Exam: AI Planning for Autonomy (COMP90054_2021_SM2)

Started: Nov 16 at 15:12

Quiz Instructions

The University of Melbourne
School of Computing and Information Systems

Final Exam, Semester 2, 2021
COMP90054 AI Planning for Autonomy

Duration: 150 minutes

Please note that you are permitted to write answers immediately, during the reading time, as this is not enforced.

Instructions to Students:

The test includes questions worth a total of 40 marks, making up 40% of the total assessment for the subject.

- This exam includes a combination of short-answer, long-answer, multiple-choice, and fill-in-the-blank questions. Please answer all questions in the fields provided.
- This is a timed quiz. The time remaining is shown in the quiz window and will continue to count down even if you leave the Canvas site.
- Open this quiz in *only one* browser window at a time. Opening the same Canvas quiz in multiple browser windows may cause problems with the auto-save features and some answers may be overwritten or lost.
- At the end of the time limit, your answers will be submitted automatically.

Authorised Materials: This exam is open-book. While undertaking this assessment you ***are permitted to***:

- make use of textbooks and lecture slides (including electronic versions) and lecture recordings
- make use of your own personal notes and material provided as part of tutorials and practicals in this subject
- make use of code that has been provided as part of this subject, or that you have written yourself
- use calculators, code, or mathematical software to compute numeric answers

While you are undertaking this assessment you ***must not***:

- make use of any messaging or communications technology
- make use of any world-wide web or internet-based resources such as Wikipedia, Stack Overflow, or Google and other search services
- act in any manner that could be regarded as providing assistance to another student who is undertaking this assessment, or will in the future be undertaking this assessment.

The work you submit ***must be based on your own knowledge and skills***, without assistance from any other person.

Technical support

This exam is a Canvas Quiz. Technical support for this exam can be accessed at: **https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/technical-support** **(https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/technical-support)**

Additional information about Canvas Quizzes, including troubleshooting tips, can be found **here** **(https://students.unimelb.edu.au/your-course/manage-your-course/exams-assessments-and-results/exams/exam-types)** (scroll down to the Canvas Quiz section).

Academic Integrity Declaration

By commencing and/or submitting this assessment I agree that I have read and understood the **University's policy on academic integrity.** **(https://academicintegrity.unimelb.edu.au/#online-exams)**

I also agree that:

1. Unless paragraph 2 applies, the work I submit will be original and solely my own work (cheating);
2. I will not seek or receive any assistance from any other person (collusion) except where the work is for a designated collaborative task, in which case the individual contributions will be indicated; and,
3. I will not use any sources without proper acknowledgment or referencing (plagiarism).
4. Where the work I submit is a computer program or code, I will ensure that:
    a. any code I have copied is clearly noted by identifying the source of that code at the start of the program or in a header file or, that comments inline identify the start and end of the copied code; and
    b. any modifications to code sourced from elsewhere will be commented upon to show the nature of the modification.

Troubleshooting

In case you cannot upload your files as requested (due to technical difficulties), please follow the steps below:

1. Name your file with your Question number followed by your Name and Student ID e.g. for Q7 for Jane Bloggs with Student ID 123456 you would upload file: **Q7 Jane Bloggs 123456.jpg**

2. Upload your files by opening the OneDrive link below - clicking this link will open a new Tab in your browserand will prompt you to select your files for upload: **https://unimelbcloud-my.sharepoint.com/:f:/g/personal/adrianrp_unimelb_edu_au/Ehw9ar9QNVtOlN7qlIHLO_oBBEylbGxwxGlwY5** **(https://unimelbcloud-my.sharepoint.com/:f:/g/personal/adrianrp_unimelb_edu_au/Ehw9ar9QNVtOlN7qlIHLO_oBBEylbGxwxGlwY5i1lF6_Qw)**

