UNIT-3

Knowledge
Intelligence
Reasoning
Knowledge Representation
Knowledge Based Agents
The Wumpus World.

- Syntax - Grammatical error (wrong inference)
- Semantic - Meaning (I watched book while reading TV)

Machine acts acc. to knowledge fed into it.

Logic 1. Propositional (T/F)
2. Predicate (∀)

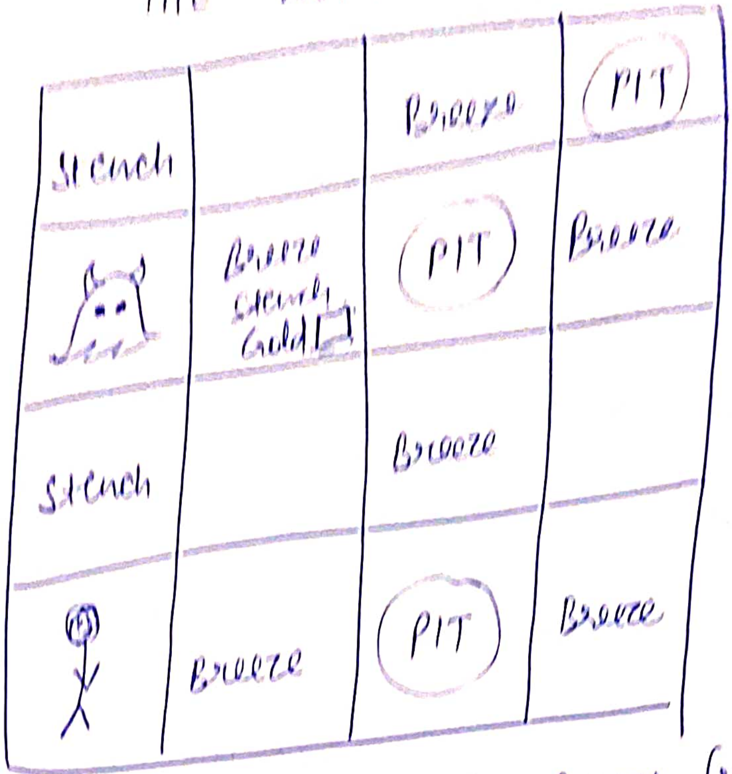
Rules → if
then

Semantics Net → Google Graph
→ meaning graph e.g. Cat has 4 legs
Tiger also has.

Frames → Slots & fillers
↓ ↓
obj attributes

Script Advancement of Frames
↳ given to agent so that it can take decisions.

THE WUMPUS WORLD



Intro :- Wumpus, Stench, Agent, Pit, Breeze, Gold, Arrow

Awards & Punishments :-

1. +100 Points if Agent comes back with Gold
2. -1 point on agents every action
3. -10 pts. if arrow is used
4. -200 pts. if agent dies

PEAS properties :-

Performance Measure :- (see above 4 pts.)

Environment :- Empty Rooms, Room with Wumpus, Stenchy Rooms, Breezy rooms, Rooms with Gold, Arrow.

Sensors :- Camera, odour sensors, Audio sensors

Effectors :- Motor to move R, L, U, D., Robotic arm to grab gold & robotic mechanism to shot arrow.

Predicate logic

1. Marcus was a man

   man (Marcus)

2. Marcus was a Pompeian

   Pompeian (Marcus)

3. All Pompeian were Romans

   $\forall x$: Pompeian $(x) \rightarrow$ Roman (x)

4. Caesar was a ruler

   ruler (caesar)

5. All Romans were either loyal to Caesar or hated him

   2 1 3

   $\forall x$: $\underbrace{\text{Roman}(x)}_{\text{if } x \text{ is roman}} \rightarrow$ loyal to $(x$, Caesar$)$

   $\vee$ hate $(x$, Caesar$)$

   2 1 3

6. Everyone is loyal to someone.

   $\forall x$, $\exists y$: loyal to (x, y)

   everyone someone 2 1 3

7. People only try to assassinate rulers they are not loyal to

   $\underbrace{\forall x}_{\text{People}}, \underbrace{\forall y}_{\text{ruler}}$: People(x) $\wedge$ ruler (y) $\wedge$ try assassinate $\underbrace{(x, y)}_{x \text{ try to ass } y}$

   $\underbrace{\neg}_{\text{who are}}$ loyal (x, y)

(8) Marcus tried to assassinate Caesar

try assassinate (Marcus, Caesar)

Prove : Marcus hated Caesar

hate (Marcus, Caesar)

Resolution in Predicate logic

(1) John likes all kinds of foods

PL ⇒ $\forall x : food(x) \rightarrow likes(John, x)$

Clause form $\neg food(x) \vee likes(John, x)$

if $a \rightarrow b \equiv \neg a \vee b$

(2) Apples are food

food (apple)

Clause form, same

FORWARD CHAINING & BACKWARD CHAINING

Q. Prove that Robert is a criminal.

FACTS :-

- It is crime for an American to sell weapons to the enemy of America.
- Country China is an enemy of America.
- China has some missiles.
- All the missiles were sold to China by Robert.
- Missile is a weapon
- Robert is American

FACTS TO FOL :-

⇒ American (x) ∧ Weapon (y) ∧ sell (x, y, z) ∧ enemy (z, America) ⇒ criminal(x)

⇒ Enemy (China, America)

