Date: 9th July, 2020

The exam paper will be available at 3pm on Thursday, July 9th, 2020 on Moodle. The submission deadline will be 9pm on Thursday, July 9th, 2020. The responses are to be handwritten, scanned as a single PDF file and submitted on Moodle.

This exam is only for graduating students taking the AGP option.

The typical space requirement for each question will be indicated in the exam paper. Students can prepare responses by either (i) writing on blank paper sheets writing the question number or (ii) printing out the exam paper and writing in the space provided.

Please write answers legibly using dark blue/black ink. Please write your name, entry number and page number on all sheets.

Clarifications will not be provided during the exam. Necessary information is provided for each question. If a question seems under-specified to you, please make an assumption, specify it and solve given that assumption.

Question	Points	Score
1	10	
2	10	
3	10	
4	10	
5	10	
6	10	
7	10	
8	5	
9	5	
10	20	
Total:	100	

Name and Entry No.: ____________________

1. (10 points) A navigating robot r is moving in a 2D grid world (with 1×1 sized cells) towards a goal g, Figure 1. The robot can move to any unoccupied adjacent grid cell (including the ones along the diagonal). The robot is considering an A* search with either of the following heuristic functions: $h_0 = (|x_g - x_r| + |y_g - y_r|)$ or $h_1 = \max(|x_g - x_r|, |y_g - y_r|)$, where (x_r, y_r) and (x_g, y_g) denote the robot and the goal positions respectively. Please indicate if an A* search, for each choice of the heuristic functions will enable the agent to find (i) a feasible path to the goal if it exists? and (ii) the optimal path to the goal?

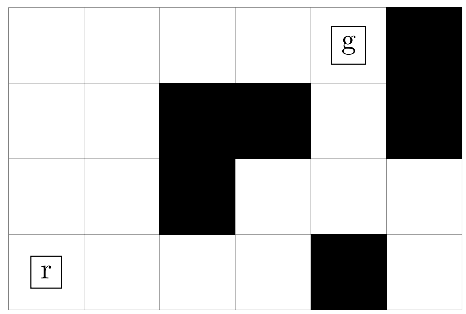

Figure 1: A grid world agent (r) navigates to the goal (g).

Name and Entry No.: ______________________________

2. (10 points) An underwater robot is submerged in a lake for surveying underwater planktons (organisms that live in water bodies). The robot is submerged in the lake from a designated point. The lake has a cliff on one side and the lake's bottom is inclined as shown in Figure 2. Let p_x and p_y denote the vehicle's x-y position. The robot moves using its thrusters that can impart a velocity v_x and v_y along the x and y axes respectively. The robot can be considered as a point mass and cannot move without the use of its thrusters. The vehicle is descending vertically to the lake bottom. However, there is noise in its motion and its sensor observations and hence it may deviate from its direct path requiring its thrusters to provide corrections during the descent. The robot is fitted with two range sensors. One sensor measures the distance d_x to the cliff and the second measures the distance d_y to the lake's bottom.

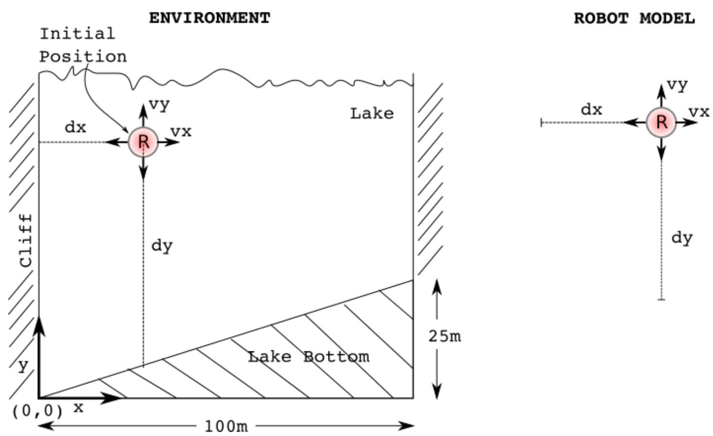

Figure 2: A robot is descending vetically to the lake bottom. Let p_x and p_y indicate the vehicles position, d_x and d_y denote distance measurements and v_x and v_y indicate velocity imparted by thrusters keeping the vehicle on course during descent.

Now consider the problem of estimating the vehicle. Design a Kalman Filter to estimate the vehicle state (x,y position) using its sensor observations at discrete time steps at intervals Δ_t. Let X_t denote the vehicle state, U_t denote actions/control input and Z_t denote the measurement received at time t.

(a) Write a motion model as appropriate for this problem setup defining each term in the equation:

$$X_{t+1} = A * X_t + B * U_t + \epsilon \tag{1}$$

Here, A and B are matrices that relate the current vehicle state and applied actions with the subsequent state. Let ϵ denote the noise in the vehicle's motion assumed to be Gaussian distributed (zero mean and covariance Q). Specify X_t, U_t, A, B and ϵ as per the scenario.

(b) Write an observation model as appropriate for this problem setup defining each term in the equation:

$$Z_t = C * X_t + \delta \tag{2}$$

Here, C is a matrix and δ denotes the noise in the measurement which is Gaussian distributed (zero mean and covariance R). Specify Z_t, X_{t+1}, C and δ as per the scenario.