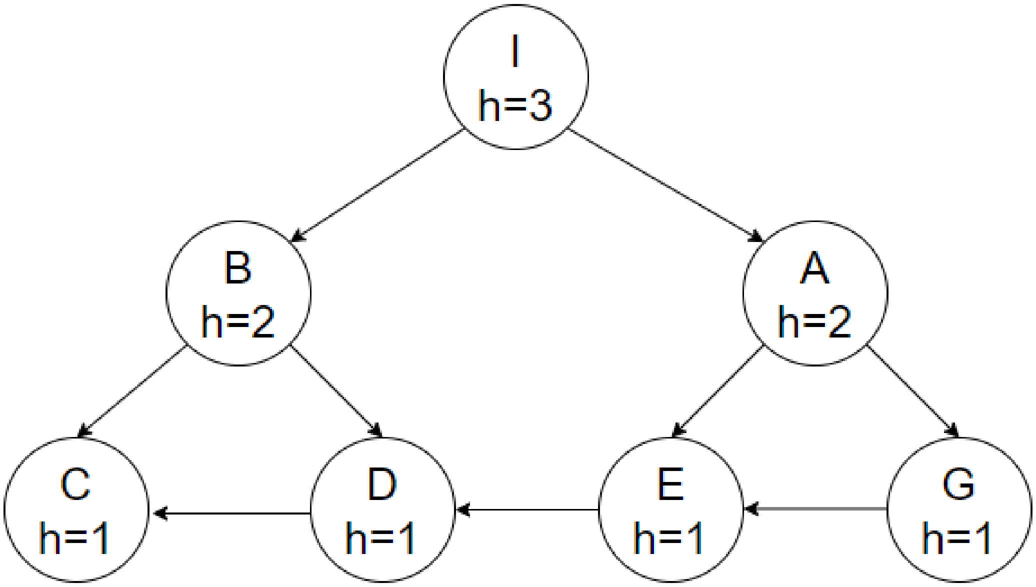
Late file upload policy: For timed exams, a deduction of 1 mark from the **final mark** (not exam mark) for each minute late up to 30 minutes. The time stamp on the server will be used as the submission time.

Search

Question 1

1 pts

We wish to use the A* algorithm to traverse the search tree below. Assume a fixed cost of 1 to transition between nodes and assume that ties are broken alphabetically, i.e. If node f(M) = f(N) then M will be expanded first. The first node to be expanded will be the initial node I. Which will be the parent of the 6th node expanded?



○ B

○ D

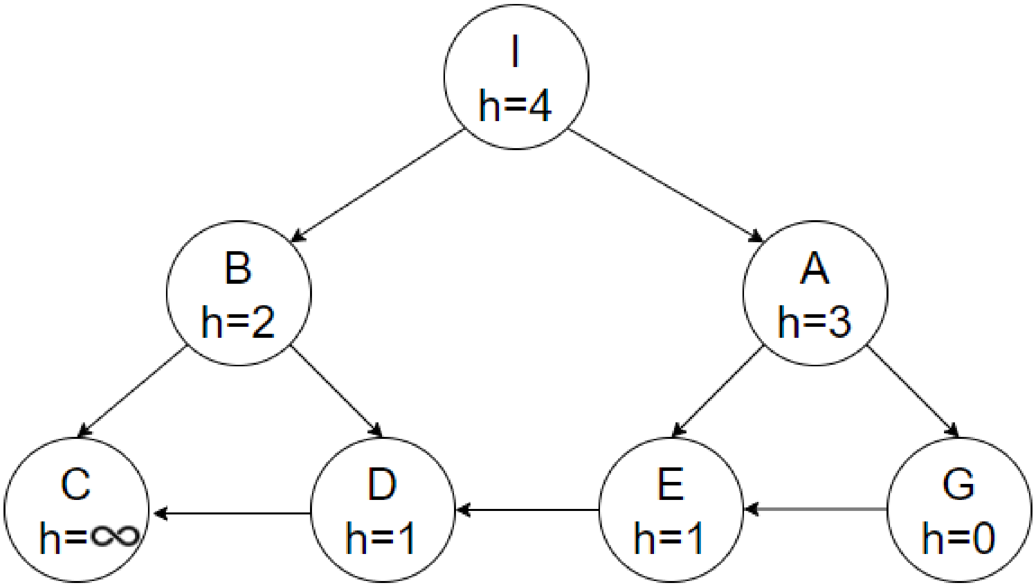
○ E

○ G

○ A

Question 2

3 pts



Assume a cost of 1 to move between nodes, initial state I, and Goal state G. With reference to the diagram above, you cannot change the heuristic values, but you can add or remove edges and add nodes with an associated heuristic value of your choice. Which nodes (with their h value) and edges do you need to add so the heuristic becomes:

(i) Safe
(ii) Admissible
(iii) Consistent
(iv) Goal aware

For each property, explain which nodes/edges you need to add to the graph. If you add new nodes/edges, explain why they are needed. If some property is unachievable, explain why.

p

0 words </>

Classical Planning

Question 3

1 pts

Which of the following statements is false?