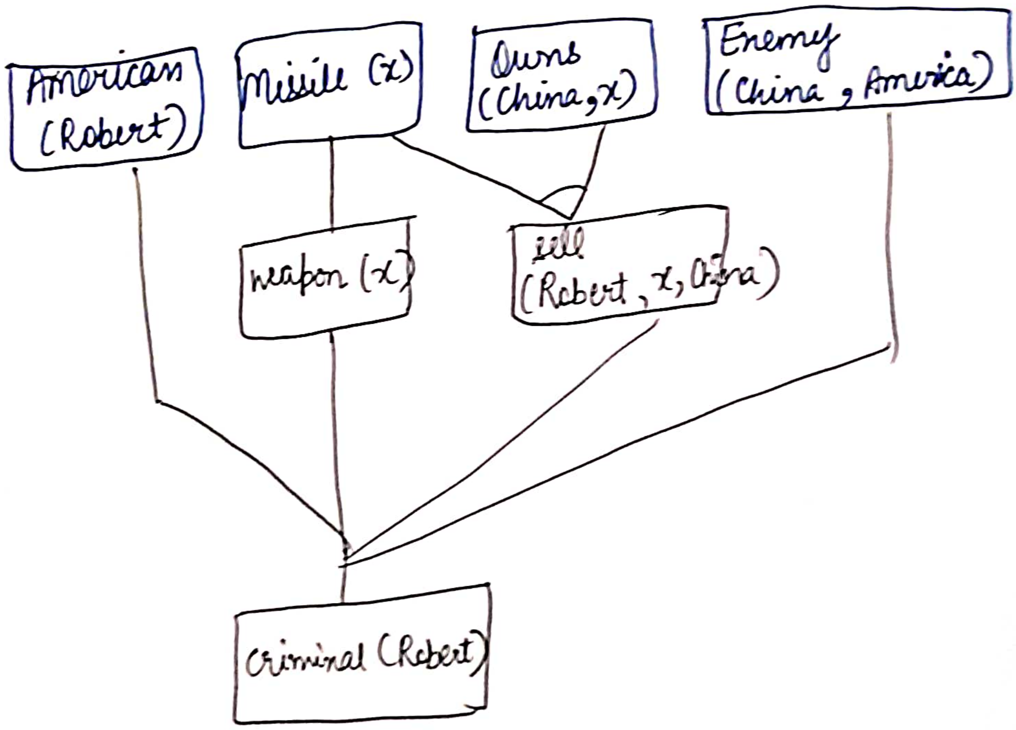
⇒ Owns (China, x)
Missile (x)

⇒ ∀x Missile (x) ∧ Owns(China, x) ⇒ sell (Robert, x, China)

⇒ missile (x) ⇒ weapon (x)

⇒ American (Robert)

Proof by Forward Chaining :- (Initial state to Goal State) – Data Driven

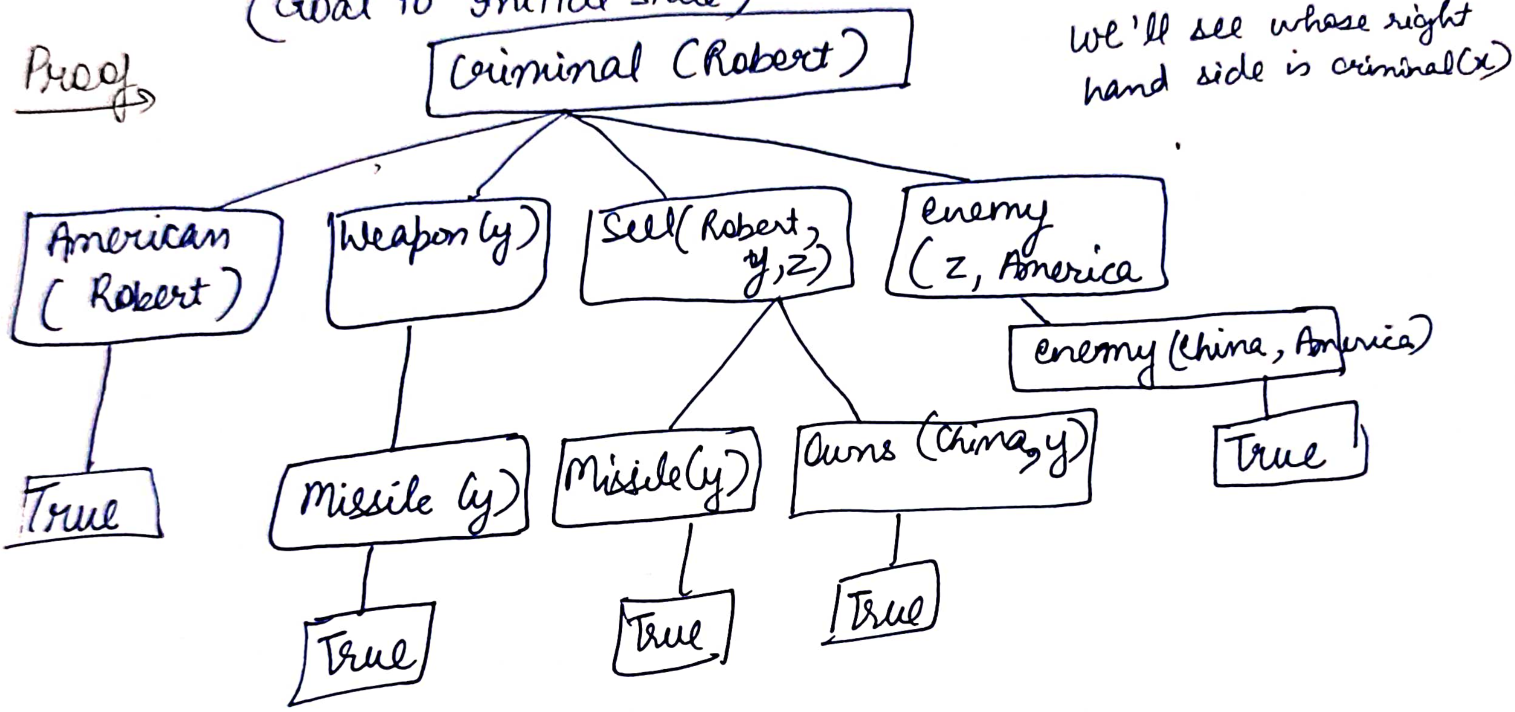


When based on the available data, a decision is taken then the process is called forward chaining.
It works from Initial state to goal state.

Backward Chaining

Decision driven / goal driven
(Goal to Initial State)

Proof ⇒



UNIFICATION & RESOLUTION

Unification stands for making of expressions to looks identical
It can be done with the process of substitution.

e.g. $p(x, f(y))$ —— ①

$p(a, f(g(z))$ —— ②

① and ② will be identical if x is replaced with a,

$$p(a, f(g(z))$$

and y is replaced with $g(z)$.

