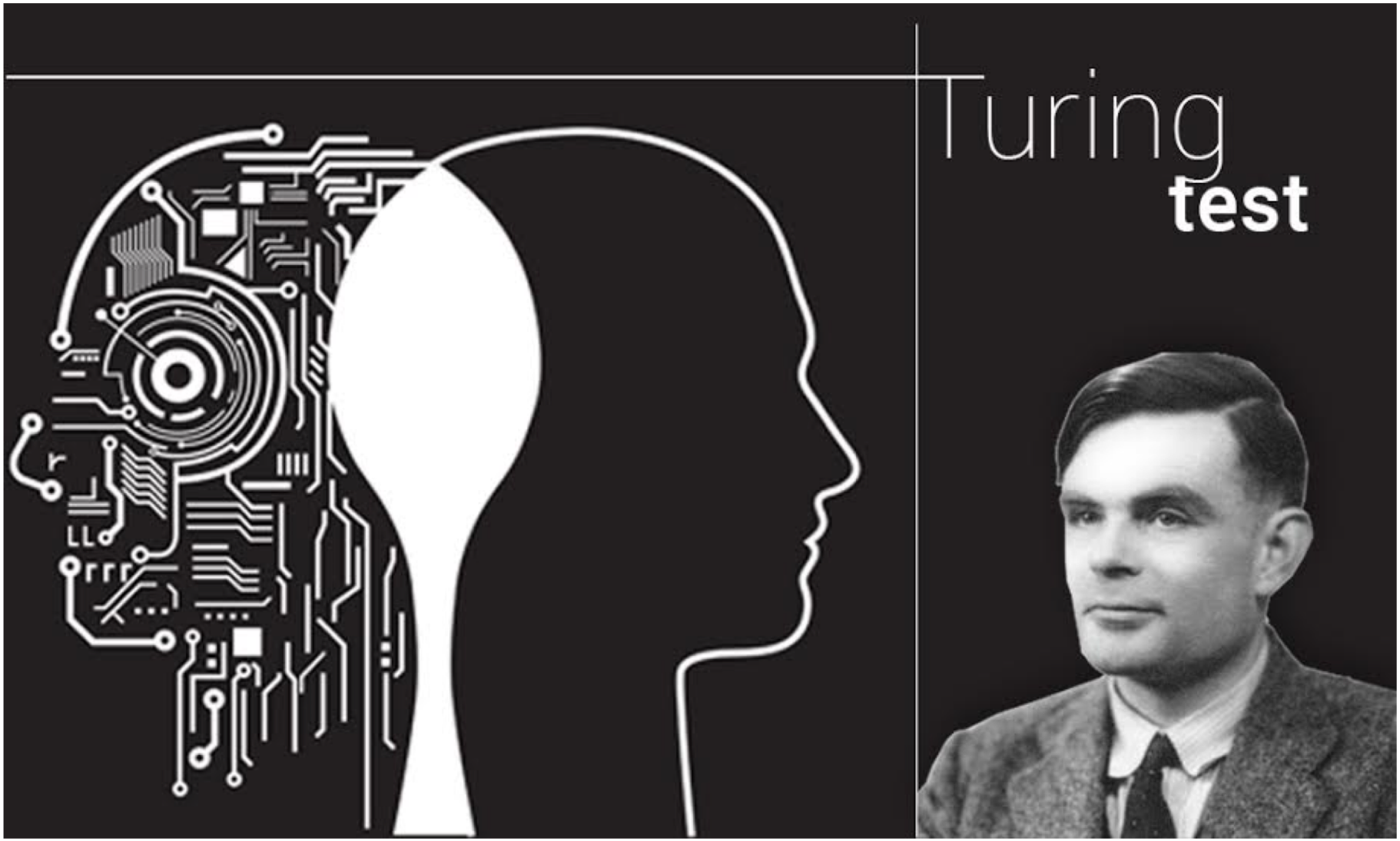
Artificial Intelligence & Machine Learning

A study material for the students of GLS University

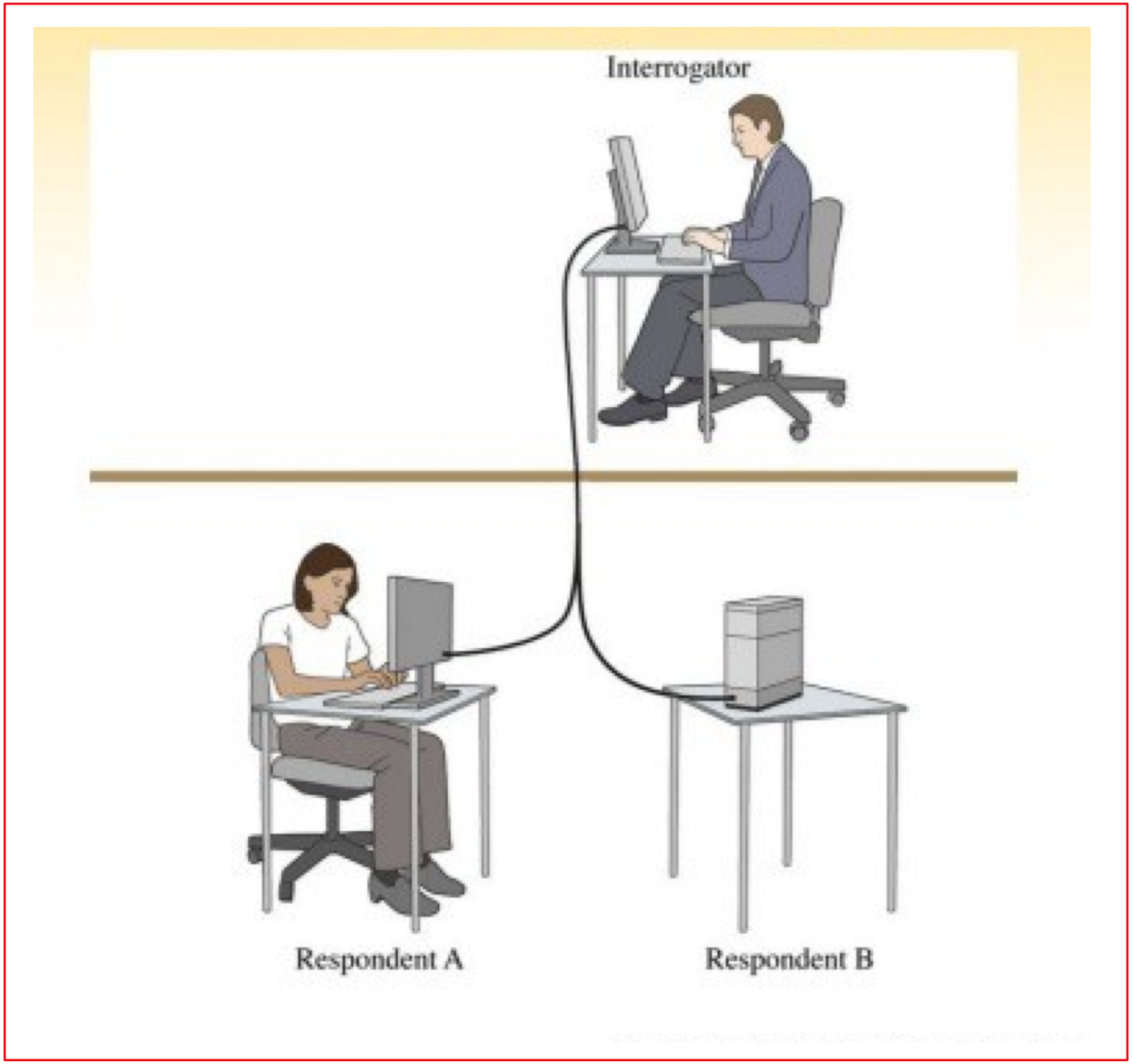
ARTIFICIAL INTELLIGENCE

- AI is the study of how to make computers do things which at the moment, people do better.
- *John McCarthy coined the term Artificial Intelligence in 1956* and defined it as ”The science and engineering of making intelligent machines and especially intelligent computer programs.”
- AI is the study of computations that make it possible to perceive, reason and act.
- The art of creating machines that perform functions which require intelligence when performed by people.
- And many more…

TURING TEST



TURING TEST



TURING TEST

- The turing test was focused in the areas of strong equivalence and weak equivalence.
- Weak equivalence is a system for humans and computers that are equivalent in results.
- Strong equivalence is a system for humans and computer that uses the same internal processes to produce results.

