

FINAL EXAM

Academic Year: ODD Semester (2024-1)

Subject : Artificial Intelligence

Lecturer : Cutifa Safitri, Ph.D

Student Name / ID : Michel Stanley/001202300106

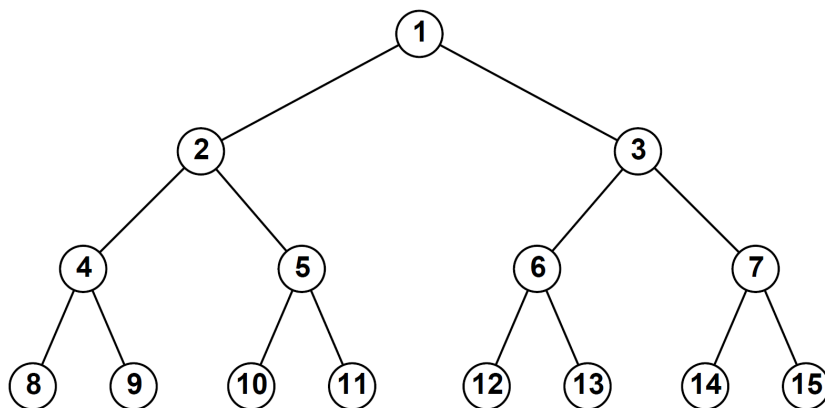
Instruction to Student:

1. This examination consists of 9 question(s).
2. Time allowed for this examination is ____ hour(s).
3. Sanctions will be given to those students who are not following the examination rules.
4. All answers to be written in answer sheet provided. The number to the corresponding question must be written correctly.
5. Students are to use proper English and are required to write neatly and clearly.
6. This is a closed book (an open book) examination.
7. Students are (not) allowed to use calculator.
8. Students are (not) allowed to use any dictionary.
9. Students need to turn-off any communication devices.
10. Students are not allowed to use electronic devices or borrow any equipment, etc. from fellow students, nor can they communicate with each other while the examination is going on.

Students are not allowed to communicate with each other while the midterm is on going.

Answer all questions in this paper.

1. Refer to this figure for questions below. Suppose the goal state is 11



List the order in which nodes will be visited for breadth-first search

1-2-3-4-5-6-7-8-9-10-11

List the order in which nodes will be visited for depth-first search

1-2-4-8-9-5-10-11

List the order in which nodes will be visited for depth-limited search with limit 3

1-2-4-8-9-5-10-11

List the order in which nodes will be visited for iterative deepening search.

1-2-3-4-5-6-7-8-9-10-11(depth 3)

2. Explain the process of data ingestion, data preparation and data segregation in machine learning pipeline.

ANSWER:

data ingestion

Data ingestion is the process of inputting the data into the machine learning pipeline either in a form of database, files, etc. After obtaining the data, the data is then formatted into a readable format for example a table. The data is then stored within the memory

Data preparation

When inserting a large batch of data for the machine, we cannot check the data one by one and correct/proofread the entirety of our data. Data preparation is a process that does this for us by cleaning several errors such as a missing value, duplicate value, outliers, and error value that might give a mismatch to the data format. There are multiple way to handle these error such as complete deletion, auto fill by trend, or leave it as null

data segregation

After collecting and processing the data, data segregation will split the data into different subsets for training, validating, and testing in order to evaluate and optimize the model. This data is usually split with almost 70% used for training and the rest split equally for both validation and testing.

If some of the answers are taken from the Internet; provide the source here: [link](#)

3. Two of the uninformed searches we learn so far are breadth-first search (BFS) and depth-first search (DFS). Define their characteristics. Which of these two are more optimized and why?

ANSWER:

BFS is a traversing algorithm that walks through each node within the current layer it's on before moving on to the next layer. DFS however, is a traversing algorithm that starts from the root and goes into the leaves of the first node until there are no more nodes before returning one step

backwards and checks the same thing for the next children and so on until the target is found. As for optimization, both algorithms have almost the same chance depending on the target and the structure of the nodes since both of the algorithms are an uninformed search algorithm, making it essentially a brute force method. In a specific scenario each method can shine on its own, for example a BFS is good to find the shortest path and DFS is good within a deep tree. BFS can be applied into GPS and DFS can be applied into web crawling or file system navigation

4. Two of the informed searches we learn so far are Greedy best-first search (GBFS) and A*. Define their characteristics. Which of these two are more optimized and why?

ANSWER:

GBFS is a traversing algorithm that applies the concept of heuristic where it'll decide based on the most optimal distance of the next node against the target where A* is almost a similar case but it puts the cost of travelling onto the next node within consideration. Between the two, A* is the most optimized one as A* is enhanced with the overall cost instead of solely focusing on minimizing the heuristic. This allows for the most optimal answer and prevents cycling the same path.

5. Two of the Evolutionary algorithms searches we learn so far are Genetic Algorithm and Genetic Programming. Define the characteristics, possible advantage and disadvantages of their technique.

ANSWER:

Genetic algorithm and genetic programming are both algorithms that are taken off the concept of how evolution and genetics works. Genetic algorithm is an optimization method that operates within a candidate of potential solution. These values will go through multiple processes such as natural selection, crossover, and mutation in order to achieve an optimized solution. The advantages of this algorithm is creating a very optimized solution and flexibility when implemented into another problem as the individual candidate can be manipulated into any data type such as string, array, bits, etc.