~~p (a/x~~

Unification $\rightarrow$ $\underbrace{[a/x,\ g(z)/y]}_{\text{substitution set}}$

Conditions for Unification :-

① Predicate symbol must be same. (p in both equations above)

② Number of arguments in both expressions must be identical. $\begin{bmatrix} p(\_,\_) \\ p(\_,\_,\_) \end{bmatrix}$ X unif. can't be performed because arg. no. is not same

③ If two similar variables present in same expⁿ, then unification fails. [x, a in above e.g.]

Algorithm for unification :-

Unify (A1, A2)

① If A1 or A2 is variable / constant
↳ if A1 and A2 are identical
return NIL

↳ Else if A1 occurs in A2
return FAIL

↳ Else return $\boxed{A2/A1}$ (A1 replaced by A2)

• Check for A2 in A1
↳ fail if A2 occurs in A1
↳ else return $\boxed{A1/A2}$ (A2 replaced by A1)

② If Predicate is not same

③ / if diff. no. of arguments
then FAIL

④ else SUBSTITUTE to NIL

⑤ LOOP

⑥ Return

D.Y.S.
Q Prime (11), Prime(y)
Substitute y with 11
Prime (11), Prime (11)
Successfully unified.

e.g. $\underbrace{Q(a, g(x,a), f(y))}_{A_1}$, $\underbrace{Q(a, g(f(b), a), x)}_{A_2}$

Substitute x with f(b) [f(b)/x]

Q (a, g (f(b), a), f(y)), Q (a, g(f(b), a), f(b))

Substitute (b/y) i.e. y is substituted with b

Q (a, g(f(b), a), f(b)), Q (a, g(f(b), a), f(b))

UNIFIED SUCCESSFULLY

Resolution

- To do or not to do
- Decide with the help of facts

Resolution is a single inference rule which can efficiently operate on the conjunctive normal form (CNF) or clausal form.
Resolution method is an inference rule which is used in both Propositional as well as First order Predicate Logic in dif. ways.
This method is basically used for proving the satisfiability of a sentence. In resolution method, <u>Proof of Refutation</u> (contradiction) technique is used to prove the given statement.

Q Use propositional logic and apply resolution method to prove that the goal is derivable from the given knowledge base.

Sol^n Lets construct propositions of the given sentences one by one :-

① Let, P : Humidity is high
Q : Sky is cloudy
It will be represented as P V Q

② Q : Sky is cloudy —— from ①
let R : It will rain
Representation :- Q?R

③ P : Humidity is high —— from ①
let S : It is hot
P?S

④ ¬S : It is not hot

Now, Apply Resolution method :-

In ②, Q ? R will be converted as $(\neg Q \vee R)$ — ④

In ③, P ? S " " " " $(\neg P \vee S)$ — ⑤

Negation of Goal $(\neg R)$: It will not rain.

$P \vee Q$ (given in ①) $\quad$ $\neg Q \vee R$ ④

$P \vee R$ $\quad$ $\neg P \vee S$ ⑤

$R \vee S$ $\quad$ $\neg S$

R $\quad$ $\neg R$

ϕ (Null)

After applying proof by Refutation (Contradiction) on the goal, the problem is solved, and it has terminated with a Null clause (ϕ). Hence, the goal is achieved. Thus, it is not raining.

<u>Rules</u> ① Convert the given axiom into clausal form i.e. disjunction form

② Apply and proof the given goal using negation rule

③ Use those literals which are needed to prove.

④ Solve the clauses together and achieve the goal.

Knowledge Repⁿ Using Semantic Nets :-

Graphical notation for representing knowledge in interconnected nodes pattern is Semantic Net. It is popular in AI & NLP since it represents knowledge or support reasoning. It can be an alternative of predicate logic sometimes.

- Nodes represent objects
  - circle (physical)
  - ellipse (concept)
  - rectangle (situation)
- Arcs represent relation between objects] links / arrows
- Link labels] specify relationships
- Also known as Associative Nets] Nodes are associated.

Components

1. Lexical Nodes, links, labels
2. Structural Directed link or nodes
3. Semantic Defⁿ related to links / nodes. Based on facts
4. Procedural
   - Constructor] creation of new links
   - Destructor] Removal of links.

We can store our knowledge in the form of a graph, with nodes representing objects in the world, and arrow representing relationships b/w those objects.

Eg.

Tom is a cat

Tom caught a bird

Tom is owned by John

Tom is ginger in colour

Cats like cream.

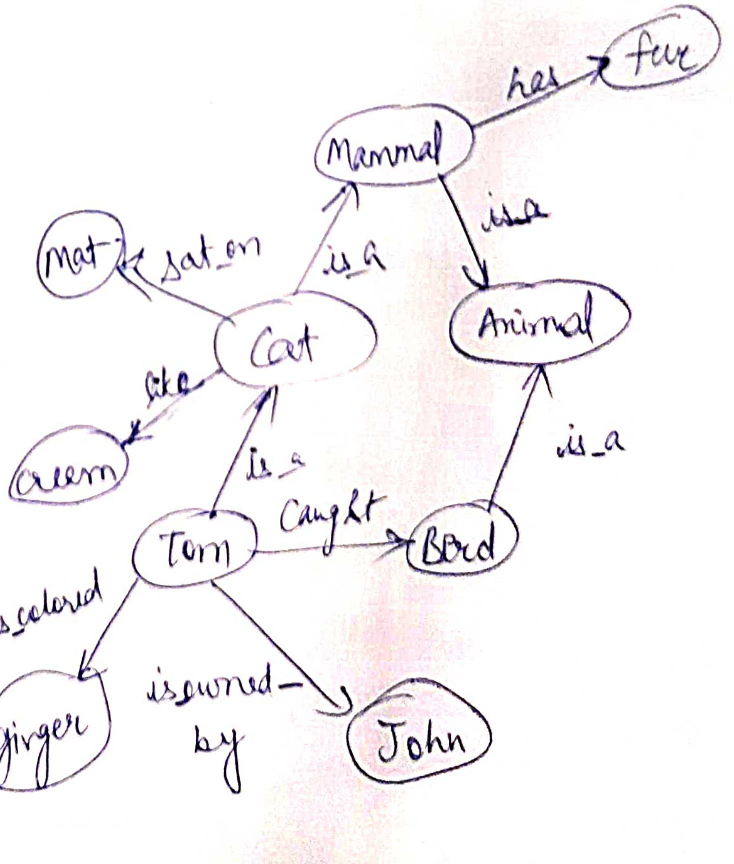
The cat sat on the mat

A cat is a mammal

A bird is an animal

All mammals are animals.