STATE SPACE APPROACH

- All the problems stated in AI commonly use the term ‘state’.
- We can summarize the solution of a problem by a collection of problem states.
- The problem solving procedure used is to apply an operator to a state to get the next state.
- Then another operator is applied to the new state to get the next state.
- The process of deriving new state from the current state by applying the operator is continued till the desired state is reached. This form of problem solving is known as State Space Approach.

WHAT ARE AI PROBLEMS? HOW ARE AI PROBLEMS DIFFERENT FROM OTHERS?

1. If a solution requires symbolic representation then it is an AI problem. Ex: Processing images.
2. If there is a combinatorial explosion in outputting the result then it is an AI problem. Ex: Chess board, Card Game
3. It is difficult to quantify data in AI. So we use fuzzy set theory. Ex: white, beautiful etc.

WHAT ARE AI PROBLEMS? HOW ARE AI PROBLEMS DIFFERENT FROM OTHERS?

4. The knowledge base of an AI problem is voluminous. It increases with time.
5. The size of the knowledge base keeps changing fast and increases day by day.
6. Jobs in which humans get fatigue and tired, machine can do it tirelessly. Ex: ATM vs Cashier

CHARACTERISTICS OF AI PROBLEMS

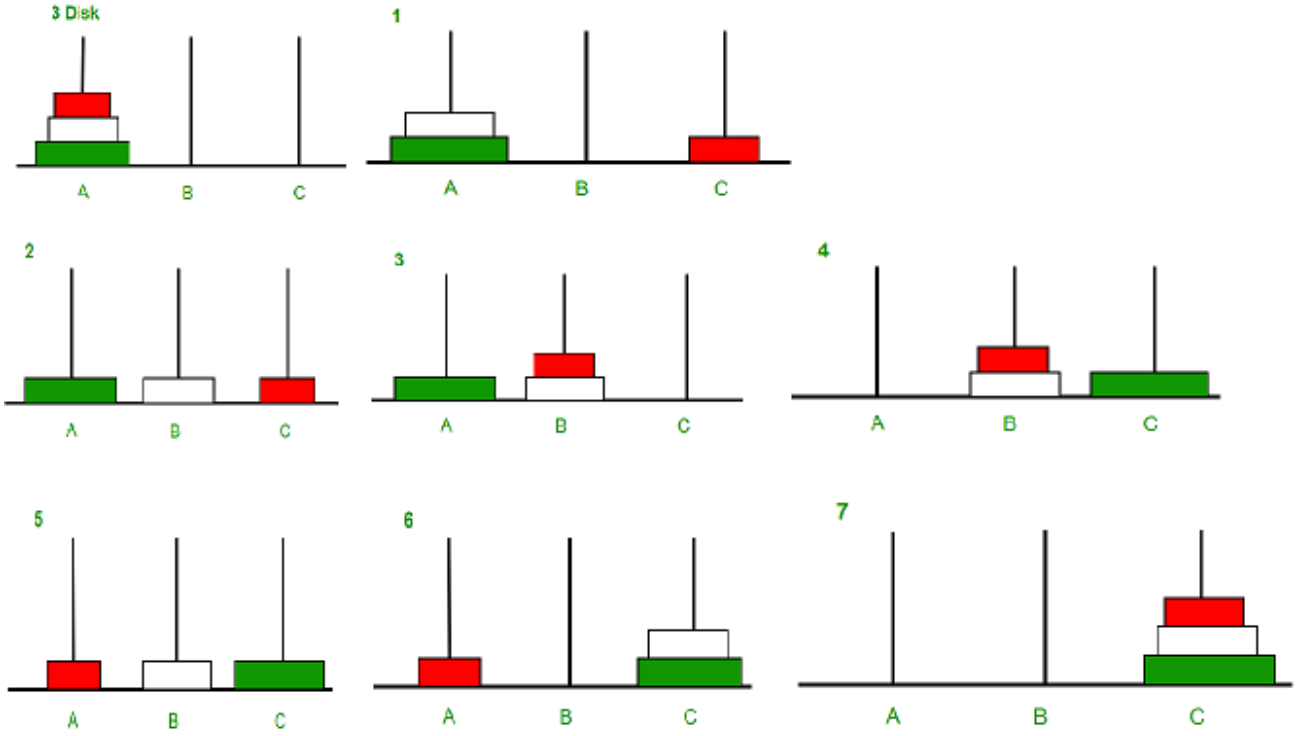
- Is the problem decomposable or not? Ex: Solving an equation requires dividing it into subparts and solving it.
- Can the solution steps be ignored? Recovered or is not recoverable? Ex: Theorem proving- Ignored, 8 puzzle- Recovered, Chess- non recoverable.
- Is the universe predictable? 8 puzzle problem-Predictable, Playing Cards-not predictable.
- Is the solution to the goal absolute or relative? Ex: In 8 puzzle problem-any path leading to the solution may be taken-absolute. In travelling salesman problem-path as well as the goal are to be considered-relative.

CHARACTERISTICS OF AI PROBLEMS

- Is the knowledge base consistent or not? If a knowledge base contains facts, then it is not consistent. Ex: Boolean Expression-consistent, Facts about production and cost-Inconsistent.
- The role of knowledge base is critical. Ex: Chess, electing a politician.
- Is the interaction with the computer necessary? Ex: Playing chess-interaction is required, Program execution-no human interaction.

DECOMPOSABLE & NON DECOMPOSABLE PROBLEM

- $\int(x^2+5x+\sin^2 x)$ is an example of decomposable problem.
- Tower of Hanoi is an example of non decomposable problem.



8-PUZZLE PROBLEM

Operations: Up, Down, Left, Right

1	2	
4	5	3
7	8	6

Initial State

1	2	3
4	5	6
7	8	

Goal State

TOPICS OF AI

- Learning Systems
- Knowledge Representation and Reasoning
- Planning
- Knowledge Acquisition
- Intelligent Search
- Logic Programming
- Soft Computing
- Management of Imprecision and Uncertainity

LEARNING SYSTEMS

- Learning is any process by which a system improves performance from experience.
- Ex: Child starts to learn pronunciation from mother.
- Adaptive Learning Process adjusts parameters continuously to get the results close enough to training dataset.
- Other Learning methods:

1. Inductive Learning
2. Analogue based Learning

KNOWLEDGE REPRESENTATION & REASONING

- Knowledge is an expertise and skill acquired by people which is very hard to represent in computers.
- Knowledge base is similar to a database which is necessary for expert system building and reasoning.
- A reasoning system is a software system that generates conclusions from available knowledge using various techniques such as deduction and induction.

PLANNING

- Planning is a significant area in AI.
- Problem Solving involves the determination of methodology from which a successful result can be obtained.
- It decomposes the original problem into appropriate subproblem and uses interaction on these subunits during the problem solving process.

KNOWLEDGE ACQUISITION

- It involves making new pieces of knowledge from a given knowledge network, setting active structures for existing knowledge.

INTELLIGENT SEARCH

- The search problems encountered in Computer Science are of deterministic nature i.e. the order of visiting elements of the search space is known.
- But the search problems in AI are non deterministic in nature and the order of visiting elements is completely dependent on the given data sets.
- Ex: DFS, BFS etc.

LOGIC PROGRAMMING

- Mathematicians are used to designing various tools for representing logical statements by symbolic operators.
- Ex: ∀x: computer-science-student(x) → smart(x) means "All computer science students are smart."

SOFT COMPUTING

- Soft Computing is a collection of computing techniques and tools, shared by closely related disciplines that comprise fuzzy logic, artificial neural networks, genetic algorithms, inductive logic programming and so on.

