the sum hold and all value distinct

- Variable: F T U W R O

- Domains:

$$\{0, 1, 2, 3, 4, 5, 6, 7, 8, 9\}$$

- Constraints:

all diff (F, T, U, W, R, O)

All needs satisfy

$$O + O = R + 10 \cdot X_1$$

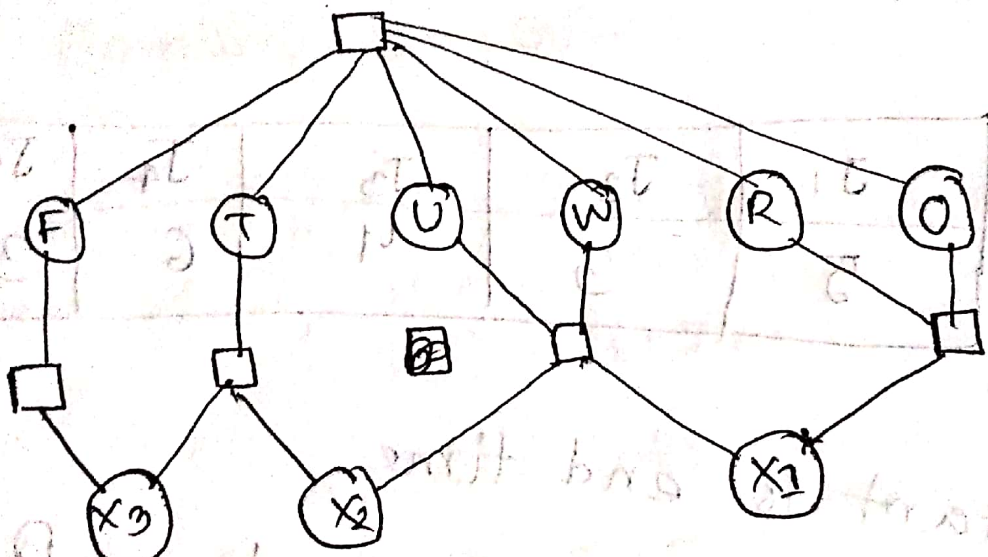
$$W + W + X_1 = U + 10 \cdot X_2$$

$$T + T + X_2 = O + 10 \cdot X_3$$

$$F = X_3$$

$$\begin{array}{r} 867 \\ +\,867 \\ \hline 1734 \end{array}$$

F	T	U	W	R	O
1	8	3	6	4	7



CSP popular in Scheduling problem

Job-scheduling (In text book)

Example- Job-scheduling

- In one day five jobs need to finish within 15 hours

J1	J2	J3	J4	J5
5	3	4	6	3

Duration (In hour)

start to end time

Constraint

- J1 must be done first
- J2 must be done after J3 is finished. (raw materials for J3)
- J2 and J4 cannot be done at the same time / in parallel (maybe same finish)

- J5 has to be done last.

Have to figure to all job's starting time.

Formulation:

Variable : T_1, T_2, T_3, T_4, T_5 (starting time)

Domain: 0-12

Constraint: (i) $T_2 >= T_1 + 5$,

$T_3 >= T_1 + 5$,

$T_4 >= T_1 + 5$,

$T_5 >= T_1 + 5$

(ii) $T_2 >= T_3 + 4$

(iii) $T_4 >= T_2 + 3$ or

$T_2 >= T_4 + 6$

(iv) $T_5 >= T_2 + 3$

$T_5 >= T_3 + 4$

Scanned with CamScanner

T5 >= T4+6

Solving CSP

Standard search formulation of CSPs

~~Intial State: the empth~~

~~Back~~t

Intially assigntment empty
all variable null

Then step এ একটা করে variable
এ value assign untill
every vanniable is
assigned and all
constraints are satisfied.

Scanned with CamScanner

Backtracking Search Algorithm

A Color-map

Intially blank

In iteration First task of every step to choose a variable and give a value and also check if it ~~any~~ doesn't break any constraint

~~suppose~~

One variable at a time
Check constrains as you go

যখনি ~~value~~ variable এ legal value assign করা যাচ্ছে তখন backtrack করতে হবে and last variable এ value assign (assignment) clear করতে হবে

Scanned with CamScanner

[ঐ variable এর জন্য ~~B~~ value বসাইতে হবে
backtrack করে]

backtrack
{ DFS থেকে different }
- ~~[illegible]~~ those siblings can satify constraint যেগুলো generate হবে
- failure occur করলে backtract করা যায়

[Backtracking ~~[illegible]~~ শুধু এক step না, ~~[illegible]~~ অনেক step করতে পারে]

Penformance:

Betton than DFS, BFS, any othen search algonithm,

- many nodes will be in the tree

- BFS will not work

- CSP will not explone all the paths, explonethe path satisfy all the constnaints

It we don't detect failune eanly, it's disasten. As bactnacting is ~~nedugtant~~ work

infisible

nedundant

Improving Backtracking

Can we detect inevitable failure early?

Technique: ~~[illegible]~~ Filtering

Filtering: Forward Checking

(to early detection method)

Cross off values that violate a constraint when added to the existing assignment.