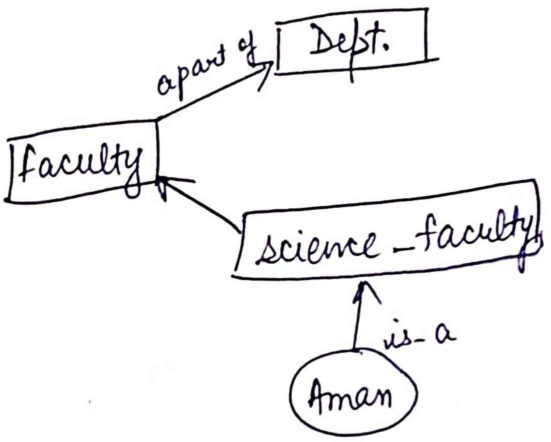
Mammals have fur.



Knowledge Repⁿ using FRAMES :-

Frame is a collection of attributes or slots & associated values that describe some real world entity. It uses the data structure (records) to represent the knowledge represented in semantic N/W.

Each frame represents the node in semantic N/W as a class or an instance & each relation as slot.



Attributes attached with each slot :-

1. Instance :- Relates slot with class
2. Definition :- Slot def^n/value
3. Default :- slot default value
4. Domain :- Slot elements domain
5. Range :- Specifies class of which elements
6. Range constraint :- logical exp^n $6 < i < 10$
7. To compute :- Value of the slot is to be computed.
8. Single valued :- $func^n$ returns single value
9. Inverse :- slot inverse used in reasoning.
10. Transfers-through :- Inheritance.

Reasoning actions that can be performed using Frames :-

1. Relating the definition :- is a, inverse links] Propagation of def^n to relate all info.
2. Inheritance :- Inherited all values incl. default values of the slot

③ Legality of Value :- Checks the legality of slot value

Range constraint

④ Consistency check Verifying slot value consistency before adding to the frame.

Domain

⑤ Maintaining consistency :- When one slot is updated its inverse should also be updated.

⑥ Computation of a Value of Slot :- to-compute transfers through

Frame is a record like structure which consists of a collection of attributes and its values to describe an entity in the world. Frames are the AI data structure which divides knowledge into substructures by representing stereotypical situations. It consists of a collection of slots & slot-types values. These slots may be of any types & sizes.
Slot have names and values which are called facets.
Facets are features of frames which enable us to put constraints on the frames
e.g. IF-NEEDED facts are called when data of any particular slot is needed. A frame is also known as slot - filter knowledge repⁿ in AI.

Applications : NLP, Machine Vision

Slots	Filters
Title	AI
Genre	Computer Science
Author	Peter Norvig
Edition	Third Edition
Year	1996
Page	1152

Advantages

① Frame knowledge representation makes the programming easier by grouping the related data

② The frame repn is comparably flexible & used by many appns in AI.

③ It is easy to add slots for new attr. & relations.

④ It is easy to include default data & to search for missing values.

⑤ Frame repn is easy to visualize.

Disadvantages

① In frame system inference mechanism is not be easily processed.

② Inference mechanism cannot be smoothly proceeded by frame representation.

③ Frame representation has a much generalized approach.

Uncertainty (Knowledge in AI)

Uncertainty :- Lack of exact infoⁿ or knowledge that helps us to find correct conclusion. 1

Sources of Uncertainty :-

1. Uncertain inputs
   - → Missing data
   - → Noisy data / irrelevant details
2. Uncertain knowledge
   - → Multiple causes lead to multiple effects
   - → Incorrect or incomplete knowledge of casualty in the domain.
3. Uncertain outputs
   - → Abduction, induction are uncertain.
   - → Default reasoning
   - → Incomplete deduction, inference

Uncertainty may be caused by problems with data such as :-

1. missing / unavailable data
2. Unreliable / unambiguous data
3. Imprecise / Inconsistent repⁿ of data
4. Guess based

Reasoning under Uncertainty :-

Probabilistic Reasoning :- It is a way of knowledge representation where we apply the concept of probability to indicate the uncertainty in knowledge.

We combine probability theory with logic to handle uncertainty :-

2 ways to solve problems :-

1. Baye's Rule
2. Baysian statistics / Networks

Bayes Theorem

Acc. to condⁿ probability,

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

prob. of event A given that B has already occurred

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

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

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

Bayes's Theorem

Prior P.

Compute entity of Bayes Th.

hypothesis

evidence/ data

Marginal P.

Posterior Prob.

Likelihood P. (prob. of evidence)

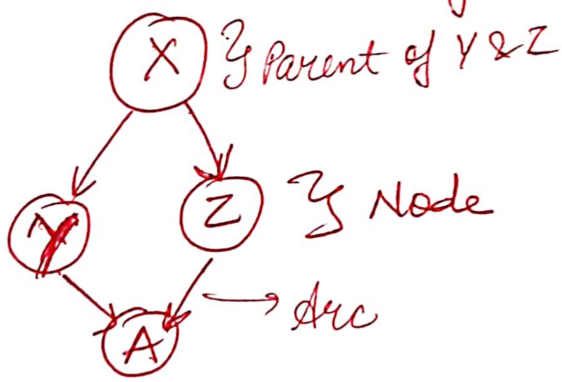
$$P(King|Face) = \frac{P(Face|King) \cdot P(King)}{P(Face)}$$

$$= \frac{1 \cdot \frac{4}{52}}{\frac{12}{52}}$$

$$= \frac{1 \cdot 1/13}{3/13} = \frac{1}{3} \text{ Ans.}$$

Bayesian Belief Network in AI :- It defines probabilistic (2)
independencies and dependencies among the variables in the N/W.