MANAGEMENT OF IMPRECISION AND UNCERTAINTY

- Data and knowledge bases in many AI problems, such as reasoning and planning, are often contaminated with various forms of incompleteness.
- This incompleteness of data can be called as imprecision. Imprecision generally occurs in databases due to lack of appropriate data and poor authentication level of sources.
- A complex problem in reasoning is the presence of imprecision of data and uncertainty of knowledge.

BRANCHES OF AI

- Logical AI: What a program knows about the world is represented by sentences of some mathematical logical language.
- Search: AI programs often examine large number of possibilities. Ex: Chess game. Various search techniques are useful in the development of intelligence application which cuts the time delays at the time of searching.
- Pattern Recognition: Ex: To find a face, a vision program may try to match a pattern of eyes and a nose.

BRANCHES OF AI

- Representation: Facts about the world are represented symbolically using mathematical logic. These representations can be used in AI algorithms for making a solution.
- Inference: It is deriving logical conclusions from a true premise known or assumed. This makes conclusions by default, but the conclusion can be withdrawn if there is evidence.
- Common Sense Knowledge and Reasoning: Since 1950s onward, this is one of the areas of AI that is the farthest from human level inspite of the fact that it has been an active research area.

INFERENCE EXAMPLE

Triangle ABC has no more than one right angle. Can you complete a proof by contradiction for this statement?

Assume $\angle A$ and $\angle B$ are right angles

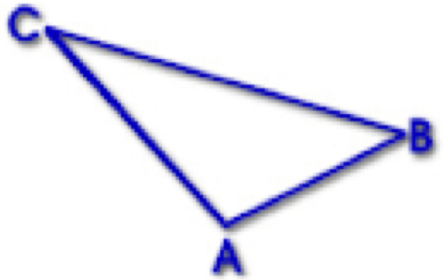
We know $\angle A + \angle B + \angle C = 180^0$

By substitution $90^0 + 90^0 + \angle C = 180^0$

$\therefore$ $\angle C = 0^0$ which is a contradiction

$\therefore$ $\angle A$ and $\angle B$ cannot both be right angles

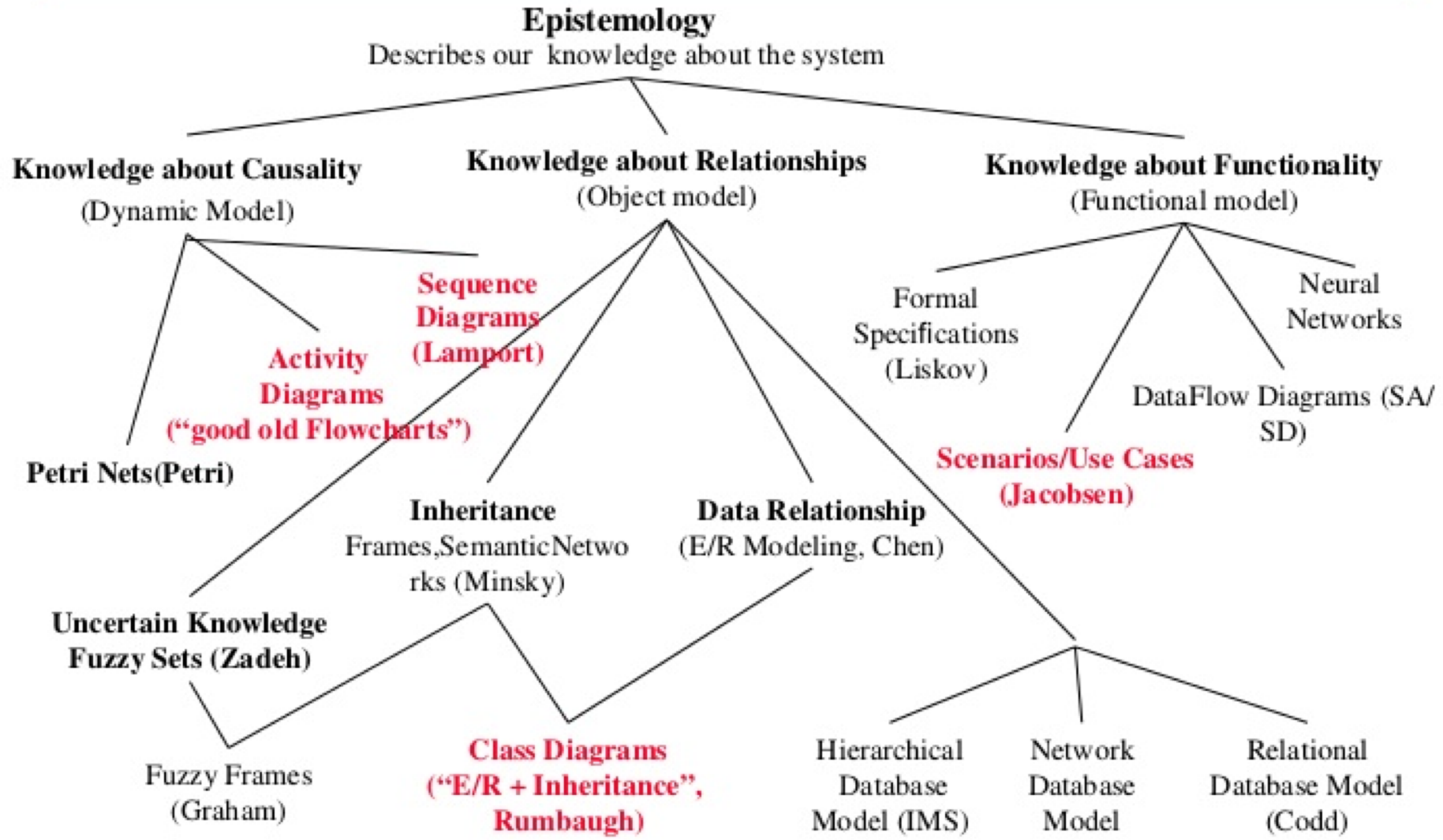
$\Rightarrow$ A triangle can have at most one right angle



BRANCHES OF AI

- Learning from experience: AI programs can learn from experience. There are many learning approaches in AI based on specialized applications.
- Planning: Planning in AI deals with general facts of the world and the effects of actions. For a particular statement and goal achievement, planning plays its facts.
- Epistemology: In AI programs, it involves the study of the kinds of knowledge that are required for solving problems in the world.

