Answer

Section A: Short Question in Artificial Intelligence

a. What is a heuristic function?

Ans. A heuristic function is a function that ranks alternatives in search algorithms at each branching step based on available information to decide which branch to follow.

---

b. Explain the role of computer vision in artificial intelligence.

Ans. Computer vision allows computers and systems to extract meaningful information from digital images, videos, and other visual inputs and then act or make recommendations based on that information.

---

c. What do you understand by feature extraction?

Ans. The process of transforming raw data into numerical features that can be processed while preserving the information in the original data set is referred to as feature extraction. It produces better results than directly applying machine learning to raw data.

---

d. What are the different properties of a task environment?

Ans. Properties of environments:

- 1. Discrete / Continuous
- 2. Observable / Partially observable
- 3. Dynamic / Static
- 4. Single agent / Multiple agents
- 5. Accessible / Inaccessible
- 6. Deterministic / Non-deterministic
- 7. Episodic / Non-episodic

---

e. Describe the first-order logic model.

Ans. Predicate logic, or first order predicate logic, is another name for first order logic. First order logic is a powerful language that can express the relationship between objects and develop information about them more easily.

---

f. Describe resolution in brief.

Ans. In predicate logic, resolution is a proof, a procedure that performs a single operation, or the variety or processes involved in reasoning with statements.

---

g. What is LISP? Why it is so popular in AI.

Ans. LISP is a programming language with an overall style based on expressions and functions. Every LISP procedure is a function that returns a data object as its value when called.

It is popular in AI because:

- 1. It is a machine-independent language.
- 2. It uses iterative design methodology and is easily extensible.
- 3. It allows us to create and update the programs and applications dynamically.

- 4. It provides high-level debugging.
- 5. It supports object-oriented programming.

---

h. What is the role of intelligent agent?

Ans. Perception and action are the two primary functions of intelligent agents. Sensors are used for perception, while actuators are used to initiate actions.

I. What is the need of classification techniques in pattern recognition?

Ans. To create a pattern recognition system, we must first select a model and prepare the data. The classification technique aids in the selection of a model and the preparation of data.

---

j. Differentiate supervised and unsupervised learning.

Ans.

S. No.	Supervised learning	Unsupervised learning
1.	Supervised learning algorithms are trained using labelled data.	Unsupervised learning algorithms are trained using unlabelled data.
2.	Supervised learning model predicts the output.	Unsupervised learning model finds the hidden patterns in data.

Section B: Long Questions of Artificial Intelligence

a. What is meant by the term artificial intelligence? How it is different from natural intelligence?

Ans. Artificial intelligence: Artificial Intelligence (AI) is an area of computer science that emphasizes the creation of intelligent machines that work and reacts like humans.

S. No.	Artificial intelligence	Natural intelligence
1.	Artificial intelligence is designed.	Natural intelligence is evolved.
2.	Artificial intelligence intends to create devices that can mock human behaviour.	Natural intelligence seeks to adapt to new environments by combining various cognitive processes.
3.	Artificial intelligence is dependent on the system receiving specific instructions.	Natural intelligence uses the brain's computing power, memory, and ability to think.
4.	Artificial intelligence can learn from data and through continuous training.	Natural intelligence is all about learning from past experiences and incidents.
5.	Artificial intelligence can work without stopping.	Natural intelligence needs rest and sleep.

---

b. Discuss branch and bound search algorithm.

Ans. The branch and bound strategy is a method for solving a series of subproblems, each of which has multiple possible solutions, and where the solution chosen for one subproblem affects the possible solutions of subsequent subproblems.

Principle:

1. Suppose it is required to minimize an objective function.

2. Assume we have a method for determining the lower bound on the cost of any solution in the set of solutions represented by some subset.

3. If the best solution found so far costs less than the lower bound for this subset:

Let Us be some subset for solution. Let

$L(S)$ = a lower bound on the cost of any solution belonging to S

C = cost of the best solution found so far

If $C \leq L(S)$, there is no need to explore S because it does not contain any

better solution.

If $C > L(S)$, then we need to explore S because it may contain a better solution.

---

c. What is the difference between knowledge representation and knowledge acquisition?