WA	NT	Q	NSW
~~R, G, B~~	~~R~~, ~~G~~, ~~B~~	R, ~~G~~, B	R, ~~G~~, B

V	SA
R, G, ~~B~~	~~R~~, ~~G~~, ~~B~~

~~Demo~~

domain empty
failure occur
need to backtrack

(WA variable এর domain update)

Filtering: Constraint Propagation

~~more stronger~~ - works better

After assigning value to any variable, needs to check all the constraints in the constraints list to make sure there is a pair of value that can fulfill the constraint

Forward → ~~checking~~ checking constraint that the current variable involve

Constraint Propagation → all constraint

→ ~~য~~ যখনি একটার জন্য অন্যটার legal value assign করা যাচ্ছে না তখনি failure

Simulation → Forward checking (In exam)

Ordering

→ to avoid backtracking

Ordering
- → which variable? → MRV → degree
- → which value? → LCV

Ordering: Minimum Remaining Values

যে variable এর value/domain কম, তাকে আগে select

→ goal domain empty

The idea used in Minimum remaining value heuristic is that we have choose a variable which has the ~~lost~~ least number of values remaining in it's domain because that variable might cost problem for in future if we don't clear it right now.

domain number tie হলে

degree heuristic

degree → কয়টা করে connect

SA → degree 5

↳ is the most involved critical/problematic variable in this model

5 [illegible] variable এর সাথে

constraint থাকে।

Least Constraining Value

Q তে

unassigned variable	SA	NS	
	B	R G B	
Q = R	✗ (not affected)	✓ (affected)	1
Q = B	✓	✓	2

2টা variable affected হচ্ছে

R is better choice here

Before all algorithm → for deterministic environment
where is no uncertainity

In environment → uncertain component
randomness
~~value may be~~
→ multiple options for value

→ uncertain
to solve
→ some model

Random variable: value surely জানি না

R = Is it raining?

T = Is it hot or cold?

D = How long will it take to drive to work?

R in {true, false}

T in {hot, cold}

D in $[0, \infty)$

→ Table

Probability Distributions: A table which associate a probability to the different value of the random variable

discreate	Continuous
{true, false}	$[0, \infty)$
{hot, cold}	

Probability distribution are applicable to ~~of~~ discreate random variables

In case of discreate random variable if we assign probability value to each of the possible value of the variable, then the table that result is know as the probability distributions for the random variable

P(T)

T	P
hot	0.5
cold	0.5

→ probability distributions for the random variable temperature

Properties of probability Distributions:

For all u

$$\forall u \; P(X=u) \geq 0$$

$$\sum_{u} P(X=u) = 1$$

Join Distributions:

$\boxed{2\times2=4}$ → join distribution table row

$P(T, W)$

T	W	P
hot	sun	0.4
hot	rain	0.1
cold	sun	0.2
cold	rain	0.3

→ for this combination

this type of ~~ta~~ table is joint distribution

sum 1

$$P(hot) = 0.4 + 0.1 = 0.5$$

Scanned with CamScanner

Join distribution is also called

$\boxed{\text{Probabilistic Models}}$

From a joint distribution, we can calculate the probability of any event

---

$$P(\text{hot or sunny}) = 0.4 + 0.1 + 0.2$$
$$= 0.7$$

Marginal Distributions:

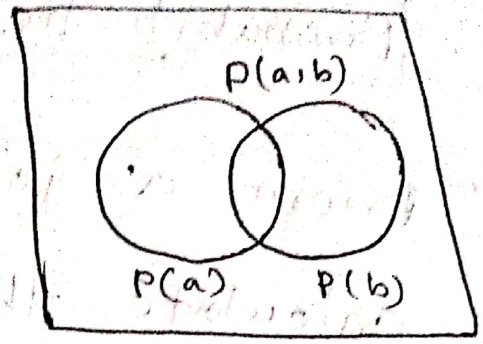
বড় table থেকে sub table create করা

$P(T, W)$ $\qquad\qquad$ $P(T)$

$$P(t) = \sum_{s} P(t, s)$$

T	P
hot	0.5
cold	0.5

Conditional Probabilities:



what is the value of this

given

temp=? $P(a|b) = \frac{P(a,b)}{P(b)}$ → join probability

→ given / prior

weather → sunny

$$P(W = s \mid T = c) = ???$$

$$= \frac{P(W = S, T = C)}{P(T = C)}$$

$$= \frac{0.2}{0.5} = 0.4$$

$$* \; P(+u \mid +y) = \frac{P(+u, +y)}{P(+y)}$$

$$= \frac{0.2}{0.6} = 0.33$$

$$* \; P(-u \mid +y) = \frac{0.4}{0.6} = 0.66$$

$$* \; P(-y \mid +u) = \frac{P(-y, +u)}{P(+u)} = \frac{0.3}{0.5} = 0.6$$

Conditional Distributions

Given value এর জন্য query variable এর সবগুলো possibility list করলে ঐ table কে conditional distributions বলে

$P(W \mid \underbrace{T = hot}_{\text{fixed}})$

W	P
sun	0.8
rain	0.2

$P(W \mid T = cold)$

sun	0.4
rain	0.6

$P(W \mid T)$ → group of conditional distributions

$P(W = s \mid T = c) = ?$ (single pro. ~~conditional~~ d value)

~~$P(W \mid T) = ??$~~

$P(W \mid T = c)$ → a conditional distributions

$P(W \mid T)$ → a group of conditional distributions

Full Joint Probability Distribution

3 random variabs

$2 \times 2 \times 3 = 12$

frequency list

8/50 → sum of all enties

8

indivual entity

normalization

Σ(sum = 1)

of all probability

$$\frac{8}{50} + \frac{6}{50} + \cdots \frac{2}{50} = 1$$

(normalization

$$P(L|T) = \frac{2}{50} + \frac{2}{50}$$

$$= \frac{4}{50}$$

Probabilistic Inference

Inference by Enumeration

3 steps:

3 types variabls.

Evidence variables $= E_1 \ldots$

$E_k = e_1 \ldots e_k$

(Right)

দেওয়া আছে (Given value)

Query* variable → Q (Left)
(Calculate)

Hidden variable → (Not right or left)
irrelevant for
for the problem

$2 \times 2 \times 3 = 8$

weather

P(W| winter)

W	P
sun	
rain	

E → S

Q → W

H → T

1st: select the enties constraint withe evidence

2nd: Hidden variable element বাদ দিতে হবে

~~3rd: common raw~~

Evidence থেকে select করা table থেকে summer

W	P	
sum	0.25/0.5	= 0.5
rain	0.25/0.5	= 0.5

3rd: Normalize

P(W) = ??

