CCS3440-Artificial Intelligence

Assignment 01

Answer all Questions

Student Name: Sahan Hansaja

Student Registration Number: 22ug2-0035

Due date: on or before 13^{th} of December 2024

1. A factory floor robot is tasked with assembling parts while navigating obstacles to reach specific assembly stations.

   a. Explain how depth-limited search (DLS) can prevent the robot from entering infinite loops.

      DLS stops the robot from going too deep into paths by setting a maximum depth it can search. This way, if the robot encounters a loop (like going back and forth between two places), it won't keep going forever because the depth limit will stop it.
      It prevents if that have circle path or loop. For example, in this case the factory robot moves some limited floors image many floors have that factory but robot can two floors only because depth-limited set to two.

   b. Describe how iterative deepening search (IDS) combines the advantages of BFS and DFS for this task.

- IDS do the following:
  - It searches like Depth-First Search (DFS), going deep into paths, but only up to a certain depth.
  - Then it increases the depth limit and searches again.
  - It repeats this process until it finds the goal.

- This combines:
  - Since it does not retain every level, DFS has a relatively low memory requirement.
  - DFS spends more time checking outermost level's nodes but rarely goes to the deepest levels like the transverse BFS which can be convenient because it minimizes the number of nodes of interest and expands the level one at a time to attain a desired distance from the current source node.
  - To the robot, it is simply going to nearby areas first (similar to BFS) while conserving resource memory (similar to DFS).

   c. Suggest an improvement to uninformed search strategies to make them more efficient for large-scale problems.

Uninformed search checks every path, which takes a lot of time. To improve:

- Use a guide (heuristics): Add clues like "Which path looks closer to the goal?" For the robot, it could choose paths that seem closer to the assembly station.
- Search from both sides: Start from the robot and the goal at the same time and meet in the middle.
- Avoid repeating work: Remember places already checked so the robot doesn't waste time re-checking them.

This makes the search faster and smarter without checking everything randomly.

[Total = 20 Marks]

2. A city plans to deploy AI-powered surveillance drones to monitor crowded areas during events for safety purposes.

    a) What levels of AI intelligence (computational, perceptual, cognitive) would the drones need to identify and respond to potential threats?

    - Computational Intelligence:
        - Necessary for real-time data processing, path planning, and obstacle avoidance.
        - Example: Calculating efficient flight paths to navigate crowded areas.
    - Perceptual Intelligence:
        - Required to interpret visual and audio inputs from cameras and sensors.
        - Example: Detecting suspicious objects or analyzing crowd behavior for unusual activities.
    - Cognitive Intelligence:
        - Needed to analyze complex situations and make decisions.
        - Example: Differentiating between harmless actions (e.g., people waving) and threats (e.g., someone carrying a weapon).

b) How would you ensure the AI adheres to ethical standards, such as privacy and data security, during operations?

- Privacy:
    - Limit the data collected to what's strictly necessary.
    - Use techniques like face blurring to anonymize individuals unless a confirmed threat exists.

- Data Security:
    - Encrypt all data during storage and transmission to prevent breaches.
    - Regular system updates ensure protection against cyberattacks.

- Accountability:
    - Maintain a human-in-the-loop system where critical AI decisions are reviewed by humans.
    - Establish clear guidelines for drone deployment and use.

- Transparency:
    - Inform the public about the scope and purpose of drone operations.
    - Publish regular audits to ensure compliance with ethical standards.

- Bias Reduction:
    - Train the AI on diverse datasets to avoid discriminatory or biased decisions.
    - Periodically review and improve the system's performance.

c) Propose a method for testing the system's reliability in recognizing non-standard scenarios, such as detecting objects in poor lighting or adverse weather conditions.

- Simulations:
    - Test the AI in virtual environments that simulate poor lighting, heavy rain, fog, or crowded areas.
- Controlled Real-World Testing:
    - Conduct field trials in adverse conditions such as nighttime, rain, or high winds.
    - Introduce unique test cases like unusual objects, non-standard crowd behaviors, or rare clothing styles.
- Stress Testing:

        - Overload the system by simulating multiple tasks simultaneously (e.g., detecting threats in a chaotic crowd) to evaluate its limits.
    - Feedback and Retraining:
        - Use insights from testing to retrain and improve the AI model’s ability to handle complex scenarios.

By combining these methods, the drones can be better equipped to operate effectively in unpredictable situations.

[Total = 20 Marks]

3. An AI-powered chatbot is being developed to assist in mental health therapy sessions.

    a) How could cognitive modelling be used to ensure the chatbot interacts with patients in a human-like and empathetic manner?

    Cognitive modeling is an approach to making a chatbot interact in a human-like way by taking a simulation of how a human thinks and responds on emotional levels. For example, this might use NLP for emotion detection in what is being said or written, hence adapting its tone toward providing supportive or encouraging responses when it sees fit. Cognitive Behavioral Therapy techniques may also be embedded within the framework of the chatbot to help it guide users through structured problem-solving or thought reframing processes. Moreover, training the chatbot on real-life therapeutic conversations improves its ability to maintain a natural flow, reflect on user input, and ask relevant questions. Finally, memory models can help the chatbot recall details from past sessions, fostering a sense of continuity and trust in its interactions.

    b) Suggest ways to measure the chatbot's performance using variations of the Turing Test.