(c) If the lake bottom is undulating and curved (not linearly increasing in height) then would the proposed estimator accurately estimate the vehicle position? Yes/No. Justify.

Name and Entry No.: ____________________

3. (10 points) Consider the underwater robot surveying a lake as described in the last problem. We now address the task of reaching a target site S in the lake for surveying, see Figure 3. Assume that the lake can be approximated as a 2D grid and a scientist has provided the location of the site for exploration as well as places where there are rocks which must be avoided. The robot can take 4 symbolic actions GoLeft, GoRight, GoUp and GoDown which can move the robot to the adjacent unoccupied grid cell. Actions are stochastic and succeed in reaching the intended cell only 70% of time. For example, if the robot takes GoUp action then there is a 10% chance of landing in the cells to the left, right or below the current cell. Assume that using the solution to the previous problem the robot can localize itself to a specific cell in the grid.

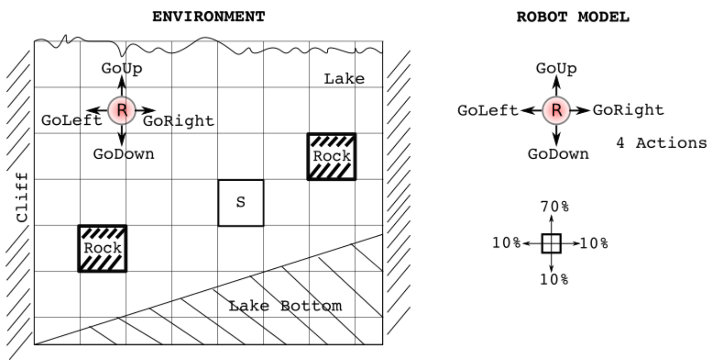

Figure 3: The robot's task is to plan its approach towards an identified site while avoiding rocky areas. The lake can be represented as a grid. The robot can take four actions which are stochastic with a 70% chance of moving to the intended cell.

(a) Formulate a model for the robot to plan its actions to reach the survey site.

(b) Now assume that the robot has a finite battery life and hence can plan only till a finite time horizon. Do you expect the agent to behave differently as the battery reaches exhaustion? Yes/No. Justify (1 line).

4. (10 points) Consider the underwater robot surveying a lake as described in the previous two problems. We now address the task of planning visits to many sites for exploration. The robotic agent is provided a set of n candidate sites $S_0, S_1, \ldots, S_{n-1}$ in the lake to survey planktons, see Figure 4. The robot's actions include exploring a specific site i from the set of n candidate sites. The exploration proceeds by deciding a site to explore, moving to the site (using the solution to the previous problem) and scanning the site for planktons. If the agent observes a new plankton from its camera, it takes a photograph and immediately returns to its initial position to uplink since communication underwater is difficult. In the case of no plankton sightings, the robot returns back to its initial position. After each survey, the agent must pick the site for exploration (it can visit sites any number of times). The plankton distribution at any site is stationary over time and unknown to the agent.

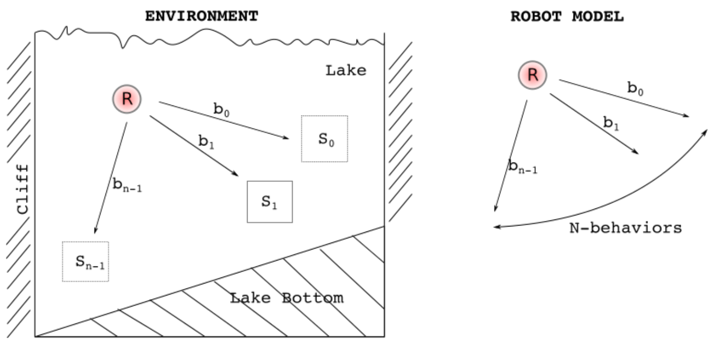

Figure 4: The robot can explore any of the n sites for planktons. It must decide how to explore the sites to maximize its objective of plankton sightings.

(a) How should the rational agent decide its site visits to maximise the scientific yield of plankton observations? Formulate a model for the agent to decide which site to explore at each iteration. Assume the absence of a time deadline for the task and the agent can ignore any cost of taking actions.

Name and Entry No.: ______________________________

(b) Now, assume that the scientist has some prior knowledge about which sites are likely to have higher plankton observations. How can such information be incorporated in your model? Describe briefly.

5. (10 points) A robot (with an arm) is tasked with stacking blocks on a table into a tower assembly. The robot is given a set of N blocks placed randomly on the table. Let x_i^p denote the position of the i^{th} block (in 3D). Let x_i^g denote the intended goal position of the i^{th} block in a tower arrangement. The robots hand position is denoted as x_r. The robot models the environment as an MDP accounting for errors in its actions. Let r denote the reward function for the MDP.