Ans. Knowledge representation:

- 1. Knowledge representation is the study of various methods of picturing knowledge and how well they resemble the representation of knowledge in the human brain.
- 2. Knowledge representation has two entities:
  - **a. Facts:** Facts are the truth in some relevant world.
  - **b. Representation:** Representation is the presentation of facts in some chosen formalism.
- 3. For example:
- Fact: Charlie is a dog.
- Representation of fact using mathematical logic: Dog (Charlie)
- 4. Knowledge representation should possess following characteristics:
  - a. Representation scheme should have a set of well-defined syntax and semantic.

- b. It should have good expressive capacity.
- c. It must be effective.

Knowledge acquisition:

- 1. Knowledge acquisition refers to the process of acquiring knowledge from a human expert for an expert system, which must be meticulously organised into IF-THEN rules or some other form of knowledge representation.
- 2. Knowledge acquisition is the process of absorbing and storing new information in memory, the success of which is determined by the ease with which the information can be retrieved later.
- 3. The process of storing and retrieving information is heavily reliant on the information's representation and organisation.

Procedure for knowledge acquisition:

- **1. Identification:** Break the problems into parts.
- **2. Conceptualisation:** Identify the concepts.
- **3. Formalisation:** Represent the knowledge.
- **4. Implementation:** Programming.
- **5. Testing:** Validate of knowledge.

---

d. Explain clustering and differentiate between supervised and unsupervised learning.

Ans. Clustering:

- 1. Clustering is the process or grouping of objects that are classified based on a close association or shared characteristics.
- 2. The objects can be physical or abstract entities, and the characteristics can be attribute values, object relationships, or a combination of the two.
- 3. Clustering is essentially a discovery and learning process in which similarity patterns among a group of objects are discovered.

Difference:

S. No.	Supervised learning	Unsupervised learning

|		
1.	Supervised learning algorithms are trained using labelled data.	Unsupervised learning algorithms are trained using unlabelled data.
2.	Supervised learning model takes direct feedback to check if it is predicting correct output or not.	Unsupervised learning model does not take any feedback.
3.	Supervised learning model predicts the output.	Unsupervised learning model finds the hidden patterns in data.
4.	In supervised learning, input data is provided to the model along with the output.	In unsupervised learning, only input data is provided to the model.
5.	The goal of supervised learning is to train the model so that it can predict the output when new data is provided.	Unsupervised learning seeks to discover hidden patterns and useful insights in an unknown dataset.
6.	Supervised learning needs supervision to train the model.	Unsupervised learning does not need any supervision to train the model.
7.	Supervised learning can be categorized in classification and regression problems.	Unsupervised learning can be classified in clustering and associations problems.

e. Describe the N-queens problem in detail. Also explain with an example.

Ans. N-queens problem:

- 1. N-Queen problem is based on chess games.
- 2. The problem is based on arranging the queens on chessboard in such a way that no two queens can attack each other.

- 3. The N-Queen problem states as consider a n x n chessboard on which we have to place in queens so that no two queens attack each other by being in the same row or in the same column or on the same diagonal.

- 4. 2-Queen's problem is not solvable because 2-Queens can be placed on 2 x 2 chess board as shown in Fig. 1.

Q	Q
Ilegal

Q	
Q	
Ilegal

Q	
	Q
Ilegal

Q	Q
Ilegal

	Q
	Q
Ilegal
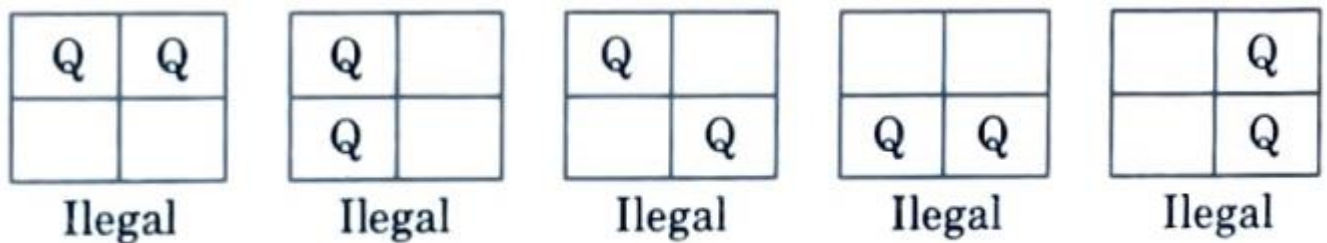
Fig. 1.