- It is a probabilistic graphical model which represents a set of variables & their conditional dependencies using a directed acyclic graph.
- Built from Probability distribution
- Consists of (1) DAG
  (2) Table of conditional Probabilities
- <u>Node</u> Corresponds to a random variable → Continuous / → Discrete
- <u>Arc</u> (Directed arrows) : Represent casual relationship or conditional Probabilities among random variables.



A is indep. of X

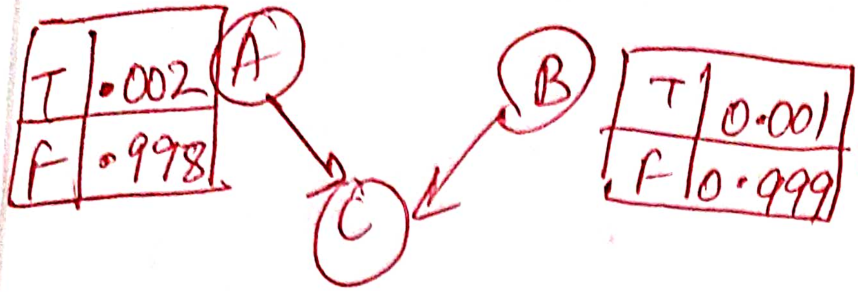
So, Bayesian Belief network shows dependencies & independencies

Here, there is a direct influence of X on Y

We will also create table :- Cond. Prob. of all nodes w.r.t. of Parent nodes

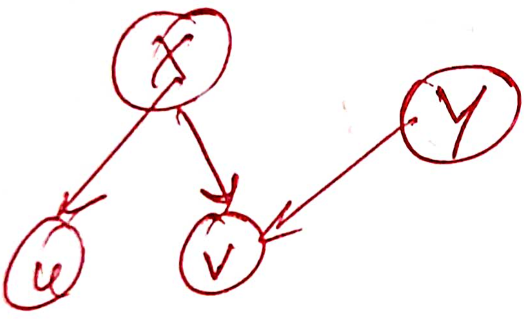
To propagate belief in Bayesian N/W, initial "DAG" is converted into undirected in which the arcs can be used to transmit probabilities in direction of evidence. e.g. →

Shivani Rohilla, MNIT Jaipur



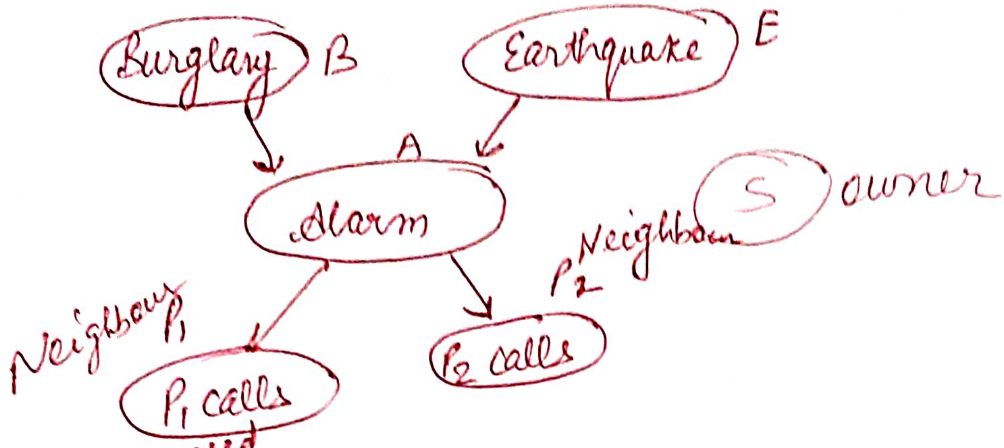
Cond n prob.

A	B	P(C=T)	P(C=F)
T	T	–	–
T	F	–	–
F	T	–	–
F	F	–	–



for correctness, there should not be a cycle

Example of Bayesian Belief Network :-



Burglary occurred $P(B=T) = 0.001$

Not occurred $P(B=F) = 0.999$

$P(E=T) = 0.002$

$P(E=F) = 0.998$

Observed values

B	E	P(A=T)	P(A=F)
T	T	0.95	0.05
T	F	0.99	0.06
F	T	0.29	0.71
F	F	0.001	0.999

A	$P(P_1=T)$	$P(P_1=F)$
T	0.90	0.10
F	0.05	0.95

A	$P(P_2=T)$	$P(P_2=F)$
T	0.70	0.30
F	0.01	0.99

True event — false event

• $P(P_1, P_2, A, \sim B, \sim E)$

Prob. of A w.r.t. no burs & no earthquake

$= P(P_1/A) \cdot P(P_2/A) \cdot P(A/\sim B, \sim E) \cdot P(\sim B) \cdot P(\sim E)$

$= 0.90 \times 0.70 \times 0.001 \times 0.999 \times 0.998$

$= 0.00062$

Shivani Rohilla, MNIT Jaipur

A

Probabilistic Reasoning
Probabilistic Reasoning over time
Forward & Backward Reasoning
Other uncertain techniques - Data Mining
Fuzzy logic
Dempster - shafer theory

Probability

$$0 \leq P(A) \leq 1 \rightarrow \text{certainty}$$

uncertainty

$$P(\sim A) + P(A) = 1$$

Conditional Probability :-

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

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

Probabilistic reasoning is a way of knowledge representation where we apply the concept of probability to indicate the uncertainty in knowledge. Here, we combine probability theory with logic to handle the uncertainty.

The sentences where we are not sure about their occurrence or certainty, we use Probabilistic reasoning like :-

e.g. Behaviour of someone in some situations.

