PBCA604:ARTIFICIAL INTELLIGENCE AND EXPERT SYSTEMS

UNIT-III TOPIC COVERED

1. Semantics Nets
2. Frames
3. Conceptual Dependency
4. Scripts
5. Monotonic Reasoning
6. Default Reasoning
7. Statistical Reasoning
8. Bayes' theorem
9. Certainty factors
10. Dempster-Shafer Theory
11. Fuzzy Logic

Semantics Nets: A Semantic Net is a formal graphic language representing facts about entities in some world about which we wish to reason. Semantic nets are a simple way of representing the relationships between entities and concepts.

- A semantic network
  - a classic AI representation technique used for propositional information
  - a propositional net
- A proposition a statement that is either true or false
- A semantic net a labeled, directed graph
- The structure of a semantic net is shown graphically in terms of nodes and the arcs connecting them.
  - Nodes are sometimes referred to as objects
  - arcs as links or edges
  - The links are used to express relationships
  - Nodes are to represent physical objects, concepts, or situation

Representation in a Semantic Net
The physical attributes of a person can be represented as in Figure

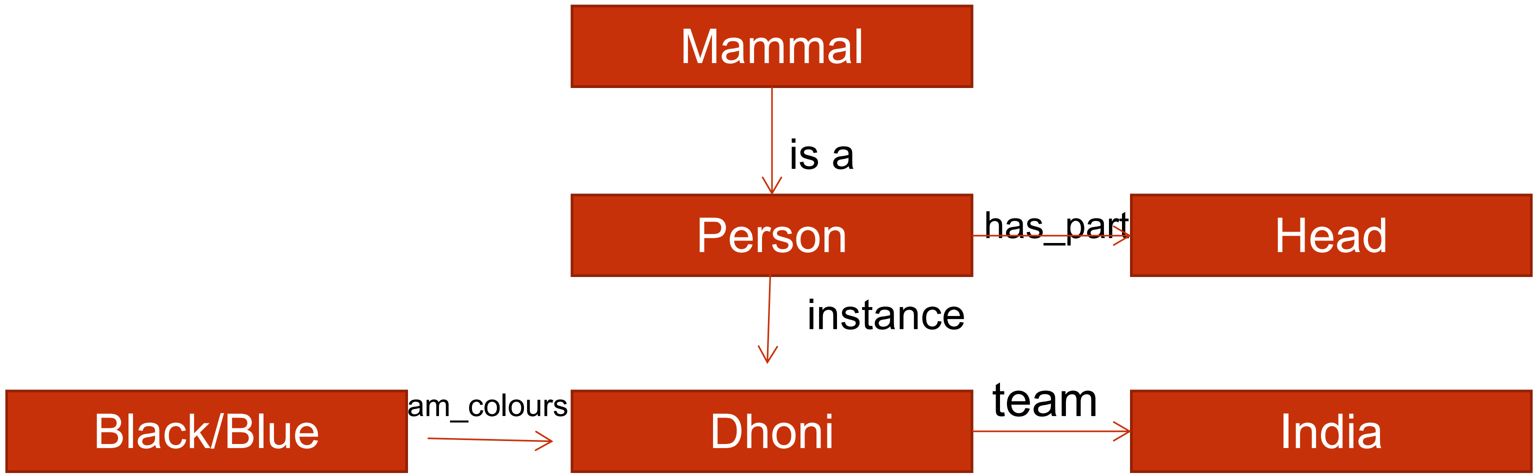


Fig. A Semantic Network

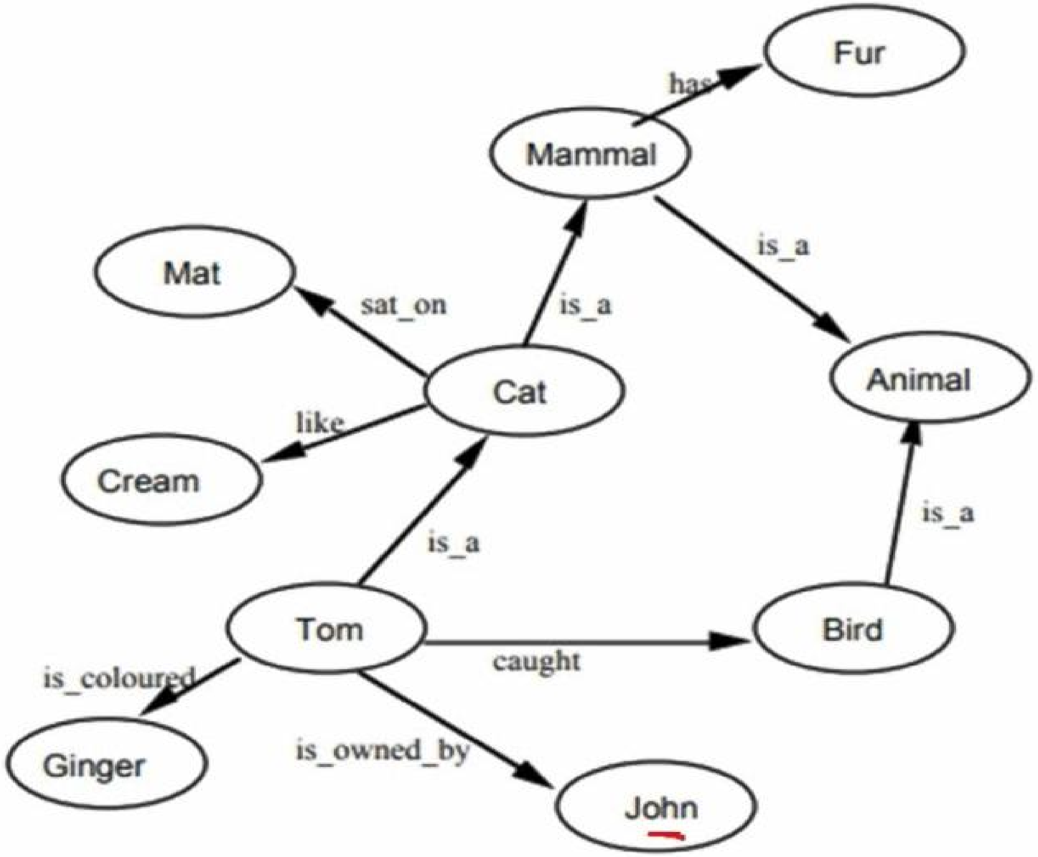
These values can also be represented in logic as: isa (person, mammal), instance (Dhoni, person), team(Dhoni, India).

Semantic Network examples

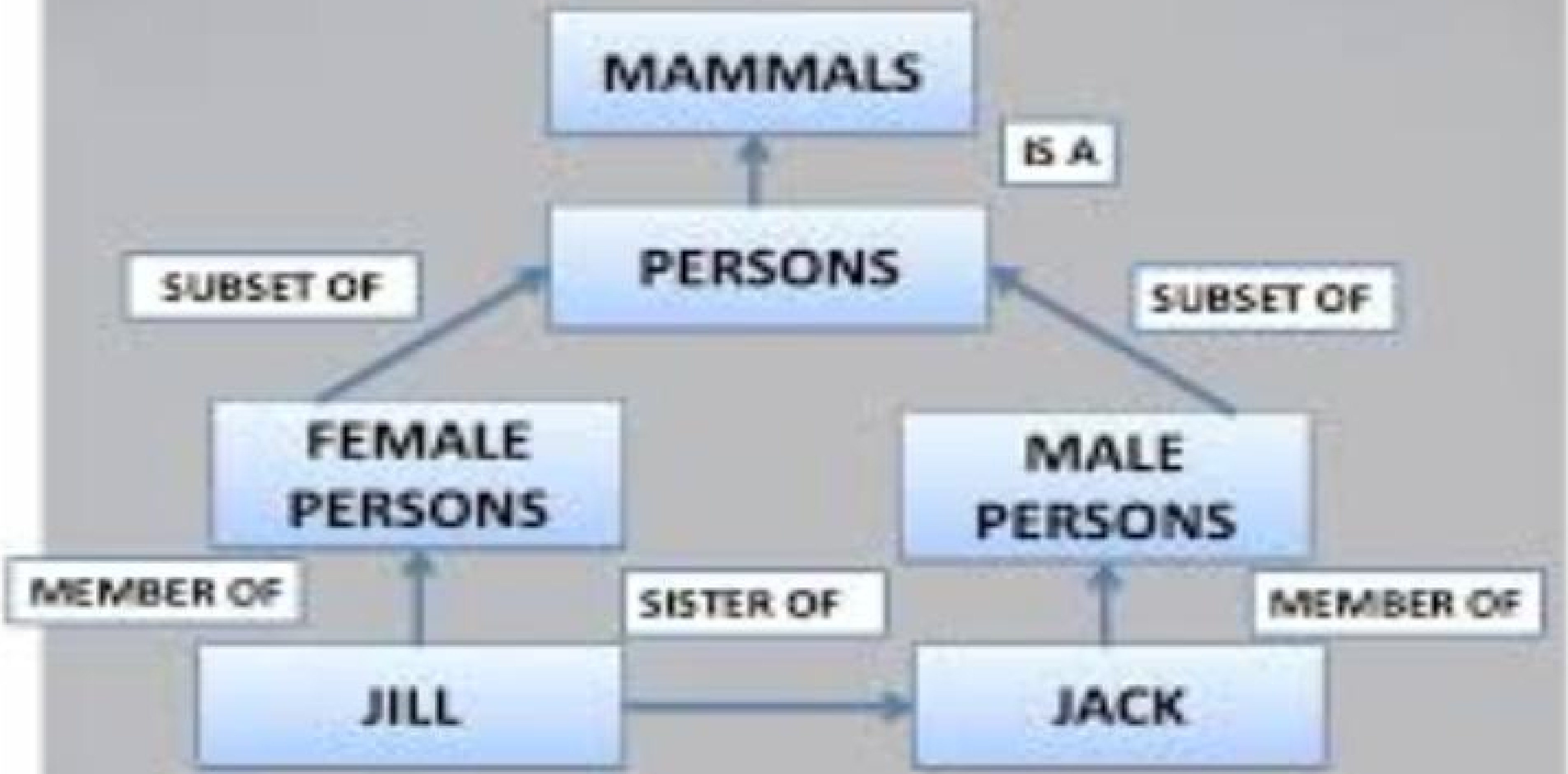
Semantic Networks:

Example :

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.



Example of semantic network



FRAMES

- Frames can also be regarded as an extension to Semantic nets.
- Indeed it is not clear where the distinction between a semantic net and a frame ends.
- Initially, semantic nets were used to represent labelled connections between objects. As tasks became more complex, the representation needs to be more structured. The more structured the system, the more beneficial to use frames.
- A frame is a collection of attributes or slots and associated values that describe some real-world entity.
- Frames on their own are not particularly helpful but frame systems are a powerful way of encoding information to support reasoning.

Set theory provides a good basis for understanding frame systems.

- Each frame represents:
  - a class (set), or
  - an instance (an element of a class)
- Instead of properties a frame has slots. A slot is like a property, but can contain more kinds of information, sometimes called facts of the slot:
  - The value of the slot
  - A procedure that can be run to compute the value (and if needed, procedure)
  - Procedures to be run when a value is put into the slot or removed
  - Data type information constraints on possible slot fillers
  - Documentation

Frames can inherit slots from parent frames. For example, man might inherit properties from Ape or Mammal (A parent class of man).

Properties

- Frames implement semantic networks.
- They add procedural attachment.
- A frame has slots and slots have values.
- A frame may be generic, i.e. it describes a class of objects.
- A frame may be an instance, i.e. it describes a particular object.
- Frames can inherit properties from generic frames.

We could represent some knowledge about elephants in frames as follows:

- Mammal
  - subclass: Animal
  - warm_blooded: yes
- Elephant
  - subclass: Mammal
  - * colour: grey
  - * size: large
- Clyde
  - instance: Elephant
  - colour: pink
  - owner: Fred
- Nellie:
  - instance: Elephant
  - size: small

- A particular frame (such as Elephant) has a number of attributes or slots such as colour and size where these slots may be filled with particular values, such as grey.
- We have used a ‘*’ to indicate those attributes that are only true of a typical member of the class, and not necessarily every member. Most frame systems will let you distinguish between typical attribute values and definite values that must be true.
- In the above frame system, we would be able to infer that Nellie was small, grey and warm blooded.
- Clyde is large, pink and warm blooded and owned by Fred.

Advantages of frames

1. A frame collects information about an object in a single place in an organized fashion.
2. By relating slots to other kinds of frames, a frame can represent typical structures involving an object; these can be very important for reasoning based on limited information.
3. Frames provide a way of associating knowledge with objects.
4. Frames may be a relatively efficient way of implementing AI applications.
5. Frames allow data that are stored and computed to be treated in a uniform manner. For example, class, a computed grade or marks of a student might be stored.

CONCEPTUAL DEPENDENCY (CD)

- Conceptual Dependency originally developed to represent knowledge acquired from natural language input.
- The goals of this theory are:

- To help in the drawing of inference from sentences
- To be independent of the words used in the original input
- That is to say: For any 2 (or more) sentences that are identical in meaning there should be only one representation of that meaning.

CD provides:

- a structure into which nodes representing information can be placed
- a specific set of primitives
- a given level of granularity

Sentences are represented as a series of diagrams depicting actions using both abstract and real physical situations.

- The agent and the objects are represented.
- The actions are built up from a set of primitive acts which can be modified by tense.

Examples of Primitive Acts are:

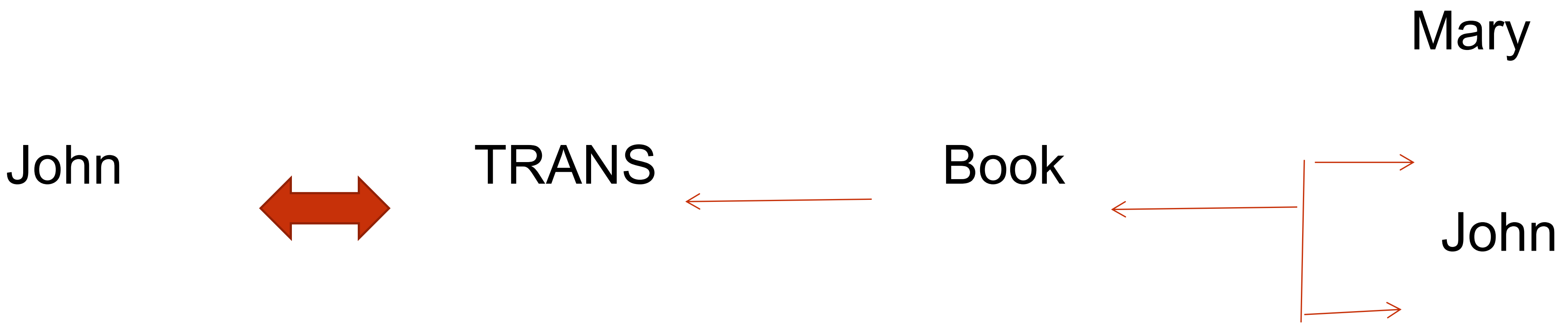
- ATRANS- Transfer of an abstract relationship. e.g., give.
- PTRANS- Transfer of the physical location of an object. e.g., go.
- PROPEL- Application of a physical force to an object. e.g., push.
- MTRANS- Transfer of mental information. e.g., tell.
- MBUILD - Construct new information from old. e.g., decide.
- SPEAK - Utter a sound. e.g., say.
- ATTEND - Focus a sense on a stimulus. e.g., listen, watch.
- MOVE - Movement of a body part by owner. e.g., punch, kick.
- GRASP - Actor grasping an object. e.g., clutch.
- INGEST - Actor ingesting an object. e.g., eat.
- EXPEL - Actor getting rid of an object from the body.

Six primitive conceptual categories provide building blocks which are the set of allowable dependencies in the concepts in a sentence:

ACTs	Actions
PPs	Object (Picture Procedure)
AAs	Modifiers of actions (Action aiders)
PAs	Modifiers of objects, i.e. PPs (picture aiders)
AA	Action aiders are properties or attributes of primitive Actions
PP	Picture procedures are actors or physical objects that perform different acts or producers.

How do we connect these things together?

- Consider the example: John gives Mary a book.



Arrows indicate the direction of dependency. Letters above indicate certain

relationships:

o object.

R recipient-donor.

I instrument e.g. eats with a spoon. D destination e.g. going home.

- Double arrows (⇔) indicate two-way links between the actor (PP) and action (ACT).
- The actions are built from the set of primitive acts (see above).

  These can be modified by tense etc.
- Primitives (ATRANS, PTRANS, PROPEL) indicates transfer of possession.
- O indicates object relation.
- R indicates recipient relation.

	Examples of their use	English version
1. PP⇔ACT	John⇔PTRANS	John ran
2. PP⇔PA	John⇔height (> average>)	John is tall
3. PP ↑ PA	Boy ↑ Nice	A nice boy

where, ↑ - arrow indicates direction of dependency

⇔ - double arrow indicates two way link between actor and action

⇔ - indicates the possession by a person.

Advantages of Conceptual Dependency

- Using these primitives involves fewer inference rules.
- So, Many inference rules already represented in CD structure.
- Moreover, The holes in the initial structure help to focus on the points still to established.

Disadvantages of Conceptual Dependency

- ✓ Knowledge must decompose into fairly low-level primitives.
- ✓ Impossible or difficult to find the correct set of primitives.
- ✓ Also, A lot of inference may still require.
- ✓ Representations can be complex even for relatively simple actions.