E → X

Q → W

H, T, S

W	P
sun	0.65
rain	0.35

0.65 + 0.35 = 1

So, no need to normalize

$E \longrightarrow S$, hot.

$Q \longrightarrow W$

$H \longrightarrow X$

W	P
sun	0.1/0.15
rain	0.05/0.15

The Product Rule

$$P(y) \cdot P(x|y) = P(x, y)$$

$\underbrace{P(y)}_{\text{Prior prob}} \cdot \underbrace{P(x|y)}_{\text{Cond}^n \text{ prob}} = P(x,y)$ Join probability

P(W)

W	P
sun	0.8
rain	0.2

P(D|W)

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

sum = 1

sum = 1

P(D, W)

D	W	P
wet	sun	0.08
dry	sun	0.72
wet	rain	0.14
dry	rain	0.06

$P(w,s) = P(w|s) \cdot P(s)$
$= 0.1 \times 0.8$
$= 0.08$

$P(d, s) = P(d|s)\, P(s)$
$= 0.9 \times 0.8$
$= 0.72$

The chain Rule

- to join more than two random variable

$$p(x_1, x_2, x_3) = p(x_3 | x_1, x_2) \; p(x_1, x_2)$$

$$= p(x_3 | x_1, x_2) \cdot p(x_2 | x_1) \; p(x_1)$$

$$p(x_1, x_2, x_3, x_4) = p(x_1) \; p(x_2 | x_1) \; p(x_3 | x_1, x_2) \; p(x_4 | x_1, x_2, x_3)$$

$$p(x_1) \; p(x_2 | x_1) \; \ldots\ldots \; p(x_n | x_1 \ldots x_{n-1})$$

Bayes Rule

↳ one of the most important rule of AI

$$P(x,y) = P(x|y) \cdot P(y) = P(y|x)\, P(x)$$

$$P(x|y) = \frac{P(y|x) \times P(x)}{P(y)}$$ → Bayes' Rule

Cause | effect → বের করতে পারব

$$P(+m) = 0.0001$$

$$P(+s|+m) = 0.8$$

$$P(+s|-m) = 0.01$$

does not have man

$$P(+m|+s) = \frac{P(+s|+m)\,P(+m)}{P(+s)}$$

$$= \frac{P(+s|+m)\,P(+m)}{P(+s|+m)\,P(+m) + P(+s|-m)\,P(-m)}$$

$$= \frac{0.8 \times 0.0001}{0.8 \times 0.0001 + 0.01 \times \underbrace{0.999}_{(1-P(m))}}$$

$$P(+s) = P(+s, +m) + P(+s, -m)$$

removing the hidden variable.

$$= P(+s|+m)\,P(+m) + P(+s|-m)\,P(-m)$$

Find,

$P(u|y) \to$ condn

$P(u, y) \to$ The product Rule

$P(y|u) \to$ Bayes Rule

First find the rule

p(W)

R	P
sun	0.8
rain	0.2

P(D|W)

D	W	P
wet	sun	0.1
dry	sun	0.9
wet	rain	0.7
dry	rain	0.3

$p(W \mid dry) = ?$

W	P
sun	0.92
rain	0.08

$p(sun \mid dry)$

$$= \frac{p(dry \mid s) \cdot p(s)}{p(d)}$$

$$= \frac{0.9 \times 0.8}{0.9 \times 0.8 + 0.3 \times 0.2}$$

$$= \frac{0.72}{0.78} = 0.92$$

$$p(d) = p(d, sun) + p(d, r)$$

$$= p(d \mid s)\, p(s) + p(d \mid r) \cdot p(r)$$

$$= 0.9 \times 0.8 + 0.3 \times 0.2$$

Scanned with CamScanner

Probabilistic Independence

দুইটা random variable কেউ কাউকে influence না করলে তাদের probalistic Independence বলা যাবে।

Two variables are independent if:

$$\forall x, y = P(x, y) = P(x)\,P(y)$$

$$\forall x, y = P(x \mid y) = P(x)$$

$$X \perp\!\!\!\perp Y$$

Join table = multiple table না হলে
Independence হবে না

True indepence: Coin flip

Conditional Independence

অন্য variable এর value given থাকলে ~~অবশ্য~~ তাদের independence বলা যাবে

P(Toothache, cavity, catch)

~~P~~ True Independent: $P(x,y) = P(x)\,P(y)$

$$P(x,y\mid z) = P(x\mid z)\,P(y\mid z)$$

$$P(x\mid z,y) = P(x\mid z)$$

$$X \perp\!\!\!\perp Y \mid Z$$

P(catch | Toothache, cavity) = P(catch | cavity)

P(Toothache | catch, cavity) = P(Tootache | cavity)

(Toothache, catch | cavity) = P(Toothache | cavity) P(catch | cavity)

$T \perp\!\!\!\perp U \mid R$

$Fire \perp\!\!\!\perp Alarm \mid Smoke$

Fire and Alarm are independent each other given smoke.

Conditional Independence and the chain Rule

$$P(x_1, x_2, \ldots x_n) = P(x_1)\ P(x_2 \mid x_1)\ P(x_3 \mid x_1, x_2) \ldots$$

$$P(R, T, U) = \underbrace{P(R)\ P(T \mid R)\ P(U \mid R, T)}$$

$$= P(R)\ P(T \mid R)\ P(U \mid R)$$

$U \perp\!\!\!\perp T \mid R$

$P(U \mid R, T) = P(U \mid R)$

~~[illegible]~~

থাকবে

বেশি random variable term simply

করা যাবে

[chain rule factorize

↓

Conditional independence Relⁿ]

apply

Bayes's Nets

full joint probability distribution model (pre)

→ Probabilistic model

Using it:

→ answer lot of queries

→ " " " inference

Why another (full joint p distribution model):