5. But 4-Queen's problem is solvable:

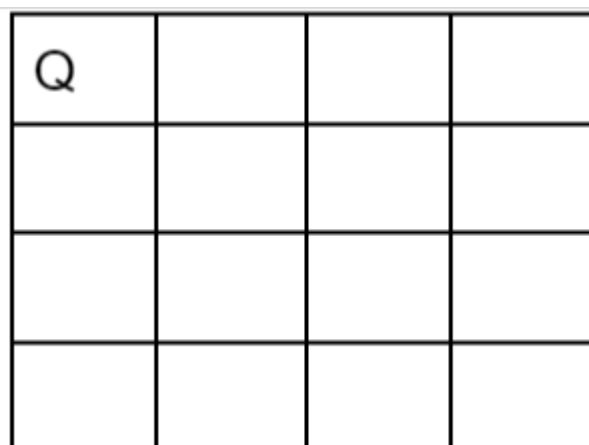|		Q	
Q			
			Q
	Q		

< = Note that no two queens can attack each other

Fig. 2.

To solve N- Queen problem:

I. Let us take the example of 4 – Queens and 4 x 4 chessboard.

ii. Start with an empty chessboard.

iii. Place queen 1 in the first possible position of its row i.e., on 1st row and 1st column.

Q			

iv. Then place queen 2 after trying unsuccessful place (1, 2), (2, 1), (2, 2) at (2, 3) i.e., on 2nd row and 3rd column.

Q			
		Q	

v. This is a dead end because a 3rd queen cannot be placed in next column, as there is no acceptable position for queen 3. Hence algorithm backtracks and places the 2nd queen at (2,4) position.

Q			
			Q

vi. Then place 3rd queen at (3, 2) but it again another dead lock end as next queen (4th queen) cannot be placed at permissible position.

Q			
			Q
	Q		

vii. Hence we need to backtrack all the way up to queen 1 and move it to (1, 2).

vi. Place queen 1 at (1, 2), queen 2 at (2, 4), queen 2 at (3, 1) and queen 4 at (4, 3).

ix. Thus the solution is obtained (2, 4, 1, 3) in row wise manner.

x. The one of the solutions to 8 – Queen problem is shown below:

xi. The solution is obtained (6, 4, 7, 1,3, 5,2, 8) in row wise manner.

Section 3: Natural Language Processing

a. What do you understand by natural language processing?

Ans.

- 1. Natural Language Processing (NLP) is the study of the problems inherent in the processing and manipulation of natural language in order to make computers understand human-written statements.

- 2. Natural language processing (NLP) is the automatic processing of human language.

- 3. Natural language processing is an AI subfield that deals with methods of communicating with a computer in one's native language.

- 4. It is used to analyse and represent naturally occurring texts at one or more linguistic levels in order to achieve human-like language processing for a variety of tasks or applications.

- 5. It is necessary to bridge the human-machine divide.

- 6. The goal of natural language is for people and computers to be able to communicate in a natural language, such as English.

- 7. The field of NLP is divided into subfields:

  - **a. NLU (Natural Language Understanding):** This investigates methods of allowing the computer to comprehend instructions given in English.

  - **b. NLG (Natural Language Generation):** This strive that computer produce ordinary English language so that people can understand computers more easily.

- 8. The study of language generation falls into following three areas:

  - a. Determination of content.

  - b. Formulating and developing a text utterance plan.

  - c. Achieving a realization of the desired utterances.

- 9. A full NLU system would be able to:

  - a. Paraphrase an input text.

  - b. Translate the text into another language.

  - c. Answer questions about the contents of the text.

  - d. Draw inferences from the text.

- 10. Applications of NLP:

  - a. Natural language interfaces to databases.

  - b. Machine translation.

  - c. Advanced word-processing tools.

  - d. Explanation generation for expert systems.

---

b. Explain the concept of alpha-beta pruning. Write alpha beta search algorithm.