Genetic programming on the other hand is also an optimization method but it solely focuses on the program rather than a parameterized solution. The advantages of this method is the provided solution is often readable by a human and often able to handle a problem that traditional programming might not be able to solve.

Both disadvantages of these algorithms however are similar. Both algorithms have a quite high computational cost, making it often slow and expensive on the progress. Moreover, parameter tuning for mutation, crossover and a generic fitness function is often challenging as you're required to find a good method to slowly progress onto the desired outcomes.

If some of the answers are taken from the Internet; provide the source here: [link](#)

6. Given a full 4-gallon jug and a full 3-gallon jug with no marker and an unlimited number of water, fill the 4-gallon jug with exactly 2-gallon of water. Show the Initial State, Transition State, Operators, Goal state and Actions.



ANSWER:

initial state (gallon1,gallon2) \rightarrow (0,0) with the max of (3,4)

operators \rightarrow Fill gallon 1

\rightarrow Pour water between each gallon

\rightarrow empty either gallon

Goal state \rightarrow (gallon1,gallon2) \rightarrow (any,2)

actions = fill gallon 1 (3,0) \rightarrow pour gallon 1 into gallon 2 (0,3) \rightarrow fill gallon 1 (3,3) \rightarrow pour gallon 1 into gallon 2 (2,4) \rightarrow empty gallon 2 (2,0) \rightarrow pour gallon 1 into gallon 2 (0,2)

7. Explain what are the relationship between entropy and knowledge?

ANSWER:

Entropy and knowledge are pretty related. Entropy is the measurement of impurity or uncertainty within a data set. It is the randomness of the data that exists within the machine, essentially an identifier for groups and a parameter to see whether your data is properly segmented or not. a lower entropy would indicate that the data is properly segmented/grouped and a higher entropy would indicate a more unstable data with the chances of overlapping or an outlier.

Knowledge on the other hand is a certain fact within the system or data. it is a determining factor that might reduces an entropy by giving us the behavior of our data or system in order to increase the certainty of the result

In short term, the relationship between knowledge and entropy is a parallel inverse relationship where the more knowledge you have on a certain system, the lesser the entropy will be and vice versa.

If some of the answers are taken from the Internet; provide the source here:[link](#)

Please refer to Article that has been posted at :

<https://spectrum.ieee.org/tech-talk/artificial-intelligence/machine-learning/deepminds-ai-shows-itself-to-be-a-worldbeating-world-builder> for questions 8 and 9.

Recall on Games: In which we examine the problems that arise when we try to plan ahead in a world where other agents are planning against us.

Quoting the articles; "To win at StarCraft II, a player builds an empire with all forethought and flexibility such an endeavor requires. Players must weigh the importance of competing objectives—like gathering resources, building structures, organizing an army, setting up defenses, and fighting battles—and shift their relative importance in real-time over the course of a game that could last an hour or more."

8. Based on the article above, discuss the agent implementations under the perspectives of **deterministic** or **chance / stochastic agents**.

ANSWER:

The agent within the article operates both deterministically and stochastically. It is deterministic due to the fact that TLO attempted to adapt to the agent's "first favorite move" which required the DeepMind group to switch between agents to adapt into their strategy. Although it can be deterministic at times, this agent is heavily required to be stochastic as StarCraft II requires "forethought and flexibility" which will "shift the relative importance". This means the agent adapts through different types of environments and switches strategy depending on its opponent to counter their strategy. This combination allows the agent to achieve a high level of performance, surpassing even the best of their crafts as shown by the 5/0 loss against TLO and 4/1 loss against MaNa

If some of the answers are taken from the Internet; provide the source here:[link](#)

9. Based on the article above, discuss the agent implementations under the environment perspectives of **perfect information (Informed Search)** or **imperfect information (Uninformed Search)**

Based on your answer, proposed the search algorithm that is suitable based on your answer at 8 and 9

ANSWER:

given by the nature of the game itself, the agent originally starts with an imperfect information hinted at "only a portion of the landscape in the game's fictional world is visible at any given time, so the odds of winning are greatly affected by the player's memory and ability to set up things that won't be continuously monitored.". This information however, provides us with an additional statement that given enough resources and memory, the agent is able to gain a full overview of the environment which leads to its victory. In conclusion, Starcraft II highly consists of imperfect information, requiring agents to handle uncertainty and dynamically adapt to multiple outcomes, slowly learning and fitting itself within the environment optimally in order to complete a complex real time strategy game.

If some of the answers are taken from the Internet; provide the source here:[link](#)

A good search algorithm that fits two of the description above would be A* search algorithm. A* is a good algorithm for strategic games which includes exploration and exploitation of a path through the evaluation of heuristic estimate to the goal and the cost to reach the next node. It is also one of the greatest informed search algorithms as time progresses with how much information it will collect and store within its memory. Therefore, a good search algorithm for a game in which strategy is required yet almost the entirety of the environment is initially unknown, A* a search algorithm would perform best in slowly adapting to the game and create an effective strategy.

THIS IS THE END OF THE FINAL EXAM QUESTIONS