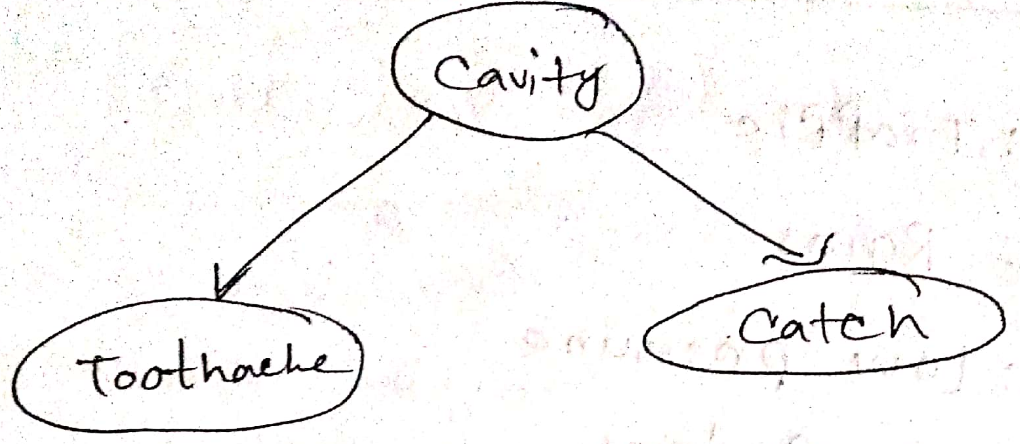
limitation

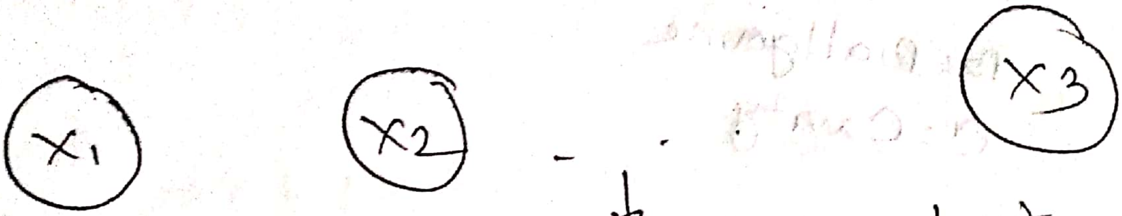
→ it's size

(# no of random variable)

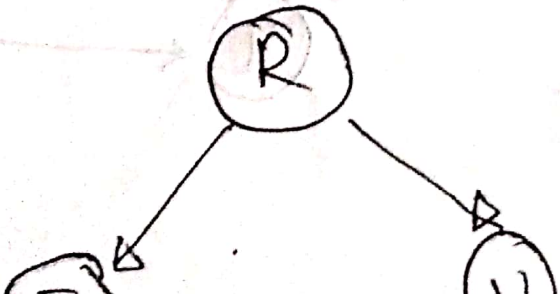
(# no of options)

unmanageable





Truely independent,
no any
no one infiluence others



Construct Baye's Net:

T: Triaffic

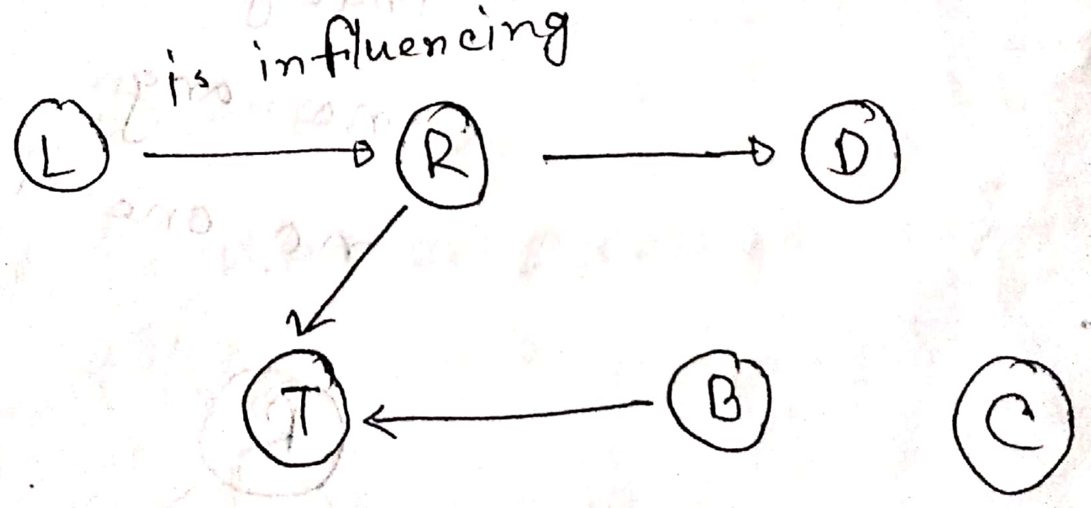
R: Rains

L: Low pressure

D: Roof drips

B: Ballgame

C: Cavity

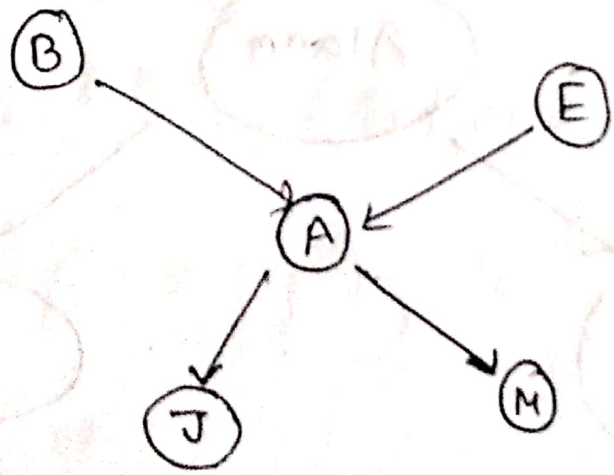


Alarm Network:

B: Burglary
A: Alarm goes off
M: Many calls
J: John calls
E: Earthquake

node

interaction → arg

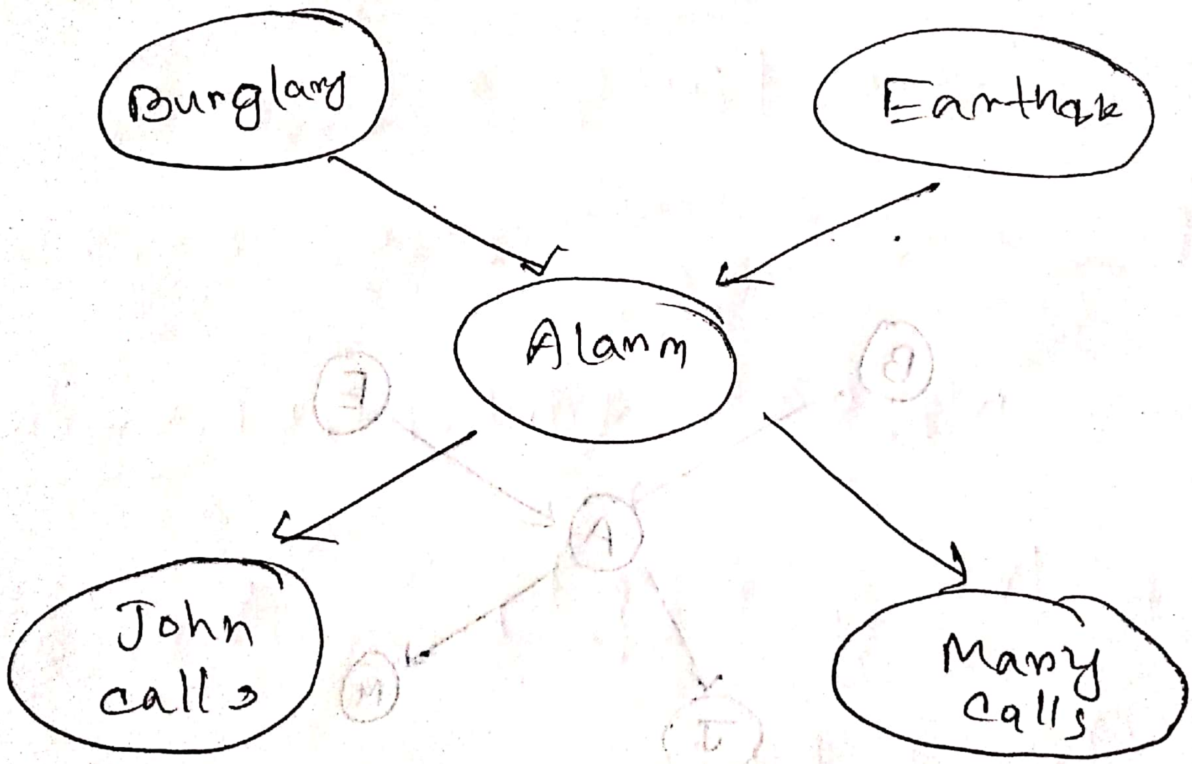


acyclecle [not form any cycle

Baye's Net Semantics:

for each random variable construct a table

$\underbrace{\text{table}}$ → conditional distribution table



$P(x_i \mid Par(X_i))$

B	P(B)
+b	0.001
-b	0999

Number of enties is much lower than for full join distribution table.

A random variable in a bayes' net is conditionally independent of it's nondesence giving it's parent

random varialbe -o শুধু তার parent এর সাথে related অন্য random variable থেকে independent

Probabilities in BNs

Chain:

$$p(x_1, \ldots, x_n) = p(x_1)\, p(x_2 \mid x_1)\, p(x_3 \mid x_1, x_2) \ldots p(x_n \mid x_1 \ldots x_{n-1})$$

$$= p(x_i \mid \text{parents}(x_i))$$

→ 5টি conditional term

$$\underbrace{p(+b, -e, +a, -j, +m)}_{} \rightarrow \text{apply chain rule}$$

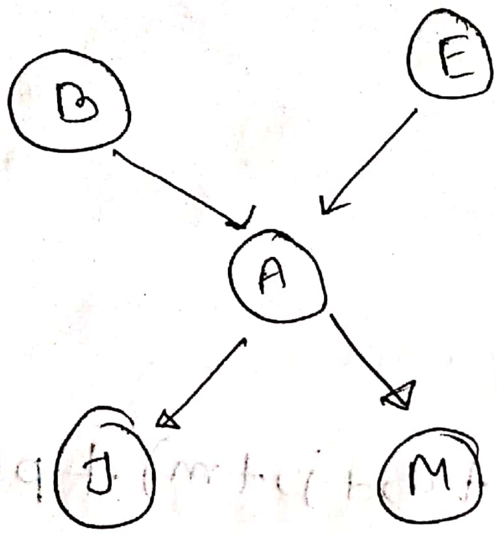
b has no parent) e has no parent

$$= p(+b) \times p(-e) \times p(+a \mid +b, -e) \times p(-j \mid +a) \times p(+m \mid +a)$$

$$= 0.001 \times 0.998 \times 0.94 \times 0.1 \times 0.7$$

Inference:

by Enumeration in Baye's Net



$$P(B \mid +j, +m) \Rightarrow \begin{array}{l} P(+b \mid +j, +m) \\ P(\neg b \mid +j, +m) \end{array}$$

B	P
+b	
¬b	

$$p(+b \mid +j, +m) = \frac{p(+b, +j, +m)}{p(+j, +m)}$$

hidden variable

A	E
+a	+e
+a	−e
−a	+e
−a	−e

$$p(+b, +j, +m)$$

$$= p(+b, +e, +a, +j, +m) + p(+b, +e, -a, +j + m) + p(-b, -e, +a, +j, +m)$$

$$= p(+b) \cdot p(+e) \cdot p(+a \mid +b, +e) \cdot p(+j \mid +a) \cdot p(+m \mid +a) + p(+b) \cdot p(+e) \cdot p(-a \mid +b, +e) \cdot p(+j \mid -a) \, p(+m \mid -a)$$