Scripts

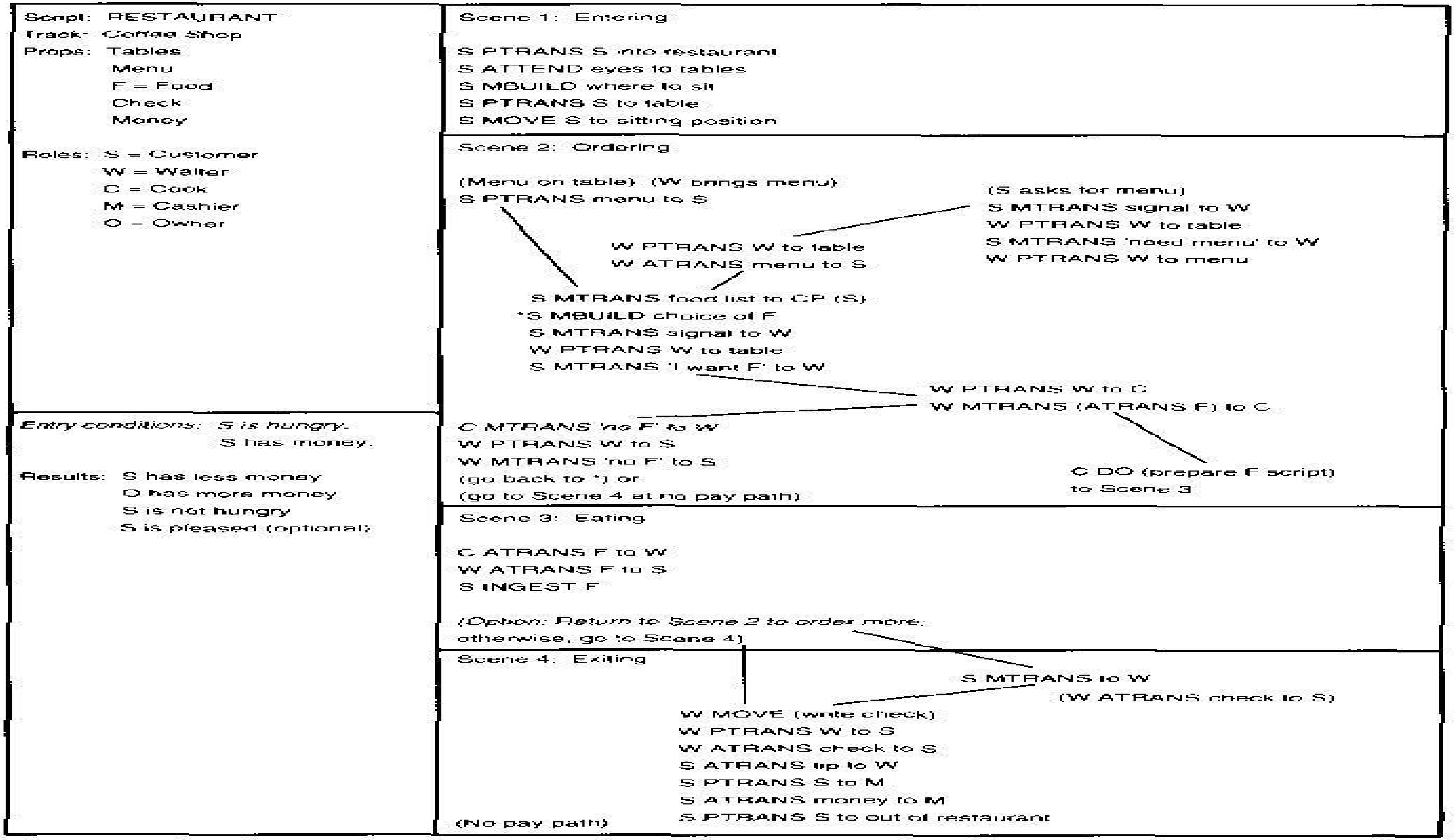
- A script is a structure that prescribes a set of circumstances which could be expected to follow on from one another.
- It is similar to a thought sequence or a chain of situations which could be anticipated.
- It could be considered to consist of a number of slots or frames but with more specialized roles.
- Scripts are beneficial because:
  - Events tend to occur in known runs or patterns.
  - Causal relationships between events exist.
  - Entry conditions exist which allow an event to take place
  - Prerequisites exist for events taking place. E.g. when a student progresses through a degree scheme or when a purchaser buys a house.

Scripts Components

- Each script contains the following main components.

Entry Conditions:	Must be satisfied before events in the script can occur.
Results:	Conditions that will be true after events in script occur.
Props:	Slots representing objects involved in the events.
Roles:	Persons involved in the events.
Track:	the Specific variation on the more general pattern in the script. Different tracks may share many components of the same script but not all.
Scenes:	The sequence of events that occur. Events represented in conceptual dependency form.

A Restaurant Script Example



- Advantages and Disadvantages of Script
- Advantages
  - Capable of predicting implicit events
  - Single coherent interpretation may be build up from a collection of observations.
- Disadvantage
  - More specific (inflexible) and less general than frames.
  - Not suitable to represent all kinds of knowledge.

Logical Reasoning

- Logic is a language for reasoning. It is a collection of rules called Logic arguments, we use when doing logical reasoning.
- Logic reasoning is the process of drawing conclusions from premises
- using rules of inference.
- The study of logic is divided into formal and informal logic.
- The formal logic is sometimes called symbolic logic.
- Symbolic logic is the study of symbolic abstractions (construct) that
- capture the formal features of logical inference by a formal system.
- Formal system consists of two components, a formal language plus a
- set of inference rules. The formal system has axioms.
- Axiom is a sentence that is always true within the system.
- Sentences are derived using the system's axioms and rules of derivation
- are called theorems.

Monotonic Logic

Formal logic is a set of rules for making deductions that seem self evident. A Mathematical logic formalizes such deductions with rules precise enough to program a computer to decide if an argument is valid, representing objects and relationships symbolically.

- Examples
- Predicate logic and the inferences we perform on it.
- All humans are mortal. Socrates is a human.
- Therefore Socrates is mortal.
- In monotonic reasoning if we enlarge at set of axioms we cannot retract any existing assertions or axioms.

Most formal logics have a monotonic consequence relation, meaning that adding a formula to a theory never produces a reduction of its set of consequences. In other words, a logic is monotonic if the truth of a proposition does not change when new information (axioms) are added. The traditional logic is monotonic.