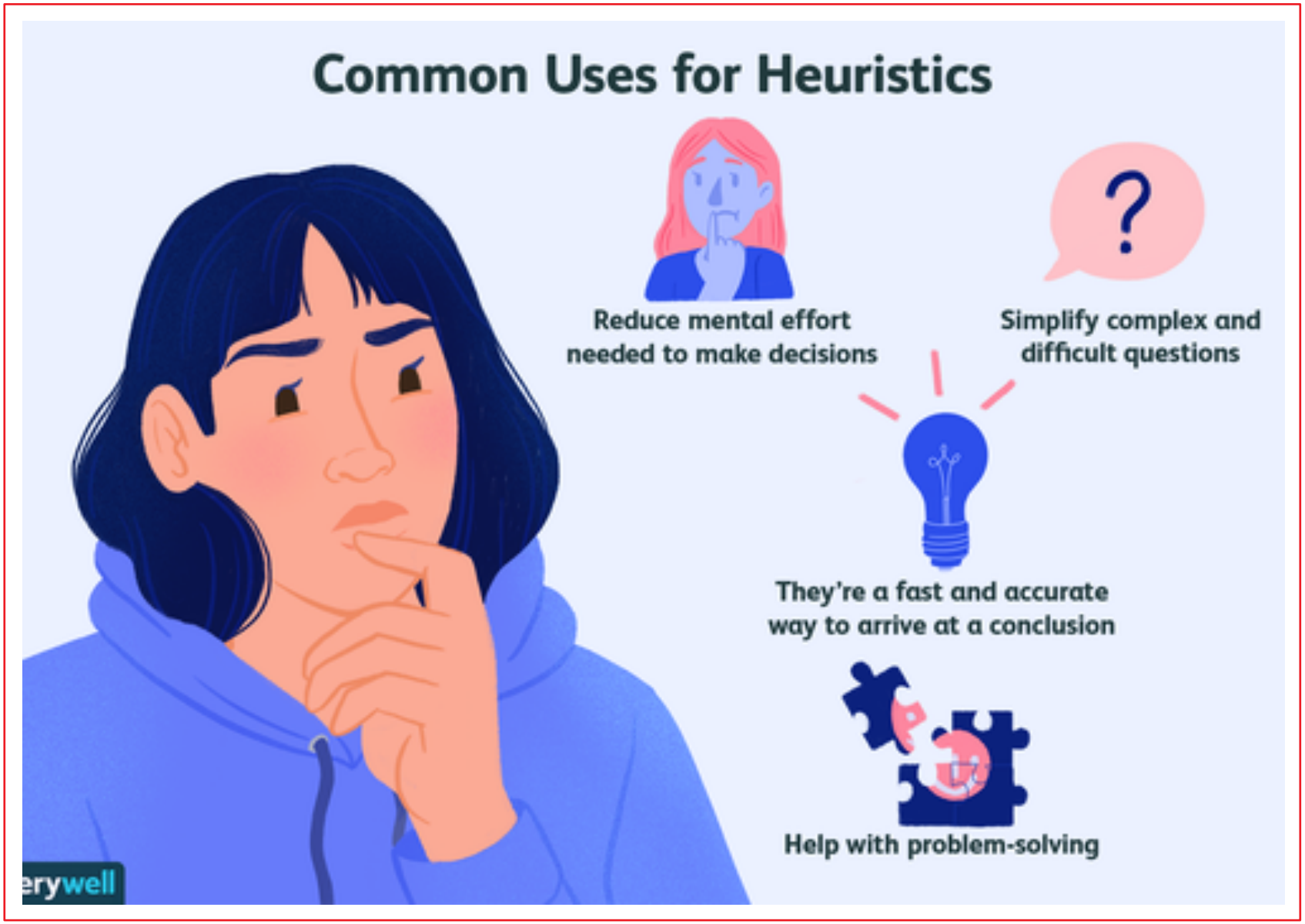
EPISTEMOLOGY



BRANCHES OF AI

- Ontology: It is the study of kinds of things that exist. In AI, the programs and sentences deal with various kinds of objects and we study what these kinds are and what their basic properties are.
- Heuristics: To discover or identify the ideas embedded in a program, we use certain functions of mathematics. Ex: In shortest path identification problem, a heuristic function is used to measure the estimate or the cost of the cheapest path from a node to the goal node.
- *Heuristic* is a fast and accurate way to make decisions in the real world, which is driven by uncertainity.

HEURISTICS


Common Uses for Heuristics
Reduce mental effort needed to make decisions
Simplify complex and difficult questions
They're a fast and accurate way to arrive at a conclusion
Help with problem-solving
erywell

BRANCHES OF AI

- Genetic Programming: It is a soft computing technique for getting programs to solve a task by mating random and selecting the fittest in millions of generations.

PRODUCTION SYSTEMS

- A system in which knowledge is represented in the form of rules is called a production system.
- A production system consists of four basic components:

1. Set of rules
2. Knowledge base
3. Control strategy
4. Rule Applier

PRODUCTION SYSTEMS

1. A set of rules of the form Ci → Ai where Ci is the condition part and Ai is the action part. The condition determines when a given rule is applied, and the action determines what happens when it is applied.
2. One or more knowledge databases that contain whatever information is relevant for the given problem. Some parts of the database may be permanent, while others may temporary and only exist during the solution of the current problem. The information in the databases may be structured in any appropriate manner.

PRODUCTION SYSTEMS

3. A control strategy that determines the order in which the rules are applied to the database, and provides a way of resolving any conflicts that can arise when several rules match at once.
4. A rule applier which is the computational system that implements the control strategy and applies the rules.

TYPES OF PRODUCTION SYSTEMS

- Monotonic production system:- A system in which the application of a rule never prevents the later application of another rule, that could have also been applied at the time the first rule was selected.
- Non Monotonic production system:- A system in which the application of a rule prevents the later application of another rule which may not have been applied at the time the first rule was selected.

TYPES OF PRODUCTION SYSTEMS

In simple terms….

Monotonic production system:-

- Anything that could be concluded before a clause is added can still be concluded after it is added.

Non Monotonic production system:-

- A logic is non monotonic where some conclusions can be invalidated by adding more knowledge.

TYPES OF PRODUCTION SYSTEMS

- Partially commutative production system:- A production system in which the application of a particular sequence of rules transforms state X into state Y, then any permutation of those rules that is allowable also transforms state x into state Y.
- Commutative production system:- A program which satisfies both monotonic and partially commutative production systems.

ADVANTAGES OF PRODUCTION SYSTEM

- Provides excellent tools for structuring AI programs.
- The system is highly modular because individual rules can be added, removed or modified independently.
- Expressed in natural form.

DISADVANTAGES OF PRODUCTION SYSTEM

- It's very difficult to analyse the flow of control within a production system. As the rules of the production system are large in number and they are hardly written in hierarchical manner, it requires some forms of complex search through all the production rules for each cycle of control program.
- It describes the operations that can be performed in a search for a solution to the problem.

DISADVANTAGES OF PRODUCTION SYSTEM

- There is an absence of learning due to a rule-based production system which does not store the result of the problem for future use.
- The rules in the production system should not have any type of conflict resolution as when a new rule is added to the database it should ensure that it does not have any conflict with any existing rules.

SOME OF THE PRODUCTION SYSTEMS

- Travelling Salesman Problem
- Water Jug Problem
- The Knight's Tour Problem
- State Space Representation
  1. State Space Search
  2. Tic –Tac-Toe
  3. The Missionaries and Cannibals Problem

TRAVELLING SALESMAN PROBLEM

- Travelling Salesman Problem (TSP) is a touring problem in which n cities and distance between each pair is given. We have to find a shortest route to visit each city exactly once and come back to the starting point.