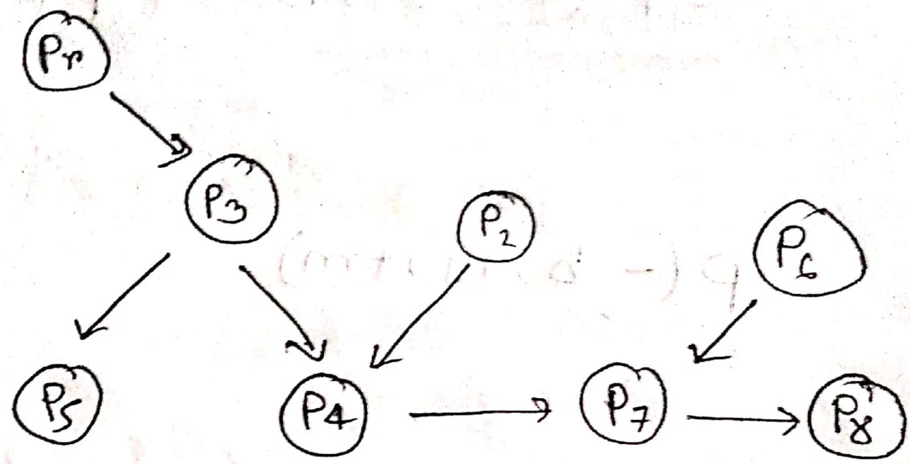
$P(+j, +m)$

$= P(+b, +j, +m) + P(-b, +j, +m)$

$P(-b, +j, +m)$

B	P
$-b$	0.229
$-b$	0.771

Independence in Bayes Nets



→ parient given

→ এরা আরেকটার chid হইতে পারবে না

conditionally independent

$$P(P_8 \mid P_3, P_7) = P_r(P_8 \mid P_7)$$

$$P_8 \perp\!\!\!\perp P_3 \mid P_7$$

P_3 , P_{8} $\Rightarrow$ check

এরা এক জন আরেক জনের child না

$P_3 \perp\!\!\!\perp P_5 \mid P_1$

chid তারা

$P_4 \perp\!\!\!\perp P_2 \mid P_3$

$P_5 \perp\!\!\!\perp P_4$ P_1

Scanned with CamScanner

Naive Bayes Classifier

→ Machine learning এর important topic classification

Classification

কোনো object কে কোনো class এ classifi করা।

▢ তিনটি Box এ দেওয়া আছে, Apple কে একটাই রাখতে হবে

→ using the features to determine which class the fruit belongs to

→ by observing level sample ~~experience~~

Two phrase:

(i) Training → (provide the classifier agent enough amount of level sample/level data so that

~~(ii) Testing~~

classifier agent can learn which features are associated with which class

<u>Testing:</u> by giving it unlavel sample

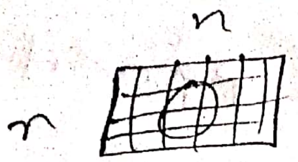
SPAM Filter

1000 inbox
1000 spam mail

Then learn which mail are belongs to which

Handwritten Digit Recognition

[Hand-written level sample দিতে হবে , সংখ্যাটা কি বলে]



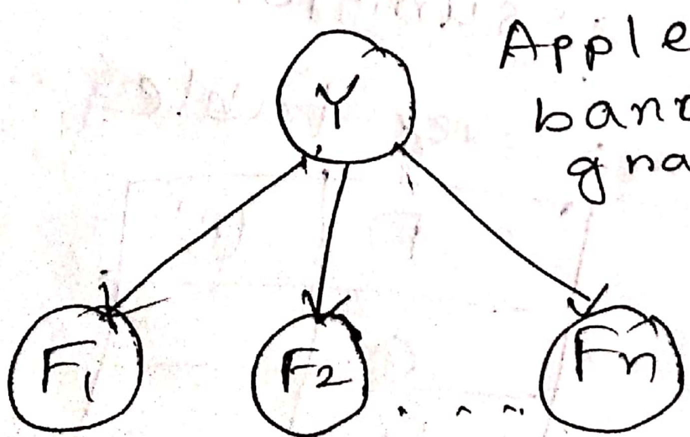
n×n pixel

কোন কোন pixel black
" " " white

} এটা থেকে determine করা যাবে কোন digit

Model-Based Classification

L→ using bayes net (probabilistic model)



Apple
banana
grapes

shape	colour	size
round	red	big
long	green	small
	yellow	
	purple	

class level is the parent of all the features.

It is assumtion the features are conditionally independent of each other given the class level

It might not be alway true. But if we use this assumption than the classifition will good and it's safe to make this assumtion

class level table

F	(P)
a	
b	
g	

features table

$P(F|Y)$

S	F	
r	a	
r	b	
r	g	

~~Fruits~~ shape given fruit

C	F	

Color given fruit

In the training stage, agent has given lot of data. Agent's job is to populate this table from the level data.

মান বসাইতে হবে table এ

unnamed sample আসবে → can determine which class the object belong

$y = \{y_1, y_2, y_3 ...\}$ $\langle f_1, ... f_n \rangle$

$p(y_1 | f_1 ...... f_n)$

$p(y_2 | f_1 ...... f_n)$

$p(y_3 | f_1 ...... f_n)$

The agent is going to classify the object in the class which has the highest probability of being selected. That's the naive bayes classfier work.