In mid 1970s, Marvin Minsky and John McCarthy pointed out that pure classical logic is not adequate to represent the commonsense nature of human reasoning. The reason is, the human reasoning is non-monotonic in nature. This means, we reach to conclusions from certain premises that we would not reach if certain other sentences are included in our premises.

The non-monotonic human reasoning is caused by the fact that our knowledge about the world is always incomplete and therefore we are forced to reason in the absence of complete information. Therefore we often revise our conclusions, when new information becomes available.

Thus, the need for non-monotonic reasoning in AI was recognized, and several formalizations of non-monotonic reasoning.

Non-Monotonic Logic

A monotonic logic cannot handle :

- Reasoning by default : because consequences may be derived only because of lack of evidence of the contrary.
- Abductive reasoning : because consequences are only deduced as most likely explanations.
- Belief revision : because new knowledge may contradict old beliefs.

A non-monotonic logic is a formal logic whose consequence relation

is not monotonic. A logic is non-monotonic if the truth of a proposition

may change when new information (axioms) are added.

Allows a statement to be retracted.

Used to formalize plausible (believable) reasoning.

Example 1 : Birds typically fly.

Tweety is a bird.

Tweety (presumably) flies.

Different Methods of Reasoning

Mostly three kinds of logical reasoning: Deduction, Induction, Abduction.

Deduction

- Example: "When it rains, the grass gets wet. It rains. Thus, the grass is wet."
- This means in determining the conclusion; it is using rule and its precondition to make a conclusion.
- Applying a general principle to a special case.
- Using theory to make predictions
- Usage: Inference engines, Theorem provers, Planning.

Induction

- Example: "The grass has been wet every time it has rained. Thus, when it rains, the grass gets wet."
- This means in determining the rule; it is learning the rule after numerous examples of conclusion following the precondition.
- Deriving a general principle from special cases
- From observations to generalizations to knowledge
- Usage: Neural nets, Bayesian nets, Pattern recognition

Abduction

- Example: "When it rains, the grass gets wet. The grass is wet, it must have rained."
- Means determining the precondition; it is using the conclusion and the rule to support that the precondition could explain the conclusion.
- Guessing that some general principle can relate a given pattern of cases
- Extract hypotheses to form a tentative theory
- Usage: Knowledge discovery, Statistical methods, Data mining.

Default Reasoning

- This is a very common from of non-monotonic reasoning. The conclusions are drawn based on what is most likely to be true.
- There are two approaches, both are logic type, to Default reasoning :

one is Non-monotonic logic and the other is Default logic.

- Non-monotonic logic
  - It has already been defined. It says, "the truth of a proposition may change when new information (axioms) are added and a logic may be build to allows the statement to be retracted."
- Non-monotonic logic is predicate logic with one extension called modal operator M which means “consistent with everything we know”.
- The purpose of M is to allow consistency.

A way to define consistency with PROLOG notation is :

To show that fact P is true, we attempt to prove ¬P.

If we fail we may say that P is consistent since ¬P is false.

Example :

$\forall x$: plays_instrument(x) $\wedge$ M manage(x) $\rightarrow$ jazz_musician(x)

States that for all x, the x plays an instrument and if the fact

that x can manage is consistent with all other knowledge then we

can conclude that x is a jazz musician.

Statistical Reasoning :

In the logic based approaches described, we have assumed that everything is either believed false or believed true.

However, it is often useful to represent the fact that we believe such that something is probably true, or true with probability (say) 0.65.

This is useful for dealing with problems where there is randomness and unpredictability (such as in games of chance) and also for dealing with problems where we could, if we had sufficient information, work out exactly what is true.

To do all this in a principled way requires techniques for probabilistic reasoning.

Bayes' theorem:

Bayes' theorem is also known as Bayes' rule, Bayes' law, or Bayesian reasoning, which determines the probability of an event with uncertain knowledge.

In probability theory, it relates the conditional probability and marginal probabilities of two random events.

Bayes' theorem was named after the British mathematician Thomas Bayes. The Bayesian inference is an application of Bayes' theorem, which is fundamental to Bayesian statistics.

It is a way to calculate the value of P(B|A) with the knowledge of P(A|B).

Bayes' theorem allows updating the probability prediction of an event by observing new information of the real world.

Example: If cancer corresponds to one's age then by using Bayes' theorem, we can determine the probability of cancer more accurately with the help of age.

Bayes' theorem can be derived using product rule and conditional probability of event A with known event B:

As from product rule we can write:

$$P(A \wedge B) = P(A|B) P(B) \text{ or}$$

Similarly, the probability of event B with known event A:

$$P(A \wedge B) = P(B|A) P(A)$$

Equating right hand side of both the equations, we will get:

$$P(A|B) = \frac{P(B|A)\,P(A)}{P(B)} \quad \text{....(a)}$$

The above equation (a) is called as Bayes' rule or Bayes' theorem. This equation is basic of most modern AI systems for probabilistic inference.