TRAVELLING SALESMAN PROBLEM

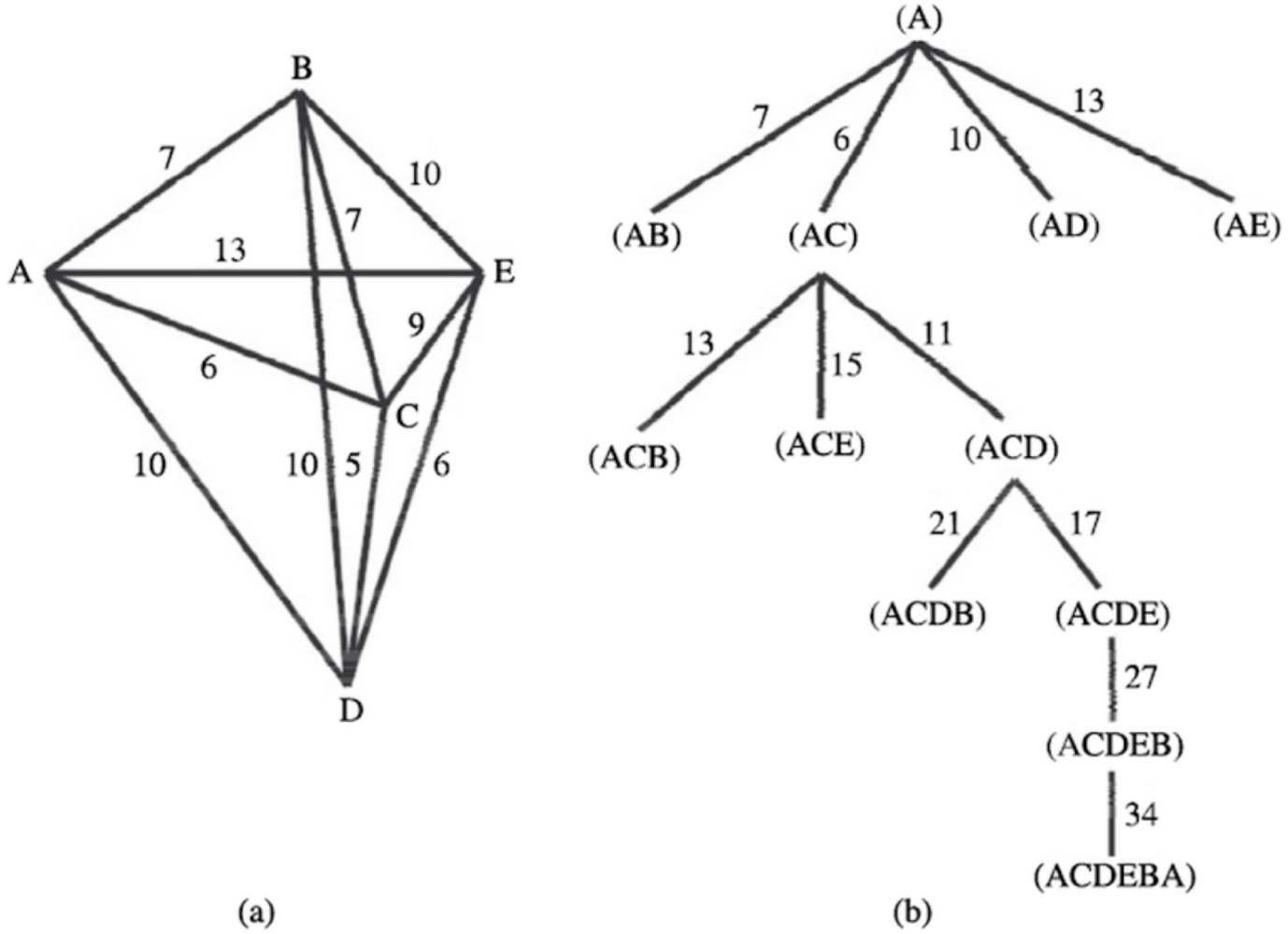
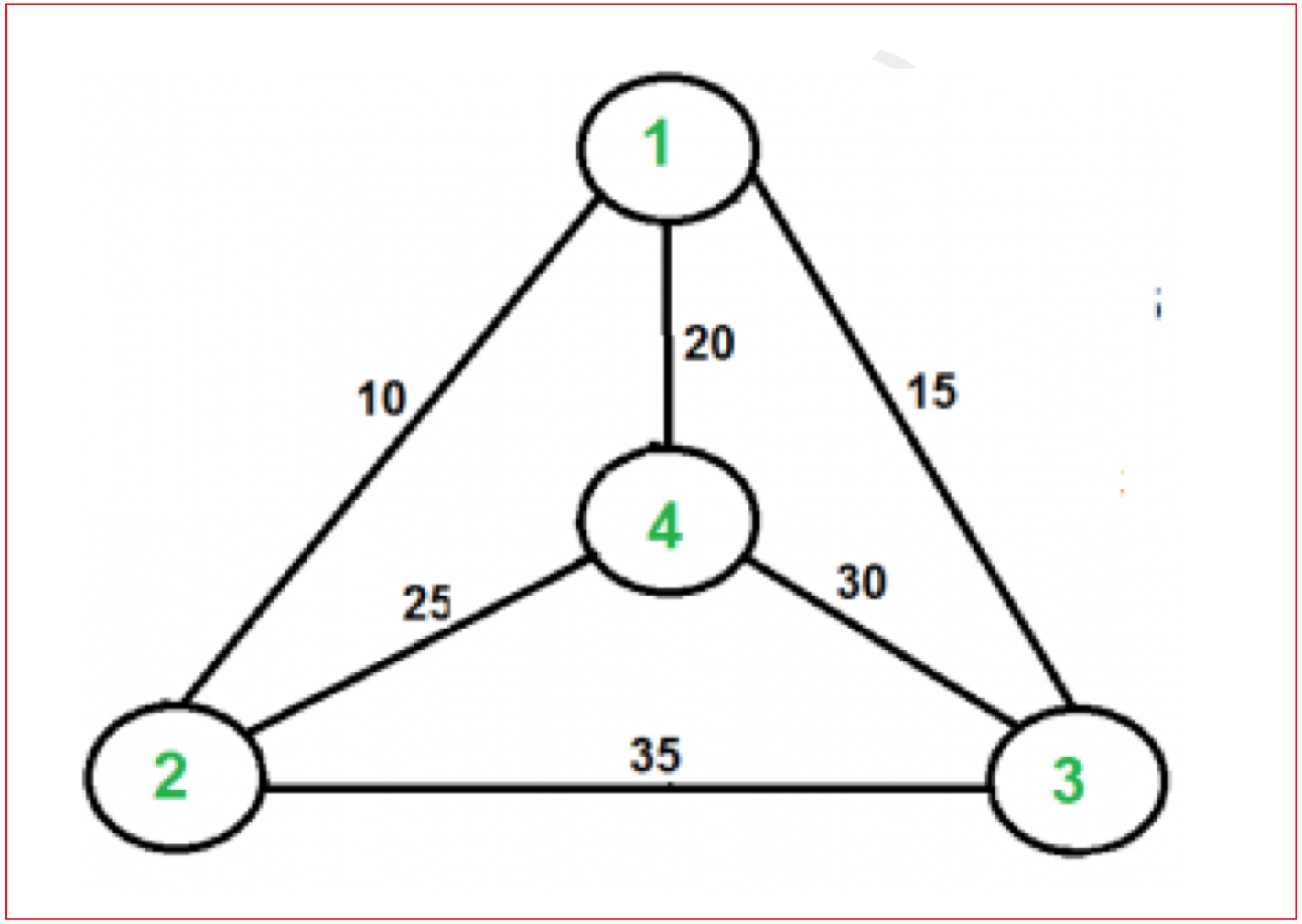


Figure 1.4 (a) 5 cities connected; (b) global database.