○ All admissible, goal aware heuristics are consistent

○ Depth first search is complete for acyclic state spaces

○ The IDA* algorithm is optimal for admissible heuristics

○ The hadd heuristic is inadmissible in general

○ The hmax heuristic is always admissible

Question 4

1 pts

Below is the Bellman-Ford Table of $h^{add}(I)$ for a particular problem where I is the initial state of the problem.

completed(A)	completed(B)	completed(C)	completed(D)	completed(E)
3	Infinity	Infinity	1	Infinity
3			1	

All actions have cost=1. The following actions are available

Action One:

- Precondition: completed(A)

- Add: completed(B)

Action Two:

- Prec: completed(B), Completed(D)

- Add: completed(C)

Action Three:

- Prec: completed(C), Completed(B)

- Add: completed(E)

Action Four:

- Prec: completed(C), Completed(D)

- Add: completed(E)

Compute the values of the next row, given the actions above. Update first the value of Completed(B), then Completed(C), and finally Completed(E), in that order. If you change the value of predicate Completed(B), then you can use this value in the computation of the next predicates: Completed(C) and Completed(E).

Question 5

1 pts

Below is the Bellman-Ford Table of $h^{max}(I)$ for a particular problem where I = {completed(D)} is the initial state of the problem.

completed(A)	**completed(B)**	**completed(C)**	**completed(D)**	**completed(E)**
2	Infinity	Infinity	0	Infinity
2			0	

All actions have cost=1. The following actions are available

Action One:

- Precondition: completed(A)

- Add: completed(B)

Action Two:

- Prec: completed(B), Completed(D)

- Add: completed(C)

Action Three:

- Prec: completed(C), Completed(B)

- Add: completed(E)

Action Four:

- Prec: completed(C), Completed(D)

- Add: completed(E)

Compute the values of the next row, given the actions above. Update first the value of Completed(B), then Completed(C), and finally Completed(E), in that order. If you change the value of predicate Completed(B), then you can use this value in the computation of the next predicates: Completed(C) and Completed(E).

Question 6

3 pts

Draw or define a graph such that IW(1) is guaranteed to terminate without expanding the goal.

Write down the order in which IW(1) expands the nodes in your graph, and justify why a node is novel or not.

Note: avoid making large examples, a graph with 4 nodes should be sufficient.

p

0 words </>

Question 7

5 pts

A robot's (R) mission is to save planet earth and make sure all carbon mines (M) are closed. The robot can move directly between mines as long as it has a snack (S) for each voyage across mines. The robot can close down a mine only if it is at the same position as the mine. Initially the mines are open, and the goal is to close all the mines.

Describe briefly in STRIPS how to model the domain described. Include a specification of the parameters of the actions, and the preconditions and postconditions of each action. Include a description of the goal state of the problem, and create 1 possible initial state where the goal is reachable, and 1 possible initial state where the goal is not reachable. Your initial states need to have 3 or more snacks and 3 or more mines. Explain clearly any assumption made.

You are allowed to use variables as arguments for the actions (action schemes), specifying the values of the variables. Note: it is not compulsory to use PDDL syntax, as long as you can convey the main ideas.

p

0 words </>

Question 8

5 pts

Create a STRIPS problem with at most three actions such that $h^{add}\left(I, G\right) \neq h^{FF}\left(I, G\right)$.

Specify your STRIPS actions using the following notation: <action name> : preconditions -> effects.

For example, action $a : p, q \rightarrow r, \text{not } t$, would stand for action a, where p and q are the preconditions, and the effects add r and delete t.

To answer this question, show your workings by 1) creating the STRIPS problem, 2) finding the value of $h^{max}(I)$, 3) then the value of $h^{ff}(I)$ using the best supporter function induced by h^{max}, and finally 4) the value of $h^{add}(I)$. You then would be able to show that $h^{add}\left(I, G\right) \neq h^{FF}\left(I, G\right)$

Edit View Insert Format Tools Table

12pt Paragraph

p 0 words

Markov Decision Proceses (MDPs)

Question 9

1 pts

Consider a policy π that takes a state and returns the action *a* that should be chosen in state *s*.

What type of policy is this?

○ A random policy

○ An initial policy

○ A deterministic policy

○ A stochastic policy

○ A local policy

Question 10

1 pts

What is the correct formula for policy extraction from a value function V? Select all correct answers.