    The chatbot’s performance can be evaluated through adapted Turing Tests. One method is to compare its interactions with those of a human therapist in a blind test, rating empathy and helpfulness. Another is to assess its ability to recognize emotions and respond appropriately. It can also be tested on solving identical mental health scenarios alongside a human therapist. User feedback and surveys can measure conversational quality, while long-term studies can track its impact on mental health compared to traditional therapy.

c) If the chatbot gives incorrect or harmful advice during a session, what steps would you take to improve its reasoning and adaptability?

If the chatbot gives harmful advice, its reasoning can be improved by reviewing incidents to identify errors and updating its training data with correct examples. Ethical safeguards and real-time feedback can help refine responses, while human therapists can review flagged cases. Using reinforcement learning, the chatbot can be trained to reward helpful advice and avoid mistakes. Sensitive situations can be referred to human professionals, and regular updates can ensure better accuracy and reliability.

[Total = 20 Marks]

4. An AI assistant is designed to help users manage their daily schedules and tasks.

a) What are the key differences between implementing this AI as a simple reflex agent versus a utility-based agent?

A simple reflex agent: works by following fixed rules or conditions, responding to tasks as they occur without evaluating their importance or long-term impact. For example, it might send a reminder for a task without considering whether it is urgent or relevant. This approach is easy to implement but lacks the ability to adapt, prioritize, or learn from past actions, making it less effective for complex scheduling needs.

A utility-based agent : on the other hand, evaluates tasks by considering factors like urgency, importance, and user preferences. It uses a utility function to calculate the best outcome and prioritize tasks accordingly. This allows it to make smarter, more personalized decisions, balancing competing priorities effectively. While more complex, a utility-based agent offers greater flexibility and better performance in managing tasks.

b) **If the assistant needs to predict the user's preferences for task prioritization, what learning components (critic, learning element, performance element, problem generator) would be essential?**

To predict user preferences for task prioritization, the following components are essential:

- Critic: Evaluates the assistant's performance by comparing predicted priorities with actual user actions or feedback. This ensures continuous improvement.
- Learning Element: Adjusts the assistant's algorithms based on user interactions and feedback to constantly improve predictions.
- Performance Element: Performs task prioritization based on current knowledge and predictions to optimize the schedule of a user.
- Problem Generator: Suggests new strategies or experiments with different prioritization methods to gather data and improve future performance.

These components work together to ensure the assistant adapts to user preferences while keeping efficiency intact.

c) **How would you design the assistant to operate in an episodic versus a sequential environment?**

In an episodic environment, tasks are independent, and the assistant handles each one separately without needing context from past or future interactions. For example, it can send reminders for single tasks without considering other tasks.

In a sequential environment, tasks are connected, and decisions affect future outcomes. The assistant must track past interactions, learn patterns, and plan to avoid conflicts, such as handling postponed tasks.

To work in both settings, the assistant should use memory for sequential tasks and simple rules for episodic ones. This ensures it can handle both individual tasks and more complex, connected schedules efficiently.

[Total = 20 Marks]

5. Analyze the task environment of a self-driving car.

a) Classify the environment using these criteria: fully or partially observable, single or multi-agent, deterministic or nondeterministic, episodic or sequential, static or dynamic, and discrete or continuous. Provide clear reasons for each classification.

Fully or Partially Observable: The environment is partially observable because the car can't see everything (e.g., other drivers' intentions or hidden obstacles).

Single or Multi-Agent: The environment is multi-agent since the car shares the road with other vehicles, pedestrians, and cyclists.

Deterministic or Nondeterministic: The environment is nondeterministic because human drivers are unpredictable, and road conditions can change.

Episodic or Sequential: The environment is sequential because each action (like braking or steering) affects future decisions.

Static or Dynamic: The environment is dynamic since conditions like traffic and weather change constantly.

Discrete or Continuous: The environment is continuous because the car must handle continuous changes in speed, position, and distance.

b) Discuss the primary challenges a dynamic and nondeterministic environment creates for a selfdriving car.

Unpredictable Behavior: The car must quickly react to sudden changes, like other drivers' unexpected moves or pedestrians crossing, making it hard to predict what will happen next.

Real-Time Decision Making: The car must process new information constantly and adapt to changes right away to avoid accidents.

Risk of Collisions: Since other drivers' actions are unpredictable, the car must decide how to handle situations like sudden stops or someone cutting in front.

Complex Sensor Interpretation: The car's sensors might give incomplete or unclear data, making it harder to interpret and make decisions.

c) Suggest how a model-based reflex agent could handle these challenges effectively.

A model-based reflex agent can effectively handle the challenges of a dynamic and nondeterministic environment by maintaining an internal model of the world, which is updated continuously based on real-time sensor inputs. This model helps the car track its position, the positions of other agents, road conditions, and obstacles, allowing for more informed decision-making. The agent also uses predefined condition-action rules to react quickly to immediate changes. For instance, if an obstacle is detected, the agent may follow a rule to stop or slow down. To handle uncertainty in a nondeterministic environment, the agent can use probabilistic reasoning or decision trees to assess possible outcomes, such as taking precautionary actions when a pedestrian is likely to cross. Additionally, the agent's model of the world is constantly updated, enabling it to adapt to changing conditions like traffic or weather. By combining a model-based approach with real-time data, the agent can swiftly respond to unpredictable events while ensuring safety and relevance in its decisions.

[Total = 20 Marks]

End of Question Paper