It shows the simple relationship between joint and conditional probabilities. Here,

P(A|B) is known as posterior, which we need to calculate, and it will be read as Probability of hypothesis A when we have occurred an evidence B.

P(B|A) is called the likelihood, in which we consider that hypothesis is true, then we calculate the probability of evidence.

P(A) is called the prior probability, probability of hypothesis before considering the evidence

P(B) is called marginal probability, pure probability of an evidence.

In the equation (a), in general, we can write P (B) = P(A)*P(B|Ai), hence the Bayes' rule can be written as:

$$P(A_i|B) = \frac{P(A_i)*P(B|A_i)}{\sum_{i=1}^{k} P(A_i)*P(B|A_i)}$$

Where A1, A2, A3,........, An is a set of mutually exclusive and exhaustive events.

Applying Bayes' rule:

- Bayes' rule allows us to compute the single term P(B|A) in terms of P(A|B), P(B), and P(A). This is very useful in cases where we have a good probability of these three terms and want to determine the fourth one. Suppose we want to perceive the effect of some unknown cause, and want to compute that cause, then the Bayes' rule becomes:

$$P(\text{cause} \mid \text{effect}) = \frac{P(\text{effect} \mid \text{cause})\, P(\text{cause})}{P(\text{effect})}$$

Certainty Factor (CF)

- The Certainty Factor (CF) is a numeric value which tells us about how likely an event or a statement is supposed to be true. It is somewhat similar to what we define in probability, but the difference in it is that an agent after finding the probability of any event to occur cannot decide what to do. Based on the probability and other knowledge that the agent has, this certainty factor is decided through which the agent can decide whether to declare the statement true or false.
- The value of the Certainty factor lies between -1.0 to +1.0, where the negative 1.0 value suggests that the statement can never be true in any situation, and the positive 1.0 value defines that the statement can never be false. The value of the Certainty factor after analyzing any situation will either be a positive or a negative value lying between this range. The value 0 suggests that the agent has no information about the event or the situation.

- A minimum Certainty factor is decided for every case through which the agent decides whether the statement is true or false. This minimum Certainty factor is also known as the threshold value.

For example

if the minimum certainty factor (threshold value) is 0.4,

then if the value of CF is less than this value,

then the agent claims that particular statement false.

Dempster-Shafer Theory

- Dempster-Shafer theory is an approach to combining evidence
- Dempster (1967) developed means for combining degrees of belief derived from independent items of evidence.
- Each fact has a degree of support, between 0 and 1:
  - 0 No support for the fact
  - 1 full support for the fact
- Differs from Bayesian approah in that:
  - Belief in a fact and its negation need not sum to 1.
  - Both values can be 0 (meaning no evidence for or against the fact)

- Set of possible conclusions: Θ Θ = { $\theta 1$, $\theta 2$, …, θn}

Where:

  - Θ is the set of possible conclusions to be drawn
  - Each θi is mutually exclusive: at most one has to be true.
  - Θ is Exhaustive: At least one θi has to be true.

Frame of discernment :

$$\Theta = \{ \theta 1, \theta 2, \ldots, \theta n\}$$

- Bayes was concerned with evidence that supported single conclusions (e.g., evidence for each outcome θi in Θ):

$$p(\theta i | E)$$

- D-S Theoryis concerned with evidences which support subsets of outcomes in Θ, e.g.,

$$\theta 1 \vee \theta 2 \vee \theta 3 \text{ or } \{\theta 1, \theta 2, \theta 3\}$$

- Frame of discernment :
- The “frame of discernment” (or “Power set”) of Θ is the set of all possible subsets of Θ:– E.g., if Θ = { θ1, θ2, θ3}
- Then the frame of discernment of Θ is:

(Ø, θ1, θ2, θ3, {θ1, θ2}, {θ1, θ3}, {θ2, θ3}, { θ1, θ2, θ3})

- Ø, the empty set, has a probability of 0, since one of the outcomes has to be true.
- Each of the other elements in the power set has a probability between 0 and 1.
- The probability of { θ1, θ2, θ3} is 1.0 since one has to be true.

Mass function m(A):

(where A is a member of the power set)

= proportion of all evidence that supports this element of the power set.

“The mass m(A) of a given member of the power set, A, expresses the proportion of all relevant and available evidence that supports the claim that the actual state belongs to A but to no particular subset of A.” (wikipedia) “The value of m(A) pertains only to the set A and makes no additional claims about any subsets of A, each of which has, by definition, its own mass.

Mass function m(A):

- Each m(A) is between 0 and 1.
- All m(A) sum to 1.
- m(Ø) is 0 - at least one must be true.

Mass function m(A): Interpetation of m({AvB})=0.3

- means there is evidence for {AvB} that cannot be divided among more specific beliefs for A or B.

Fuzzy logic

Fuzzy logic was suggested by Zadeh as a method for mimicking the ability of human reasoning using a small number of rules and still producing a smooth output via a process of interpolation.