Ans. Alpha-beta pruning:

- 1. Alpha-beta pruning is a modified version of the min-max algorithm. It is an optimization technique for the min-max algorithm.

- 2. The number of nodes (game states) that must be examined in the min-max search algorithm is proportional to the depth of the tree. We cannot completely eliminate the exponent, but we can reduce it by half.

- 3. There is a technique called pruning that allows us to compute the correct min-max decision without having to check each node of the game tree. Because this involves two threshold parameters, alpha and beta, for future expansion, it is referred to as alpha-beta pruning.

- 4. Alpha-beta pruning can be used at any depth in a tree, and it sometimes prunes not only the tree leaves but also the entire sub-tree.

- 5. The two parameters can be defined as:

  - **a. Alpha:** The best (highest-value) choice we have found so far at any point along the path of Maximiser. The initial value of alpha is $-\infty$.

  - **b. Beta:** The best (lowest-value) choice we have found so far at any point along the path of Minimizer. The initial value of beta is $+\infty$.

Section 4: Simulated Annealing Search Algorithm

a. Differentiate between local search and global search.

Ans.

S. No.	Local search	Global search
1.	Local search is the ability to search for an optimal result within a neighbourhood.	Global search is the ability to find a global solution by searching beyond the local optimal region.
2.	Performance of a local search technique is dependent on its initial position(s).	Performance of a global search technique is less dependent on its initial position(s).

3.	It is for narrow problems where the global solution is required.	It is for broad problems where the global optima might be intractable.
4.	Local search algorithms frequently grant computational complexity grants in relation to locating global optima.	Global search algorithms frequently provide very few guarantees regarding the location of the global optima.
5.	Examples of this search include, Needler-Mead algorithm, BFGS algorithm, Hill-Climbing algorithm.	Examples of this search include, particle swarm optimization, simulated annealing, and genetic algorithm.

---

b. Discuss simulated annealing search algorithm with its advantages and disadvantages.

Ans.

- 1. A stochastic global search optimisation algorithm is simulated annealing.
- 2. Simulated annealing is a metallurgy-based algorithm based on the physical annealing process.
- 3. During physical annealing, the metal is heated to its annealing temperature and then gradually cooled to change its shape.
- 4. Simulated annealing is used to optimise model parameters and has a probabilistic way of moving around in a search space.
- 5. It mimics physical annealing as a temperature parameter is used here too.
- 6. If the temperature is higher, the more likely the algorithm will accept a worse solution.
- 7. This expands the search space.
- 8. If the temperature is lower, the less likely it will accept a worse solution.
- 9. This tells the algorithm to focus on finding the global maximum by converging.

Algorithm:

- 1. Start
  - I. Initialize k= 0; L = integer number of variables;

- ii. From I → j, search the performance difference Δ.
- iii. If Δ < = 0 then accept else if exp (- Δ/T(k)) > random (0, 1)) then accept;
- iv. Repeat steps 1 and 2 for L(k) steps.
- v. k = k + 1;
- 2. Repeat steps 1 through 4 till the criteria are met.
- 3. End

Advantages:

- 1. Simulated annealing is easy to code and use.
- 2. It does not rely on restrictive properties of the model and hence is versatile.
- 3. It can deal with noisy data and highly non-linear models.
- 4. Provides optimal solution for many problems and is robust.

Disadvantages:

- 1. A lot of parameters have to be tuned as it is metaheuristic.
- 2. The precision of the numbers used in its implementation has a significant effect on the quality of results.
- 3. There is a trade-off between the quality of result and the time taken for the algorithm to run.

Section 5: Forward Chaining and Backward Chaining

a. Define forward chaining and backward chaining with examples.

Ans. Forward chaining:

- 1. Forward chaining is a method of reasoning used in artificial intelligence when using inference rules.
- 2. Forward chaining begins with available data and then applies inference rules to extract more data (from an end user) until an optimal goal is reached.
- 3. A forward chaining inference engine searches the inference rules until it finds one with the If clause known to be true.
- 4. When it finds the Then clause, it can conclude or infer it, resulting in the addition of new information to its dataset.