TRAVELLING SALESMAN PROBLEM

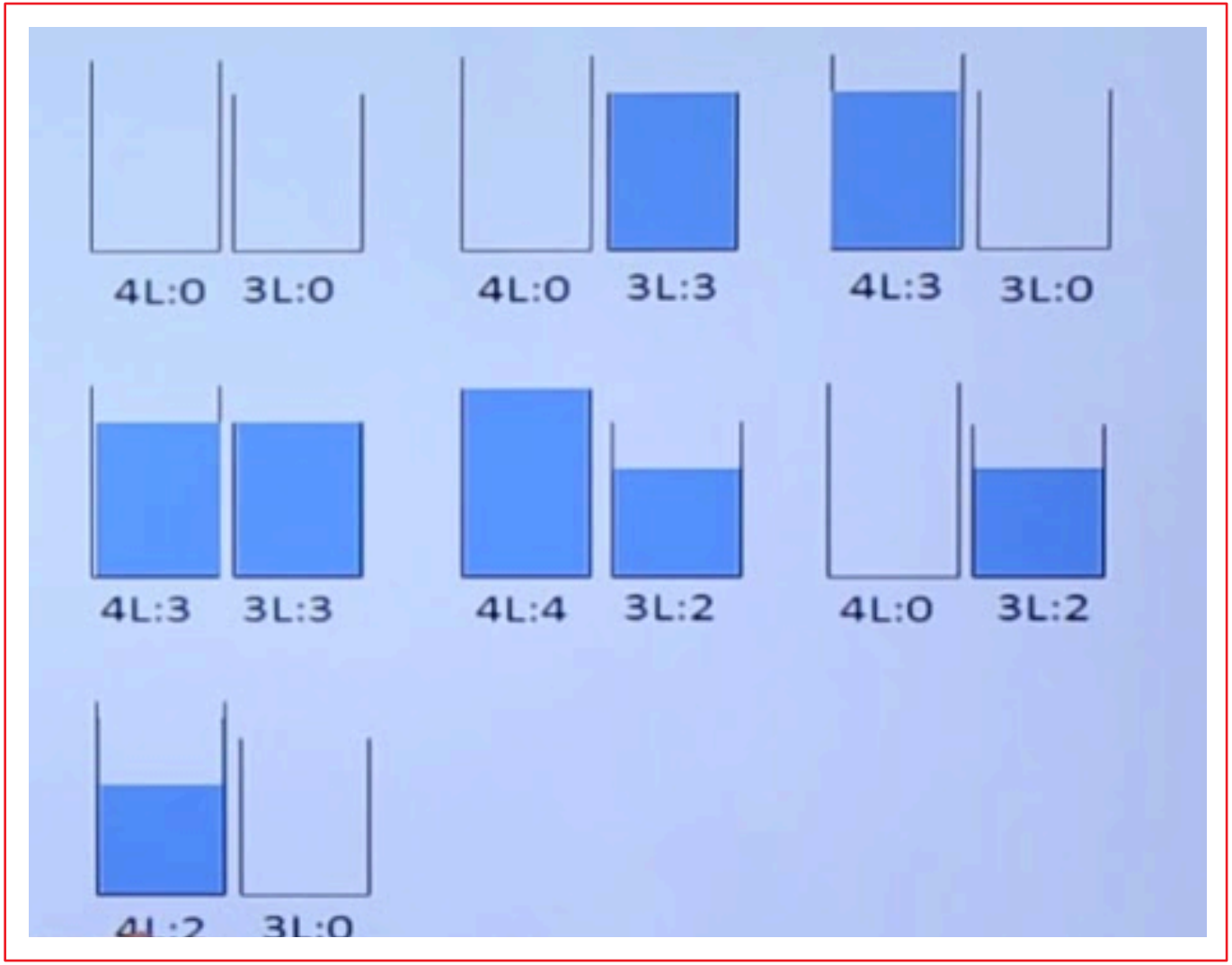


WATER JUG PROBLEM

- You are given two jugs, a 4L one and a 3L one, a pump which has unlimited water which you can use to fill the jug, and the ground on which water may be poured. Neither jug has any measuring markings on it.
- Rules of the problem are as follows:

1. We can empty the jug fully on to the ground or from one jug to the other.
2. We should not fill up a jug or empty a jug half way or fractionally from the pump or to the ground.
3. How can you get exactly 2L of water in the 4L jug?

WATER JUG PROBLEM



WATER JUG PROBLEM

- Solution:

Rule no.	Result	
	A	B
Initial state	0	0
1	4	0
8	1	3
6	1	0
10	0	1
1	4	1
8	2	3

THE KNIGHT'S TOUR PROBLEM

- In a chess game, a knight can move two squares either horizontally or vertically followed by one square in an orthogonal direction as long as it does not move off the board.
- Find out the series of legal moves in which a knight lands on each square of a chess board exactly once.

STATE SPACE REPRESENTATION

- A production system consists of a global database, a set of rules and a goal. The initial state can be represented in a way in which the computer can understand. This representation is known as State Space Representation.
- By taking up an applicable rule, we can derive another state in the solution path of the problem. The same has to be represented in the computer in a form in which it can understand.
- All the intermediate configurations obtained by applying the rules in a production system are known as set of states in the global database(space).

STATE SPACE SEARCH

- A problem in a state space is represented in terms of states and operators that change states. A state space consists of the following:

1. The representation of a state in a system can be in the form of a data structure. Ex: Board represents the current state of the game in chess.
2. A set of operators that can change from state to state. Ex: In board games, the operators are the legal moves from any given state.
3. An initial state
4. A set of final states.

TIC-TAC-TOE AS A STATE SPACE

Each state of the game can be described through the contents of the board and the next turn of the player. The tic-tac-toe board has an array of 9 cells, each of which may contain an * or o or may be empty.

State:

- **Next move of the player:** * or o
- **Configuration of the board:** It may change according to each move of the player.
- **Operators:** Change an empty cell into * or o.
- **Start State:** Initially board is empty.
- **Terminal State:** Three *'s in a row; three o's in a row; all cells full.

THE MISSIONARIES AND CANNIBAL'S PROBLEM

- Three missionaries and three cannibals are on one side of a river that they wish to cross. A boat is available that can hold at most two people and at least one. You must never leave a group of missionaries outnumbered by cannibals on the same bank.
- Find an action sequence that brings everyone safely to the opposite bank.

THE MISSIONARIES AND CANNIBAL'S PROBLEM

State Representation

- BOAT position: Original(T) or Final(NIL) side of the river.
- Number of missionaries and cannibals on the original side of the river.
- Start is (T,3,3); goal is (NIL,0,0)

Operators

- (MM 2 0) Two missionaries cross the river
- (MC 1 1) One missionary and one cannibal
- (CC 0 2) Two cannibals
- (M 1 0) One missionary
- (C 0 1) One Cannibal

APPLICATIONS OF AI

- Game Playing
- Expert Systems
- Natural Language Processing
- Image Understanding
- Robotics
- Machine Learning
- Pattern Recognition
- Virtual Reality
- Computer Vision
- Nature Inspired Computing
- Intelligent Control

GAME PLAYING

- One of the leading domains where AI has been applied with great success.
- Games can generate an extremely large search space. A large and complex powerful technique determines the alternatives to explore in the problem space.
- These techniques are based on heuristics and constitute a major area of AI research. Ex: Computer Chess Program.
- Such a game can be defined as a kind of search problem with initial state, a set of operators, terminal state and a utility function.

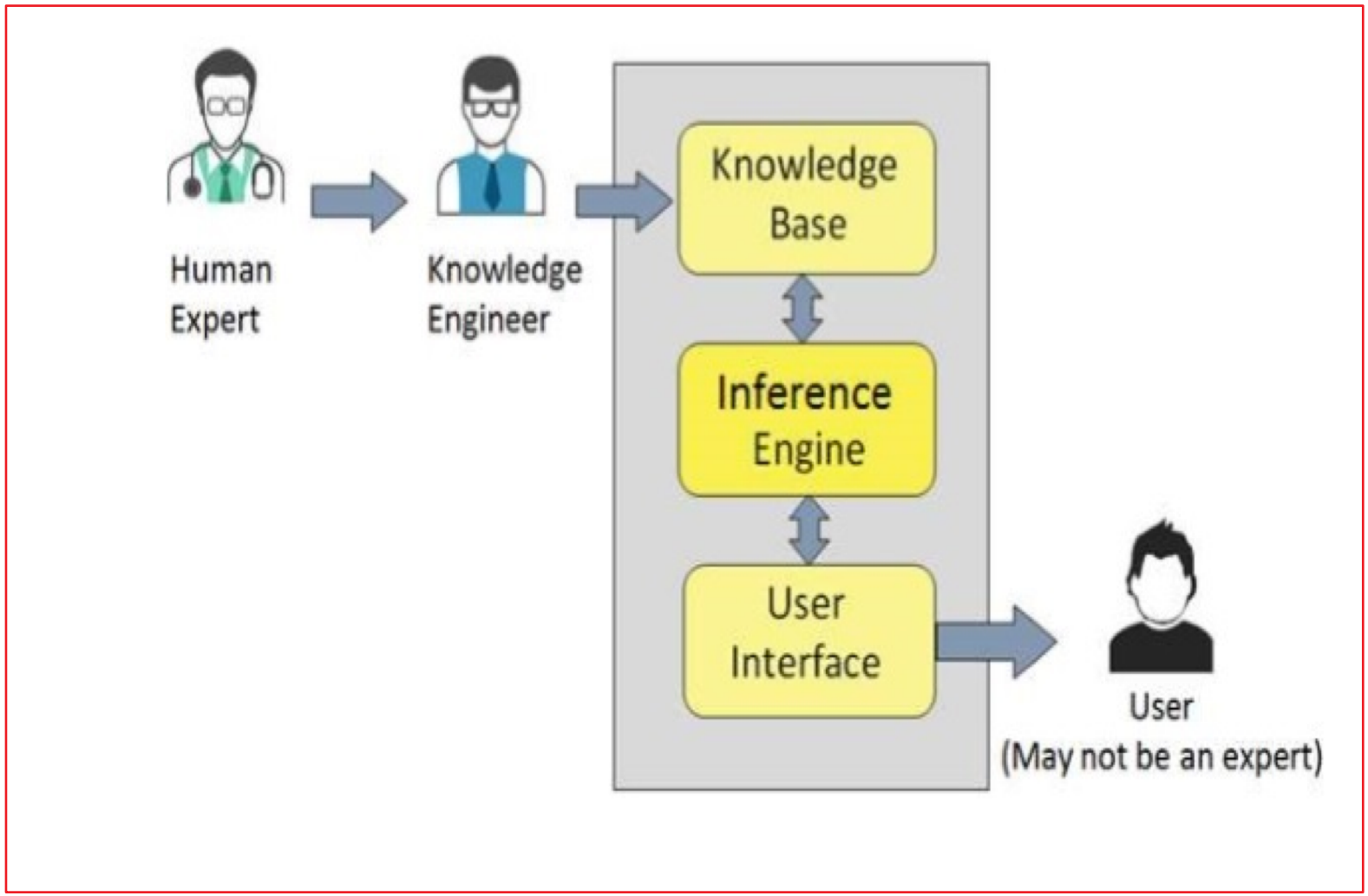
EXPERT SYSTEMS

- An ***Expert system*** is a software that manipulates encoded knowledge to solve problems in a specialized domain that normally requires human expertise.
- An ***Expert System*** is an AI program in which the system's knowledge is obtained from an expert source such that the machine can offer intelligent advice or take an intelligent decision in solving a problem.