Probabilistic reasoning over-time In the real world, we observe a lot of phenomena which change over time. Let us say we have a set of random variables which are changing over time. Some of them are observable and some are not. We need to model this uncertainty somehow by studying the past and the present & making useful predictions about the future.

Random variables :- A variable whose values depend on outcomes of a random phenomenon.

A random variable's possible values might represent the possible outcomes of a yet-to-be-performed experiment, or the possible outcomes of a past experiment whose already existing value is uncertain. They may also conceptually represent either the results of an "objectively" random process (such as rolling a die) or "subjective" randomness that results from incomplete knowledge of a quantity

Stochastic Process :- A collection of random variables $X = \{X_t : t \in T\}$ defined on a common probability space, taking values in a common set S (the state space), and indexed by a set T, often either N or $[0, \infty]$ & thought of as time (discrete or continuous respectively)

Time and Uncertainty A time interval contains random variables, some observable & some not :-

X_t : denote the set of state variables at time t, unobservable

E_t : denote the set of observable evidence variables

e.g. Security guard with umbrella. For each day t, the set E_t contains a single evidence variable U_t (umbrella) and the set X_t

contains a single state variable R_t (Rain).
Evidence starts arriving at $t = 1$. Hence our umbrella world is represented by state variables $R_0, R_1, \dots$ and evidence variables $U_1, U_2, \dots$

Markov Process

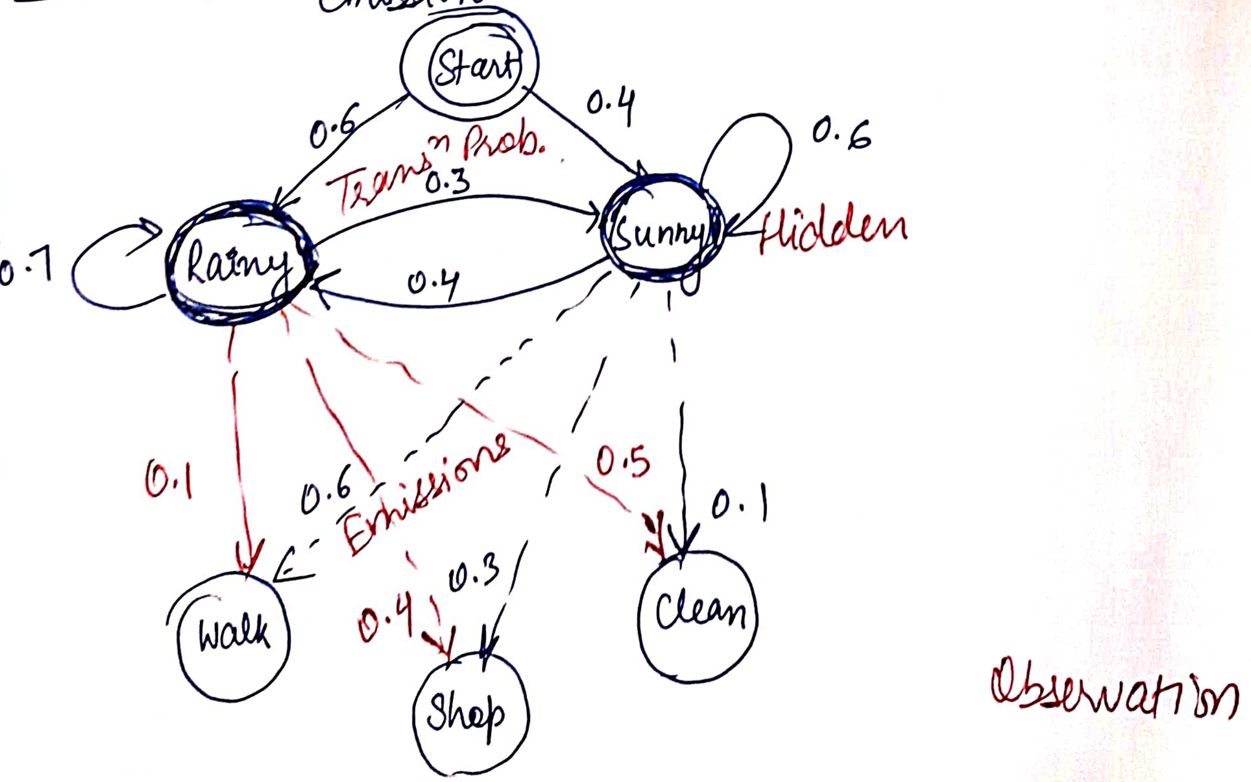
Hidden Markov Models generate models.

Markov Property :- ① It is memory less. It considers only one state. The future prediction relies on only present state (neither past nor future).

$\boxed{\text{HMM}}$ is one of the most popular models where it is a statistical markov model in which the system being modelled is similar to be Markov process with unobserved (i.e. hidden) states.

Applications Reinforcement learning, speech recognition, gesture recognition, handwriting recognition, partial discharges, bioinformatics, stock price prediction.

Terminologies Hidden, observation, transition, probability, emissions.



Shivani Rohilla, MNIT, Jaipur

Property

① Memory less

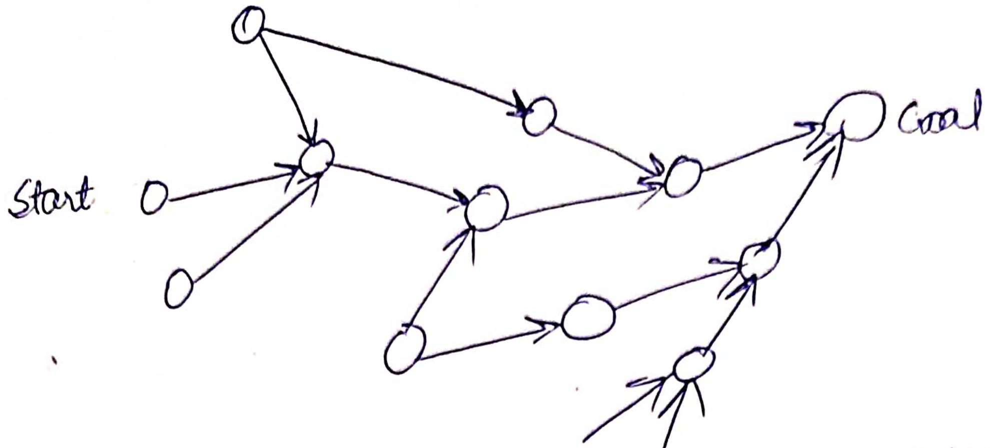
② Its future & Past are independent i.e. prediction depends on current state.