প্রত্যেক individual class এর condn probability calculate করা হবে। ~~[illegible]~~

~~given~~ যে probability টা maximum এ class এই object টাকে classify করা হবে।

$$P(y_1 \mid f_1 \ldots \ldots f_n) = \frac{P(y_1, f_1 \ldots \ldots f_n)}{P(f_1, \ldots \ldots f_n)}$$

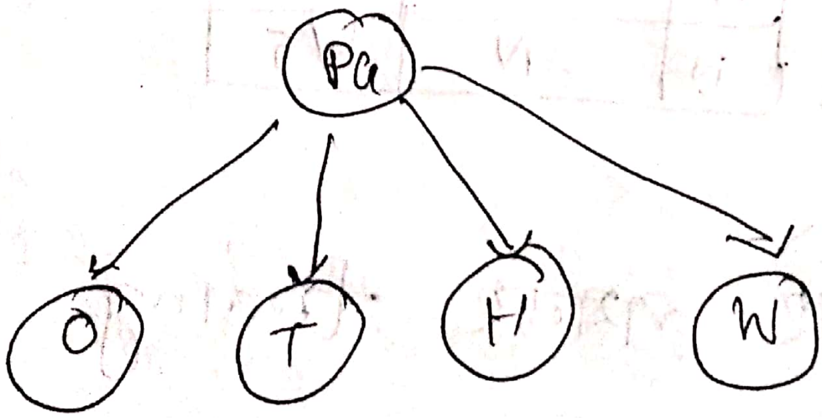
(join)

(prior)

conditional probability

Example (Golf)

base-ment table populate



PG	P
Y	9/14
N	5/14

O	PG	P
S	Y	3/9
O	Y	4/9
R	Y	2/9
S	N	2/5
O	N	0/5
R	N	3/5

Yes 9 টা
9 টার মধ্যে কয়টা
Sunny আছে

H	PG	P
H	Y	3/9
N	Y	6/9
H	N	4/5
N	N	1/5

এই পর্যন্তই traing step.

Testing:

<sunny, hot, noramal, false>

figure out that day is suitable for play golve or not

বের করতে হবে $p(y \mid s,h,n,f)$, $p(n \mid s,h,n,f)$

for denominator

$$p(y \mid s,h,n,f) = \frac{p(y,s,h,n,f)}{p(s,h,n,f)}$$

$$= p(y)\, p(s \mid y)\, p(h \mid y)\, p(n \mid y)\, p(f \mid y)$$

[conditional probability এর product]

$$= 9/5 \times \frac{3}{9} \times \frac{3}{9} \times \frac{6}{9} \times \frac{9}{14}$$

$$= 0.0141$$

$$p(r, s, h, n, f) = p(n) \times p(s|n) \times p(\bar{h}|n) \times p(n|n) \times n(f|n)$$

$$= \frac{2}{5} \times \frac{2}{5} \times \frac{1}{5} \times \frac{2}{5} \times \frac{5}{14}$$

$$= 0.0068$$

$$P(\text{Yes}|\text{today}) = \frac{0.0141}{0.0141 + 0.0068}$$

$$= 0.67$$

$$P(\text{No}|\text{today}) = \frac{0.0068}{0.0141 + 0.0068}$$

$$= 0.33$$

$$P(\text{Yes}|\text{today}) > P(\text{No}|\text{today})$$

Yes class এ classify করবে,

∴ This day is suitable for playing golf

drawback:

$$x \frac{0}{85} x$$

↳ 0 (total result)

it means, you have not enough data

It doesn't mean you can not play golve in that day.

use a Smoothing technique:

la-placian smoothing technique

add small positive value with every probability value using in the calculation

$$\frac{n+k}{y+dk}$$

↳ number of options

Scanned with CamScanner

(outlook এর ক্ষেত্রে options - 3)

<u>for every calcutation:</u>

$$\frac{3+1}{9+3} \times \frac{2+1}{9+3} \times \frac{6+1}{9+2} \times \frac{6+1}{9+2} \times \frac{9+1}{14+2}$$

Markov Models

↳ Probability Model

- Full Join Probability Distribution Table
  ↳ ~~de~~ determinate any query using inference by ~~enumitation~~ enumeration

- Bayes Net
  ↳ calculate probability " " by enumeration

∴ So, why Markov??

Actually markov model is not used for a network of multiple random variables. Markov model is used to model the heavier of a single random variable over different periods of time.

over different period of time কীভাবে variable টার value change হচ্ছে। এটাই ~~কীভাবে~~ modeling করবো

First order Markov Model

In this model it is assumed that every state of a random variable is influenced by only its previous state, not the state before that.

[different diferent time instance এ কোন একটা random variable এর state / value কী হতে পারে সেটাকে indicate করতেছি different different number দিয়ে।]

First order markov model এ, an assumtion is made that every State is dependent/ on influenced of it's previous state, and not the State before that.

~~Ideally তেমনটি হবে না~~

Ideally তেমনটি হবে না।

X_1 influences X_2,

X_2 ,, X_3

$$p(x_1)\ p(x_2|x_1)\ p(x_3|x_2)\ p(x_4|x_3)$$

Another Assumption: transition probabilities the same at all time

[X_2 কেবল X_1 এর উপর Dependant

X_3 কেবল X_2 ,, ,, ,,]

একটা single table for all

Example: Weather

$X = \{rain, sun\}$

Intial distribution: 1.0 sun (Today is the sunny day)

$P(X_1 = S) = 1$

$P(X_1 = r) = 0$

CPT $P(X_t \mid X_{t-1})$:

আগের

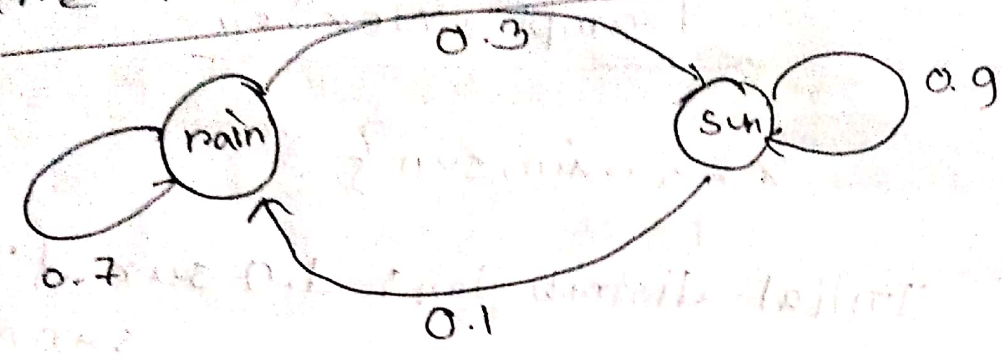
পরের দিন

$P(X_t \mid X_{t-1})$

X_{t-1}	X_t	$P(X_t \mid X_{t-1})$
sun	sun	0.9
sun	rain	0.1
rain	sun	0.3
rain	rain	0.7

state trasition Diagram

Futture Probability Calculation:



rain given sun → 0.1

sun ,, rain → 0.3

[যেটা থেকে বের হবে এইটার probability]

$$P(x_2 = sun) = P(x_2 = s, x_1 = s) + P(x_2 = s, x_1 = r)$$

$$\boxed{P(x, y) = P(x \mid y)\, P(y)}$$

$\cancel{p(x_2=sun)}=$

$$= p(x_2 = sun \mid x_1 = sun)\; p(x_1 = sun) +$$

$$\underbrace{p(x_2 = sun \mid x_1 = rain)}_{\text{(যেহেতু যাচ্ছে given}}\; p(x_1 = rain)$$

$$= 0.9 \times 1 + 0.3 \cdot 0 \cdot 0$$

$$= 0.9$$

$$p(x_2 = r) = 1 - 0.9$$

$$= 0.1$$

x_3 → directly x_1 থেকে পারবো না

so, first x_2 calculate করতে হবে

$$p(x_3 = sun) = p(x_3 = s,\ x_2 = s) + p(x_3 = s, x_2 = r)$$

$$= p(x_3 = s \mid x_2 = s) \times p(x_2 = s) +$$

$$p(x_3 = s \mid x_2 = r) \times p(x_2 = r)$$

$$= 0.9 \times 0.9 + 0.3 \times 0.1$$

$$= 0.84$$

$p(X_3 = r) = 1 - 0.84$

$= 0.16$

Day after tomorrow → first tomorrow
→ then day after tomorrow

$\cdots \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix} \begin{pmatrix} 0.75 \\ 0.25 \end{pmatrix}$

একটা সময় গিয়ে বার বার repeat করতেই থাকবে

converge ~~করে~~ same value repeat করতে থাকে

- stationary distribution

- depends on table value

Those value where the convergence going to happened (the value which is known as stationary distribution completely depends on the conditional probability table, it does not depend on the initial distribution at all as after long time the initial distribution কোনো effect ফেলতে পারে না।

How to calculate stationary value???

$$\left.\begin{aligned} P(X_\infty = s) &= 0.75 \\ P(X_\infty = n) &= 0.25 \end{aligned}\right\}$$

$$P_\infty(x) = P_{\infty+1}(x)$$

$$P_{\infty+1}(s) = P(s|s)\, P_\infty(s) + P(s|n)\, P_\infty(n)$$

$$P_\infty(s) = P(s|s)\, P_\infty(s) + P(s|n)\, P_\infty(n)$$

$$P_\infty(s) = 0.9\, P_\infty(s) + 0.3\, P_\infty(\text{rain})$$

$$\Rightarrow 0.1\, P_\infty(s) = P_\infty(n)$$

$$\Rightarrow P_\infty(s) = 3\, P_\infty(n) \quad \ldots\ldots (i)$$

$P_\infty(sun) + P_\infty(r) = 1 \quad \ldots \ldots \text{(i)}$

$P_\infty(s) = \frac{3}{4}$

$P_\infty(r) = \frac{1}{4}$

[Conditional probability table দেওয়া থাকলেই বের করা যাবে, initial value দরকার নেই]

3 variable

x in $\{a, b, c\}$

Transition matrix: conditional probability

$$
\begin{array}{c|ccc}
 & & x_t \rightarrow & \\
x_{t-1} & a & b & c \\
\hline
a & \frac{2}{5} & \frac{2}{5} & \frac{1}{5} \\
b & \frac{1}{5} & \frac{3}{5} & \frac{1}{5} \\
c & \frac{1}{5} & \frac{2}{5} & \frac{2}{5}
\end{array}
$$

$p(x_t \mid x_{t-1})$

$p(b \mid c) = \frac{2}{5}$

x_t x_{t-1}

$p(c \mid a) = \frac{1}{5}$

$P_\infty(a) = ?$

$P_\infty(b) = ?$

$P_\infty(c) = ?$

$$P_\infty(a) = P(a|a) \times P_\infty(a) + P(a|b) \times P_\infty(b) \times$$
$$P(a|c) \times P_\infty(c)$$

$$P_\infty(a) = \frac{2}{5} P_\infty(a) + \frac{1}{5} P_\infty(b) + \frac{1}{5} P_\infty(c)$$

$$\Rightarrow 5P_\infty(a) = 2P_\infty(a) + P_\infty(b) + P_\infty(c)$$

$$\Rightarrow 3P_\infty(a) - P_\infty(b) - P_\infty(c) = 0 \quad \ldots \text{(i)}$$

$$P_\infty(b) = P(b|a) \times P_\infty(a) + P(b|b) \times P_\infty(b)$$
$$+ P(b|c) \times P_\infty(c)$$

$$= \frac{2}{5} P_\infty(a) + \frac{3}{5} P_\infty(b) + \frac{2}{5}$$
$$P_\infty(c)$$

$\Rightarrow 5P_\infty(b) = 2P_\infty(a) + 3P_\infty(b) + 2P_\infty(c)$

$\Rightarrow 2P_\infty(a) - 2P_\infty(b) + 2P_\infty(c) = 0$

$\Rightarrow P_\infty(a) - P_\infty(b) + P_\infty(c) = 0 \quad \ldots \text{(ii)}$

$P_\infty(a) + P_\infty(b) + P_\infty(c) = 1 \quad \ldots \text{(iii)}$

Using calculator