EXPERT SYSTEMS

- Developing an expert system means extracting from human expert knowledge that can be considered experts, rules, strategies and procedures to solve a pre determined group of problems and aggregating this knowledge into a computer software system.
- This process is known as Knowledge Engineering and the person responsible for it is known as *Knowledge Engineer*.

EXPERT SYSTEMS



EXPERT SYSTEMS

- Main components of an expert system are: *Knowledge Base, Inference Engine and User Interface*.
- The knowledge base stores facts about the world.
- Inference engine is a component of the system that applies logical rules to the knowledge base to deduce new knowledge.
- There are various types of expert systems like rule base, knowledge base and hybrid systems.
- Here the expert's knowledge is elicited through questionnaires, interviews or a combination of both.

EXPERT SYSTEMS

- Knowledge Acquisition can be very time and effort consuming.
- Here AI performs a decision making task much faster than the humans can do.
- Ex: MYCIN is an expert system that can diagnose infections and suggest better treatments than medical practitioners.
- Modern expert systems use self learning methods to dynamically learn new rules that can be validated by experts and can be added to the knowledge base.

NATURAL LANGUAGE PROCESSING

- NLP is a technique that builds the ability for machines to read and understand the languages that humans speak.
- Powerful NLP system is able to acquire new knowledge by reading text to understand the situations.
- Major areas of NLP include:

1. Information Retrieval: Information retrieval is the activity of obtaining information resources relevant to an information need from a collection of information resources.

NATURAL LANGUAGE PROCESSING

2. Information Extraction: Information extraction (IE) is the task of automatically extracting structured information from unstructured and/or semi-structured machine-readable documents. In most of the cases this activity concerns processing human language texts by means of natural language processing (NLP).
3. Speech Processing: Also known as speech recognition that can recognize spoken words, which can then be converted to text.
4. Text-to-Speech: Text-to-Speech converts text into human-like speech.

NATURAL LANGUAGE PROCESSING

- Natural language search: It is a search using regular spoken language, such as English. Using this type of search you can ask the database a question or you can type in a sentence that describes the information you are looking for.
- Query expansion: It is a process in Information Retrieval which consists of selecting and adding terms to the user's query with the goal of minimizing query-document mismatch and thereby improving retrieval performance.

NATURAL LANGUAGE PROCESSING

- Truecasing: Truecasing is the problem in natural language processing of determining the proper capitalization of words where such information is unavailable. This commonly comes up due to the standard practice of automatically capitalizing the first word of a sentence.
- Machine learning algorithms are used in NLP tasks.

IMAGE UNDERSTANDING

- To interpret a scene, the image is passed through three basic processes like low, medium and high level vision.
- The low level visions pre process the image by filtering from noise.
- The medium level deals with the improvement of details and segmentation. The segmentation of image is partitioning the image into objects of interest.
- The high level vision system includes three steps:

1. Recognizing the object from segmented image
2. Labelling the image
3. Interpretation of the scene

- The high level vision system makes use of AI tools and techniques.

ROBOTICS

- AI is applied in robotics in order to see, hear and to react to other sensory stimuli.
- Robotic is applied in tasks that requires higher precision or that are risky for humans.
- More importantly, a robot can make a vast number of potential sequence of moves in a given amount of space.

PATTERN RECOGNITION

- It is a branch of AI that focuses on recognition of patterns and regularities in data.
- Computers can identify patterns in numbers, texts and images.
- Ex: Cyber Security, buying patterns of the customer
- The act of taking in raw data and taking an action based on the category of the pattern.
- The major aim of pattern recognition is the classification of data that is based either on apriori knowledge or on statistical information extracted from the patterns.

PATTERN RECOGNITION

Some application areas of pattern recognition are:

- Optical character Recognition
- Handwriting Recognition
- Speech Recognition
- Face Recognition

COMPUTER VISION

- Computer vision is a field of computer science that works on enabling computers to see, identify and process images in the same way that human vision does, and then provide appropriate output.
- The image data can be in the form of picture, video sequences, multidimensional data from a medical scanner or views from multiple cameras.
- It is like imparting human intelligence and instincts to a computer. In reality though, it is a difficult task to enable computers to recognize images of different objects.

COMPUTER VISION

- Computer vision is closely linked with artificial intelligence, as the computer must interpret what it sees, and then perform appropriate analysis or act accordingly.
- Organizing information for example, for indexing databases of images and image sequences.
- Interaction for example as the input to a device for computer human interaction.
- Modelling objects or environments for example, industrial inspection, medical image analysis or topographical modelling.
- Detecting events for example, visual surveillance or people counting.

INTELLIGENT CONTROL

- Intelligent control is a class of control techniques that use various artificial intelligence computing approaches like neural networks, Bayesian probability, fuzzy logic, machine learning, reinforcement learning, evolutionary computation and genetic algorithms.

VIRTUAL REALITY

- A variety of specialized input and output devices are necessary.
- A tracker that is capable of sending its location space and orientation, can be used as an input device.
- A **tracker** comes along with a joystick or mouse.
- The attempt to create a truly natural input device was fulfilled by the development of **data gloves**.
- Thus with the numerous sensors on the data gloves, user can interact with the virtual world through hand gestures.

VIRTUAL REALITY

- We can apply Virtual Reality environments in a variety of ways.

1. **In Research:** To visually explore whatever physical world phenomenon is seen.
2. **In Power Plants:** For training personnel to work in dangerous environments with expensive equipments through simulation.
3. **In Health care:** For training medical personnel to practice surgery on simulated individuals.

VIRTUAL REALITY - HARDWARE

- Variety of input and output devices are required to achieve virtual reality.
- One of the most important devices used is **tracker**. Tracker reports its location in space and its orientation. It can be combined with joystick/mouse.
- Another example is **data glove**. It is a device which contains sensors that can read the angle of each finger joints in the hand.
- The user can interact with the virtual world through hand gestures.

MACHINE LEARNING

- Machine Learning is a branch of AI that deals with the design and development of algorithms that allow computers to evolve behaviors based on empirical data (i.e. from sensors/databases).
- The major focus of the study is to make the system automatically learn from the available data and environment and train itself for making intelligent decisions.
- It thus helps to create a system with a lot of ease and reduced effort which is otherwise difficult to create manually.