□ $\text{argmax}_{a \in A(s)} \sum_{s \in S} P_a(s' \mid s)[r(s, a, s') + \gamma V(s')]$

□ $\text{argmax}_{a \in A(s)} V(s)$

□ $\text{argmax}_{a \in A(s)} \sum_{s \in S} P_a(s' \mid s)[r(s, a, s') + \gamma \max_{a' \in A(s')} Q(s', a')]$

□ $\pi(s) = Q(s, a)$

□ $\pi(s) = \text{argmax}_{a' \in A(s)} \ Q(s, a')$

Question 11

2 pts

Match the techniques below with their properties. Multiple techniques can match to one property.

Policy iteration	[Choose]
Monte-Carlo Tree Search	[Choose]
Q-learning	[Choose]

Reinforcement Learning (RL)

Question 12

1 pts

What is the difference between on-policy and off-policy learning?

- ○ On-policy learning uses its policy to do exploration, while off-policy does exploration instead of exploitation
- ○ On-policy learning does temporal-difference updates based on the best possible next action executed, while off-policy does updates assuming the actual next action
- ○ On-policy learning using its policy to do exploitation, while off-policy does exploitation instead of exploration
- ○ On-policy learning does temporal-difference updates based on the actual next action executed, while off-policy does updates assuming the best possible next action
- ○ On-policy updated based on the next state, while off policy feeds back on the current state
- ○ On-policy updated based on the current state, while off policy feeds back on the next state

Question 13

1 pts

Backward induction and multi-agent MCTS are both techniques for solving extensive form games. Under which circumstances would you choose to use backward induction instead of multi-agent MCTS?

- □ If an optimal solution is needed
- □ If the environment is not one of the players
- □ If there are only two players
- □ If the game tree is small enough to solve the problem exhaustively

Question 14

1 pts

What is the difference between reward shaping and Q-function initialisation?

○ In reward shaping the potential function is used in the update, while in Q-value initialisation, the potential is calculated in the initial step

○ Nothing -- they are equivalent

○ Reward shaping uses potential functions while Q-function initialisation uses real functions

○ Reward functions work for any Q-function representation, while for Q-function initialisation it must be a Q-table representation

Question 15

4 pts

Consider a reinforcement learning agent this is try to learn how fast a vacuum cleaning robot can travel without over-heating.

There are two states: *cool* and *warm.*

There are two actions: *slow* and *fast.*

If the robot goes fast, it is more likely to transition to a *warm* state than it is goes *slow.*

Using a learning rate of 0.4 and a discount factor of 0.8, we arrive at the following Q-table:

Q(cool, fast)	9
Q(cool, slow)	7
Q(warm, fast)	4
Q(warm, slow)	5

The agent executes the action *fast* in the state *cool,* receives a reward of 6, and is now in the *warm* state. It will execute the action *slow* next.

Calculate the new value for *Q(cool, fast)* using 1-step SARSA to 2 decimal places.

Game Theory

Question 16

2 pts

Consider the following two-player game in normal form. Select all pure strategy Nash equilibrium for this game, if any exist.

		Player 2	
Player 1	D	E	F
A	5, 10	0, 10	5, 10
B	15, 10	5, 5	15, 10
C	20, 5	10, 15	15, 15

□ A, D: (5, 10)

□ B, D: (15, 10)

□ C, D: (20, 5)

□ A, E: (0, 10)

□ B, E: (5, 5)

□ C, E: (10, 15)

□ A, F: (5, 10)

□ B, F: (15, 10)

□ C, F: (15, 15)

Question 17

2 pts

In your own words, compare the concepts of pure strategy and mixed strategy in normal form games.

Edit View Insert Format Tools Table

12pt Paragraph B I U A T² ⋮

p 0 words </>

Question 18

5 pts

Consider the following problem. There are two famous pie shops in town. Each shop can charge $2, $4, or $5 for a pie. Each day, 6000 tourists will buy pies; 3000 from each shop, irrelevant of the price. However, 4000 local people will always by from the cheapest shop. If both shops have the same price, 2000 locals will buy from each. So, if Shop 1 is the cheapest price, it will sell 3000 pies to tourists and 4000 pies to locals, making $7000x, where x is the price.

Assuming that the shops set the prices without seeing each others' price, what prices will the shops select? *Show your working.*

Not saved Submit Quiz