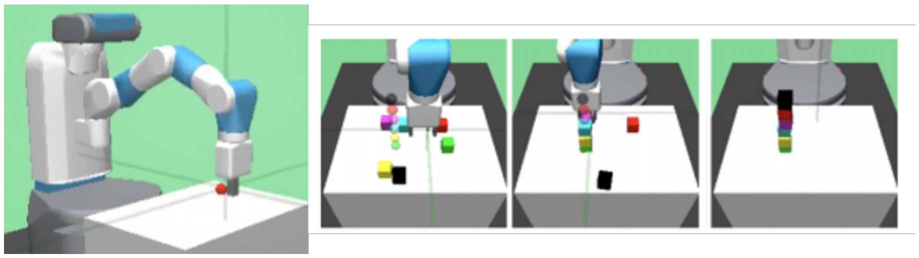

Figure 5: A robot manipulator is asked to stack blocks into a tower.

(a) Please write an expression for the reward function r that models the objective of stacking blocks into a tower at the intended positions. The stacking task is considered successful if all objects are stacked and if each object's final position is within tolerance ϵ of the prescribed goal location for each block.

(b) Assume that the robot computes a policy using the reward function you prescribed. Qualitatively speaking, what is the likely position of the robot's hand in relation to the stack of blocks at the end of the stacking? Can you modify the reward function to ensure that the robot's hand moves beyond a horizontal distance of 2ϵ from the nearest point from the tower. Let $x_{nearest}^g$ denote the nearest point on the tower at the end of the stacking behaviour?

Name and Entry No.: ______________________________

6. (10 points) Consider an embodied agent performing tasks in a household (setting similar to homework II). The agent encodes the domain using a symbolic representation where the state is represented as a list of true predicates. The predicates not mentioned in the state are considered false. The goal is assumed to only include predicates that are true. Actions are represented as a set of pre-conditions and effects where only positive predicates are allowed in pre-conditions. Effects are allowed to contain predicates that are deleted or added.

   (a) If the planning domain is represented by n predicates, then (i) how many states are possible? and (ii) how many goals are possible?

   (b) The agent is considering using the FF planner. It drops the negative effects from all its action definitions. Does the agent now have a relaxed problem?

7. (10 points) Consider a robot with an arm capable of picking objects on a table top. The robot possesses a language understanding model (Generalized Grounding Graphs) to interpret instructions. The model is trained with a corpus containing instructions for picking and placing blocks on the table. We can assume that if there is a single block on the table, the robot is able to interpret an instruction such as "pick up the block on the table".

   (a) Consider the scenario where a human operator facing the robot places an additional block on the table. Hence, there are two blocks on the table on either side of the robot. A human operator provides the following instruction, "pick up the block on the table". What would be output of the language understanding model?

   (b) Next, assume that the robot is equipped with a sensor (depth camera) for detecting and localizing the person standing in front and facing the robot. The human operator would now like the robot to correctly interpret the following instructions, "pick up the red block on my left" and "pick up the red block on your left". Briefly describe how you would enhance the language grounding model to correctly interpret the two instructions?

Name and Entry No.: ______________________________

8. (5 points) A navigation agent (e.g., an intelligent car) is mounted on the top with a visual sensor. The visual sensor captures an image of the agent's surroundings at each time step as it drives along the city. The agent's problem is to determine if it is at a place encountered before from visual appearance alone. The agent is using the FAB-MAP algorithm for place recognition. However, the algorithm has been simplified as follows. The model considers the visual words detected in an image at a location as independent. What effect will this have on the agent's ability to recognize places?

9. (5 points) An agent (robot manipulator) is learning to perform a pouring task by observing human demonstrations using Behaviour Cloning (BC) and Inverse Reinforcement Learning (IRL) methods. The human repeatedly pours water from a jug into a cup on a table top. The robot observes the hand trajectory of the human and the cup remains at a fixed position on the table during the demonstration. At test time, the robot is asked to perfrom the task under two settings. (i) the cup is in the same position on the table as in the human's demonstration, and (ii) the cup is moved to a different location on the same table. Explain whether the agent is likely to be successful at learning the task with BC or IRL in each of the two cases.

10. (20 points) Please provide responses to the following questions indicating whether the statements are true or false with short (1-2 line) justification.

    (a) A goal-directed agent that computes a value function for an MDP is a *reactive* agent and an agent that executes a given policy is a *utility* based agent.

    (b) In a Kalman Filter, the uncertainty over the state increases when measurements are incorporated.

    (c) An agent using a discrete HMM for grid world localization (like assignment 1) can predict its position only a *single* time step ahead.

    (d) The SAS+ (multivariate) planning domain representation is a more expressive and a more compact representation in comparison to STRIPS-style planning representation.

    (e) For an agent with a high degree of freedom (e.g., an anthropomorphic arm), planning in the configuration space is preferred over planning in the task space.

Name and Entry No.: ______________________

(f) DAGGER and Behavior Cloning do not use the same imitation loss.

(g) An embodied agent moving in a building with multiple rooms is asked to find a person somewhere in the building. This problem cannot be modeled as an MDP.

(h) The size of the planning graph for a symbolic planning task is exponential in the size of the planning problem.

(i) The Beta distribution is a conjugate prior for the Gaussian distribution.

(j) The inertial measurement unit directly observes the pose (position and axes rotation) of a moving vehicle.