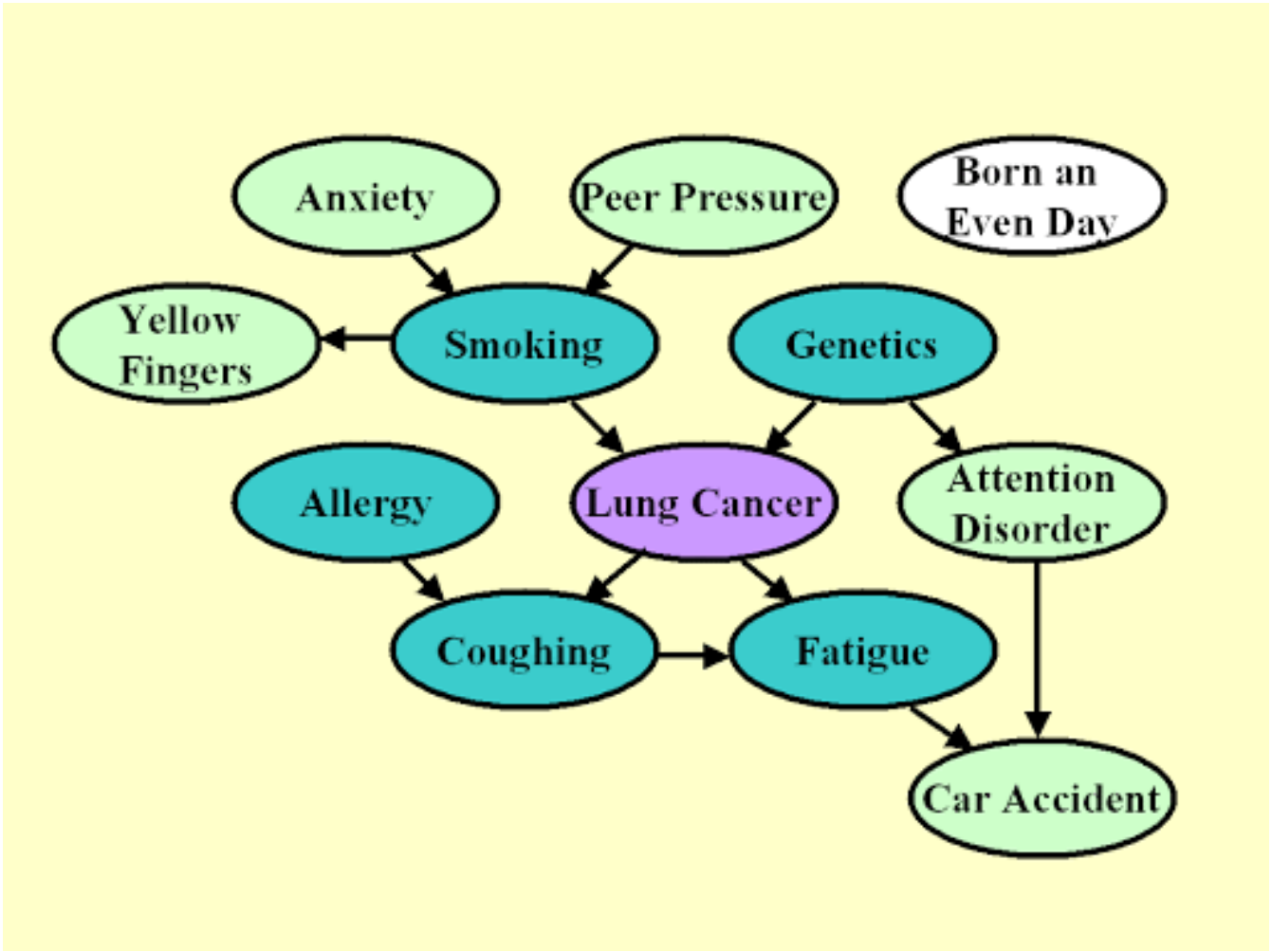
MACHINE LEARNING

- The difficulty lies in the fact that a set of all possible behaviors is too large to be covered by a set of observed samples.
- Decision tree learning, association rule learning, neural networks, inductive logic programming, fuzzy logic etc are few examples of learning algorithms.

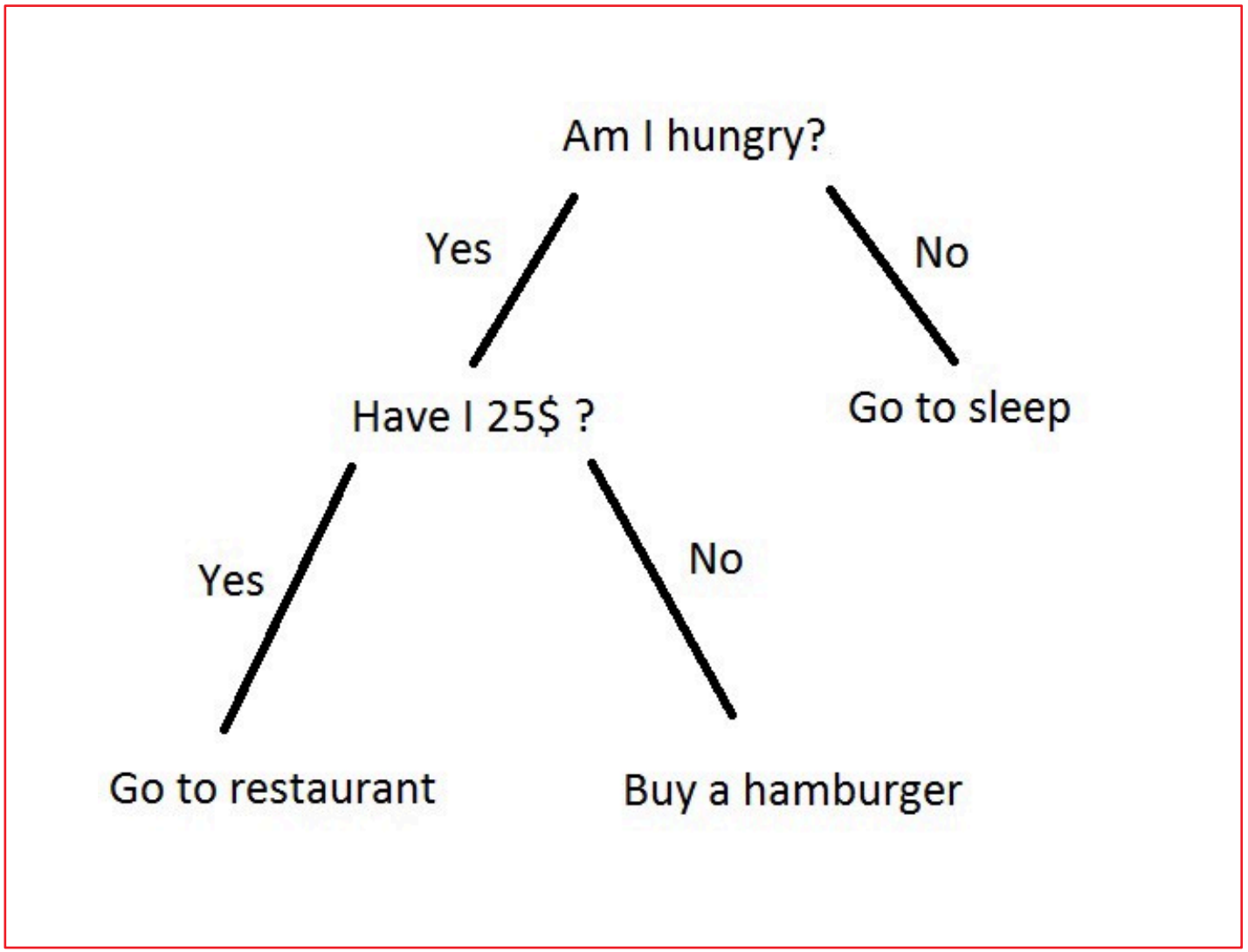
MACHINE LEARNING

- **Neural Networks:** ANN are modelled from the biological inspiration of human brain. It is a series of algorithms that endeavors to recognize underlying relationships and set of data through a process that mimics
- **Bayesian Networks**: It is a directed Acyclic graphical model that represents a set of variables and their conditional dependencies with the help of a graph.
- **Decision Tree:** This learning system uses a decision tree to go from observations about an item to conclusions about the item's target value.

BAYESIAN NETWORK



DECISION TREE



NATURE INSPIRED COMPUTING

- It is a computing technique that absorbs ideas by observing how nature behaves in various situations to solve problems.
- Research on this computing technique has opened new branches such as neural networks, artificial immune system, swarm intelligence and so on.

To Summarize…

- **AI tree**
- **Fruits:** Applications
- **Branches:** Expert Systems, NLP, speech understanding, robotics, computer vision, neural computing, fuzzy logic, genetic algorithms.
- **Roots:** Psychology, Philosophy, Electrical Engineering, Management science, Computer Science Linguistics

THANK YOU