Goal

The goal is to make a sequence of decisions, where, a particular decision may be influenced by earlier decision.

Forward and Backward Reasoning

Forward Reasoning :-



In forward reasoning, the left hand side of the rules are matched against the current state and right sides are used to generate the new state, until the goal is reached.

Reason forward from the initial state :-

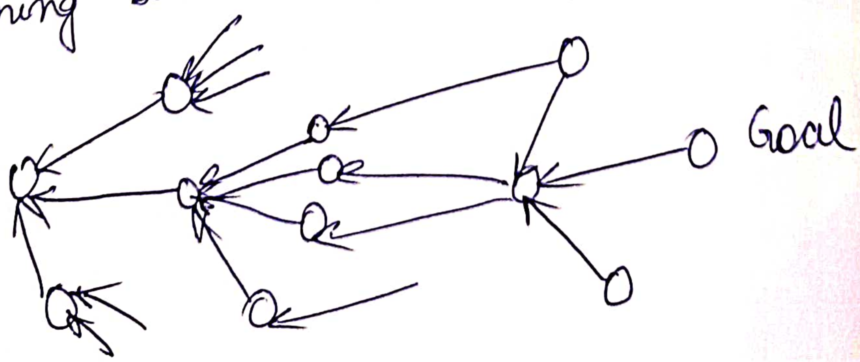
Step 1 Begin building a tree of move sequences by starting with the initial configuration at the root of the tree.

Step 2 Generate the next level of tree by finding all rules whose left hand side matches against the root node. The RHS is used to create new configurations.

Step 3 Generate the next level by considering the nodes in the previous level and applying it to all rules whose LHS match.

Backward Reasoning :

It is Reasoning backward from the goal states :-



In Backward reasoning, RHS of the rules are matched against the current states & left sides are used to generate the new state.

Step 1 Begin building a tree of move sequences by starting with the goal node config. at the root of the tree.

Step 2 Generate the next level of the tree by finding all rules whose RHS matches against the root node. The LHS is used to create new configurations.

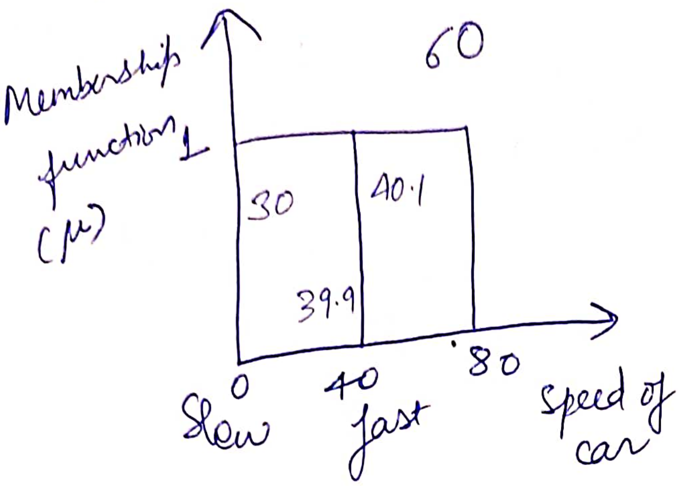
Step3 Generate the next level by considering the nodes in the prev. level and applying it to all rules whose RHS match it. Then use corresponding LHS to generate the new nodes. Continue until the node that matches the initial state is generated.

Fuzzy logic

In case of continuous values, (not discrete), we need fuzzy logic. It represents uncertainty.

- It is represented by degree

[Boolean will represent only in 0 or 1



Fast Transition

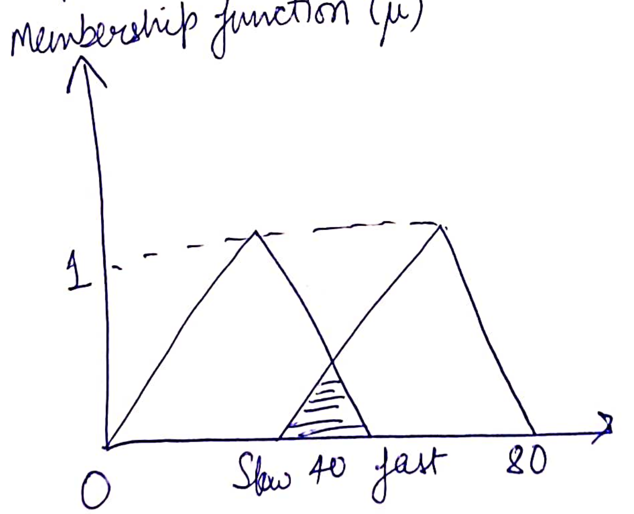
example of car

Accelerator

Brake

Knowledge Based Agent

Fuzzy logic provides flexibility [Fuzzy ~~takes~~ tells degree



speed of car

More efficient way (Fuzzy Logic)

"Check the degree of fastness" by this formule

$$\begin{cases} 0, & \text{if speed}(x) \leq 40 \\ \dfrac{\text{Speed}(x) - 40}{10}, & \text{if } 40 < \text{Speed}(x) < 50 \\ 1, & \text{if speed}(x) \geq 50 \end{cases}$$

Data Mining is a branch of A.I. A.I. is supposed to improve upon the functions of the human brain, data mining does this by identifying patterns in the data that wouldn't otherwise be seen. Data Mining is a process of using computers to identify hidden patterns in, and build models from large data sets. In order to utilize data mining the data must be clean & consistent & the data must reflect current & future trends.

e.g. Customer segmentation, Marketing & Promotion, Market Basket Analysis, Fraud Detection, Financial ~~Modelling~~ etc.