- 5. Inference engines will often cycle through this process until an optimal goal is reached.
- 6. For example, suppose that the goal is to conclude the colour of my pet Bruno given that he croaks and eats flies, and that the rule base contains the following two rules:
  - a. If X croaks and eats flies – Then X is a frog.
  - b. If X is a frog – Then X is red.

Backward chaining:

- 1. Backward chaining begins with a list of goals (or a hypothesis) and works backwards to see if any of these goals can be supported by data.
- 2. A backward chaining inference engine would search the inference rules until it found one with a Then clause that matches a desired goal.
- 3. It is added to the list of goals if the If clause of that inference rule is not known to be true.
- 4. For example, suppose that the goal is to conclude the colour of my pet Bruno given that he croaks and eats flies, and that the rule base contains the following two rules:
  - a. If X croaks and eats flies – Then X is a frog.
  - b. If X is a frog – Then X is red.

---

b. Discuss various application domains of machine learning.

Ans. Application domains of machine learning:

- **1. Image Recognition:** It is used to identify objects, persons, places, digital images, etc.
- **2. Speech Recognition:** When we use Google, we have the option of searching by voice, which falls under speech recognition. Speech recognition, also known as Speech to text or Computer speech recognition, is the process of converting voice instructions into text.
- **3. Traffic prediction:** If we want to go somewhere new, we use Google Maps, which shows us the best path with the shortest route and predicts traffic conditions.
- **4. Product recommendations:** Machine learning is widely used for product recommendation by various e-commerce and entertainment companies such as Amazon, Netflix, and others.

- **5. Self-driving cars:** Machine learning is important in self-driving cars. Tesla, the most well-known car manufacturer, is developing a self-driving car. It trains the car models to detect people and objects while driving using an unsupervised learning method.
- **6. Stock Market trading:** In the stock market, machine learning is widely used. Because there is always the risk of share price fluctuations in the stock market, this machine learning's long short-term memory neural network is used to forecast stock market trends.

Section 6: Pattern Recognition

a. Write short notes on support vector machine.

Ans.

- 1. A Support Vector Machine (SVM) is a machine learning algorithm for data classification and regression analysis.
- 2. SVM is a supervised learning method that examines data and categorizes it into one of two groups.
- 3. An SVM produces a map of the sorted data with as wide a margin between the two as possible.
- 4. The SVM algorithm's goal is to find a hyperplane in an N-dimensional space that clearly classifies the data points.
- 5. The dimension of the hyperplane depends upon the number of features.
- 6. If the number of input features is two, then the hyperplane is just a line.
- 7. If the number of input features is three, then the hyperplane becomes a 2-D plane.
- 8. Applications of SVM:
    - I. Text and hypertext classification
    - ii. Image classification.
    - iii. Recognizing handwritten characters.
   - iv. Biological sciences, including protein classification.

---

b. What is pattern recognition? Explain various steps involved in the designing of a pattern recognition system with the help of a diagram.

Ans. Pattern recognition:

- 1. Pattern recognition is the process of identifying patterns through the use of a machine learning algorithm.

- 2. Pattern recognition is the classification of data based on prior knowledge or statistical information extracted from patterns and/or their representation.

- 3. In a pattern recognition application, raw data is processed and converted into a machine-readable format.

- 4. Examples: Speech recognition, speaker identification, Multimedia Document Recognition (MDR), automatic medical diagnosis.

Steps involved in designing pattern recognition system:

- **1. Classification:** An appropriate class label is assigned to a pattern in classification based on an abstraction generated by a set of training patterns or domain knowledge. Supervised learning makes use of classification.
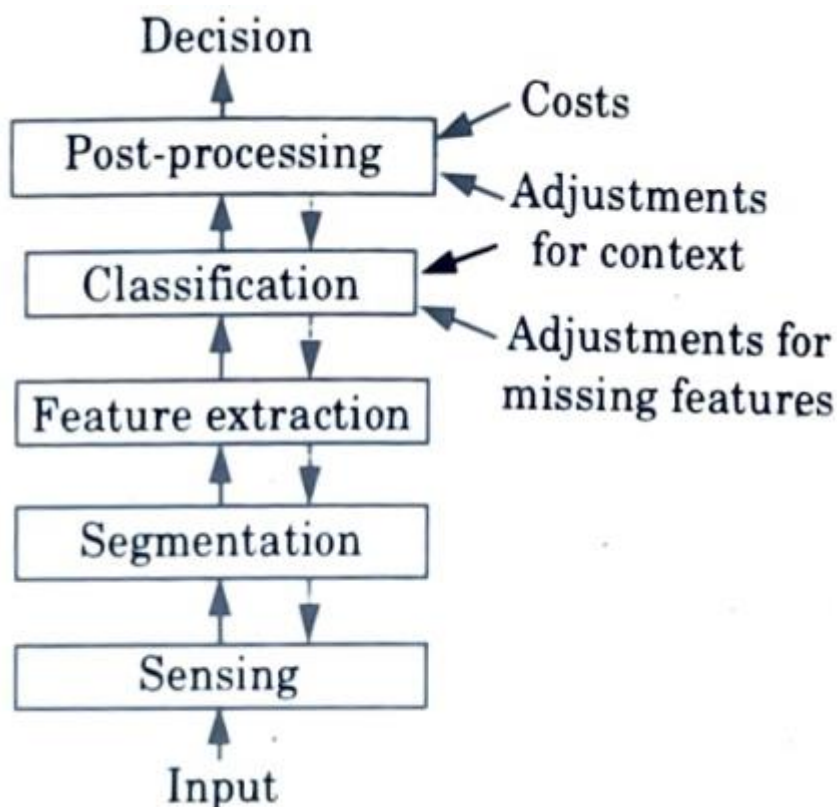
- **2. Clustering:** It generates a partition of the data which helps decision making, the specific decision-making activity of interest to us.

Clustering is used in an unsupervised learning:

- **a. Sensing:**

- I. A sensor is a device that converts images, sounds, or other physical inputs into signal data.
- ii. A pattern recognition system's input is frequently a transducer, such as a camera or a microphone array.

- **b. Segmentation and grouping:**
  - I. The process of dividing a digital image or speech into multiple segments is referred to as segmentation.
  - ii. The segment or separates sensed objects from their surroundings or other objects.
- **3. Feature extraction:**
  - I. A feature extractor measures object properties that are useful for classification.
- **4. Classification:**
  - I. The classifier component of a full system's task is to assign the object to a category using the feature vector provided by the feature extractor.
- **5. Post processor:**
  - I. The post processor uses the output of the classifier to decide on the recommended action.

Section 7: Bayesian Classifier

a. Write short notes on Bayesian classifier.

Ans.

- 1. A Bayesian classifier is a simple probabilistic classifier that is based on the Bayes theorem and strong independence assumptions.
- 2. A Bayesian classifier is based on the idea that a class's role is to predict the values of features for its members.
- 3. Examples are grouped into classes because they share feature values.
- 4. These classes are frequently referred to as natural kinds.
- 5. A Bayesian classifier is based on the idea that if an agent knows the class, it can predict the values of the other features.
- 6. If it does not know the class, it can use Bayes' rule to predict the class based on the feature values.

- 7. The learning agent in a Bayesian classifier creates a probabilistic model of the features and uses that model to predict the classification of a new example.

- 8. A Bayesian classifier is a probabilistic model in which the classification is a latent variable that is related to the observed variables in a probabilistic way.

---

b. Write short on LDA and PCA.

Ans. LDA:

- 1. Linear Discriminant Analysis (LDA) is a technique for data classification and dimensionality reduction that is widely used.

- 2. LDA handles the case where the values within the class frequencies are unequal and their performances have been tested on randomly generated test data with ease.

- 3. LDA works when the measurements on independent variables are continuous quantities for each observation.

- 4. The use of LDA for data classification is applied to a speech recognition classification problem.

PCA:

- 1. PCA is one of the most commonly used statistical techniques. Dimensionality is reduced using the Principal Component Analysis (PCA) technique.

- 2. A higher dimensional data space can be transformed into a lower dimensional space using the PCA technique. This is also known as Hoteling transformation.

- 3. PCA linearly transforms a high-dimensional input vector into a low-dimensional one with uncorrelated components by calculating Eigen vectors of the original inputs' covariance matrix.

- 4. The primary benefits of PCA are the reduction of data set dimensionality and the identification of new meaningful underlying variables.