With fuzzy logic an element could partially belong to a set represented by the set membership. Example, a person of height 1.79 m would belong to both tall and not tall sets with a particular degree of membership.

Difference between binary logic and fuzzy logic

Grade of thruth

Not tall Tall

1

0

1.8 M height x

Binary valued logic {0, 1}

Grade of thruth

Not tall Tall

1

0

1.8 M height x

Fuzzy logic [0, 1]

A fuzzy logic system is one that has at least one system component that

uses fuzzy logic for its internal knowledge representation.

Fuzzy system communicate information using fuzzy sets.

Fuzzy logic is used purely for internal knowledge representation and

externally it can be considered as any other system component.

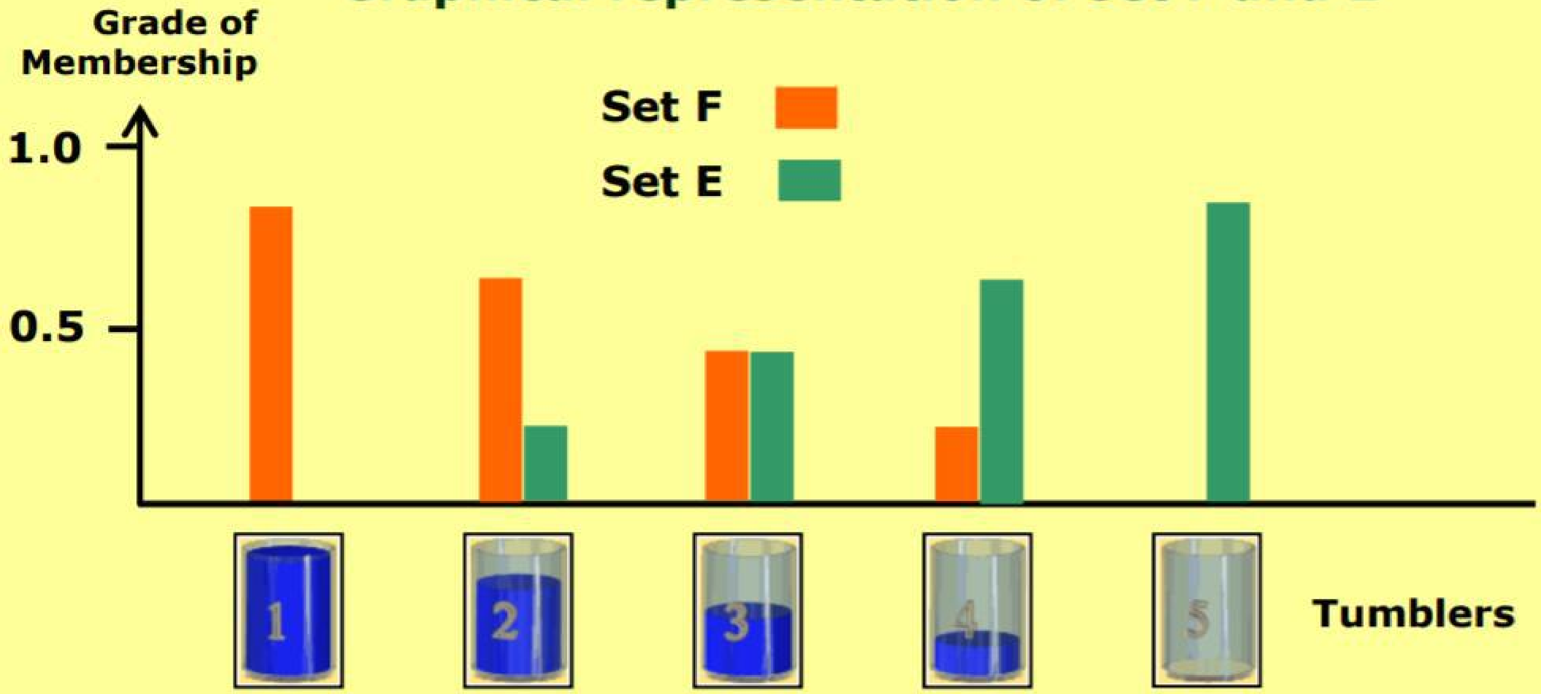
Fuzzy Membership

Example : Five tumblers

- Consider two sets: F and E.
- F is set of all tumblers belong to the class full, and
- E is set of all tumblers belong to the class empty.
- The sets F and E have some elements, having partial membership.
- Such kind of non-crisp sets are called fuzzy sets.
- The set "all tumblers" here is the basis of the fuzzy sets F and E,
- is called the base set.

Tumblers					
Grade of membership to set ***F***	**100%**	**75%**	**50%**	**25%**	**0%**
Grade of membership to set ***E***	**0%**	**25%**	**50%**	**75%**	**100%**

Graphical representation of set F and E



THANKS