Advantages :- (Applications of Fuzzy logic)

- Control systems
  - (Kitchen appliances
  - Car control systems)
  - Aerospace vehicle control systems
- Image Processing
- Decision - Making
- Route - Planning
- Testing for safety - critical systems.

e.g. Sam is tall (5ft)

fairly if (speed = 0)
slow
else if (speed = 1)
fast

tall
fairly tall }
short

e.g. speed of car

0 — 0.25 — slowest

0.25 — 0.50 — slow

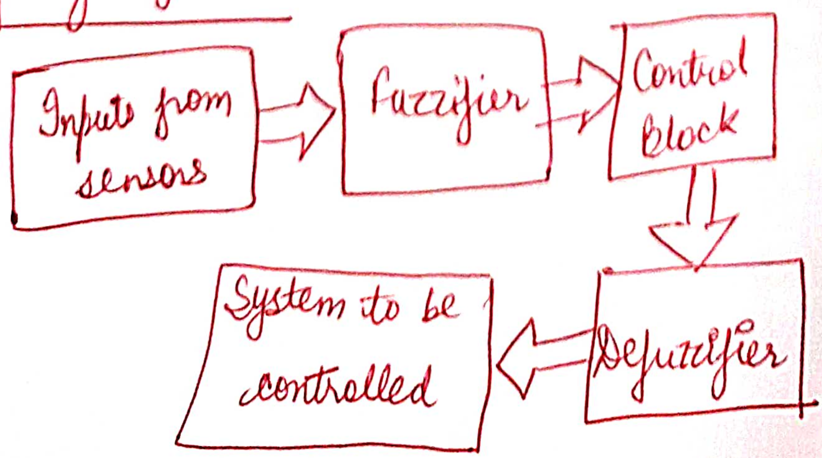
0.50 — 0.75 — fast

0.75 — 1.0 — fastest

e.g. Refrigerator

The fuzzy logic controller is used to maintain the inner temperature in spite of the environmental variations such as the outer temperature change or the volume change of the stored foods of the refrigerator.

Fuzzy logic System:-



Dempster-Shafer Theory / DST / Evidence theory

~~Selection~~ Framework for reasoning.

It is uncertain.

It can have a measure of Plausability i.e. frequency of correctness that it is true.

It fails when we work for hierarchical models.

Situation A housewife finds a stain on her husband's shirt. Her husband denies that it is a lipstick stain.

She knows that applying a little ethanol (C_2H_5-OH) to a stain & softening it with washing up almost always removes lipstick stain. She also knows that this method hardly ever works on other stains.

She applies this method to the shirt & finds that it removes the stain. As ~~too~~ a result, she concludes that it was almost certainly a lipstick stain.

$$P(X/Y) = \frac{P(X) * P(Y/X)}{P(Y)}$$

P(works / lipstick) = 0.9 (suppose)

P (works) = 0.3 (suppose)

$$P(\text{lipstick} \mid \text{works}) = \frac{P(\text{works}) * P(\text{works} / \text{lipstick})}{P(\text{lipstick})}$$

Also, P(lipstick) = 0.25 (suppose)

$$\Rightarrow \frac{0.3 \times 0.9}{0.25} = 0.75$$

$P(\text{late} \mid \text{lipstick}) = 0.6$

$P(\text{late}) = 0.5$

$$P(\text{lipstick} \mid \text{late}) = \frac{P(\text{lipstick}) \times P(\text{late} \mid \text{lipstick})}{P(\text{late})}$$

$$= \frac{0.75 \times ? \ldots}{\ldots}$$

may be = 0.6

0.9

0.5

We may have many pieces of evidences & many conditional probabilities.

So, to overcome this we have Dempster Shafer Theory.

Arthure P. Dempster 1967

Glenn Shafer 1976

So, Narrowing down the set

{strawberry stain, lipstick stain, blue ~~ink~~ stain}

= 0.8

We can consider some points & draw some results through reasoning without ~~un~~certainty.

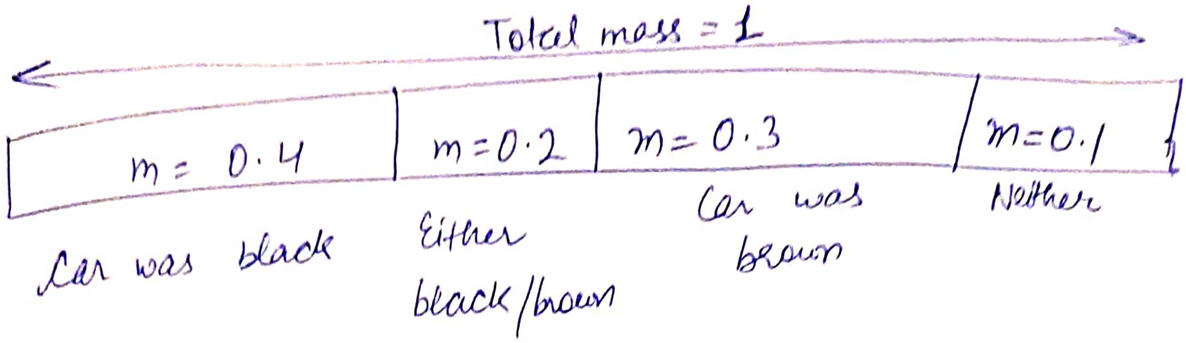
When we go through this theory we get some Measure of plausability.

We have to fed all these info in knowledge Based Agents.

This is knowledge Repⁿ & Reasoning

Belief and Plausability (Another example)

Witness : " I am fairly sure that the car was either brown or black. Probably black, though it could have been brown. I could be wrong though."



Belief

Bel (Black) = 0.4

Bel (Brown) = 0.3

Bel (not Black) = 0.3 + 0.1 = 0.4

Bel (Not Brown) = 0.5

Plausability : Pl (Black) = 0.4 + 0.2

= 0.6

= 1 − Bel (Not Black)

Pl (Brown) = 0.2 + 0.3

= 0.5

= 1 − Bel (Not Brown)