Name and Entry No.: ____________________

Date: 27th August, 2020

The responses are to be handwritten, scanned as a single PDF file and submitted on Moodle. This exam is only for the Non-AGP students enrolled in the course.

Typical space requirement for each question is indicated. Responses can be prepared by either (i) writing on blank paper sheets indicating the question number or (ii) printing out the exam paper and writing in the space provided for each question.

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

If writing on blank paper sheets, please ensure that responses to each question begin on a fresh page and that all sub-parts of a question appear together.

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.

You may refer to your class notes during the exam. Cases of copying in the answer scripts will be awarded zero points for this exam.

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

1. (10 points) A robot is trying to estimate the state of a container to be stored or discarded. The robot has tactile sensors that can measure forces during object interaction and a classifier that can predict if the container is FULL from the given force measurements. The classifier false positive rate is 10% and the false negative rate is 20%. Further, there is a 20% chance that the robot's lifting is jerky causing the container's contents to fall out during the motion. Please write down a model for this problem. Indicate how to estimate the belief over the container's state given the observation sequence FULL, ¬ FULL and FULL. Please provide how you would formulate the problem without calculating the exact value of the belief.

Name and Entry No.: ______________________________

2. (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?

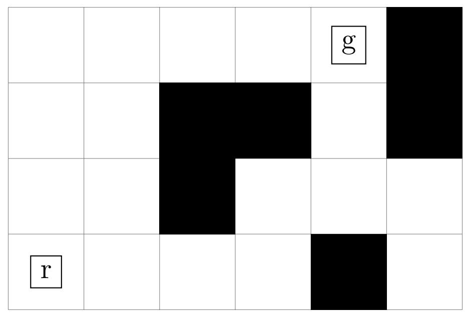

Figure 1: A grid world agent (r) navigates to the goal (g).

3. (10 points) Consider an embodied agent performing table top manipulation. The agent encodes its task, for instance, block assembly, as a planning problem encoded in a STRIPS/PDDL-style representation. The state is represented as a list of true logical literals (ground predicates). For example, a state consisting of literals: $On(A, B)$ and $On(B, C)$, implies that block A is on top of B and B is on top of C respectively. The literals not mentioned in the state are considered false. The goal is assumed to only possess positive literals. Actions are represented as a set of pre-conditions and effects where only positive literals are allowed in pre-conditions. Effects are allowed to contain predicates that are deleted or added. Please answer the following questions providing concise (1-3 line) explanations.

   (a) If there are n literals characterizing the domain, then (i) how many states are there in the planning domains? and (ii) how many goal states are possible?

   (b) Now consider that the agent is provided with an impossible goal: $On(A, B)$ and $On(B, C)$ and $On(C, A)$. Will the resulting planning graph directly encode that the goal is impossible? Justify.

Name and Entry No.: ____________________

4. (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 robot's 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.

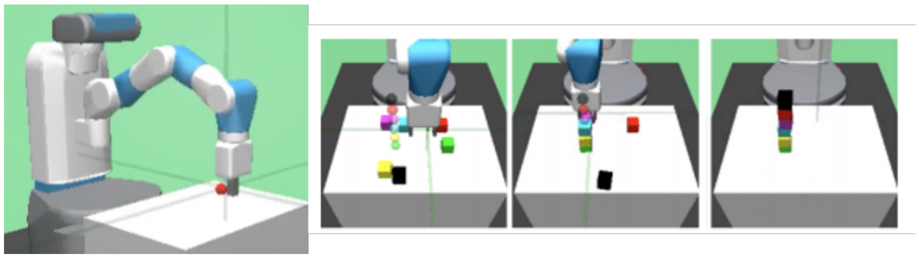

Figure 2: A robot manipulator is asked to stack blocks into a tower.

(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? Let $x_{nearest}^g$ denote the nearest point on the tower at the end of the stacking behaviour. 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?

5. (10 points) An agent equipped with an arm and a gripper is learning to grasp or pick up objects in its manipulation area, Figure 3 (left). The robot can view the scene with a depth camera that provides the 3D position of the object, the table and the plastic wall-type object on the table. The agent is equipped with three types of picking behaviors, Figure 3 (right): (i) *Crane grasp*: the hand is positioned on top of the object, lowered to the object and picked up joining its fingers (like the movement of a crane). (ii) *Scoop*: the agent slides the object along the table towards a vertical wall and then lifts or scoops against a vertical wall and (iii) *Pick along edge*: the object is slid to the edge of the table and then the agent grasps it from the side. The grasping strategies are stochastic and may not be successful in picking the object (the object may fall from the end). Assume that the agent can interact with the object multiple times and the object is returned to its original position in case of a picking failure.

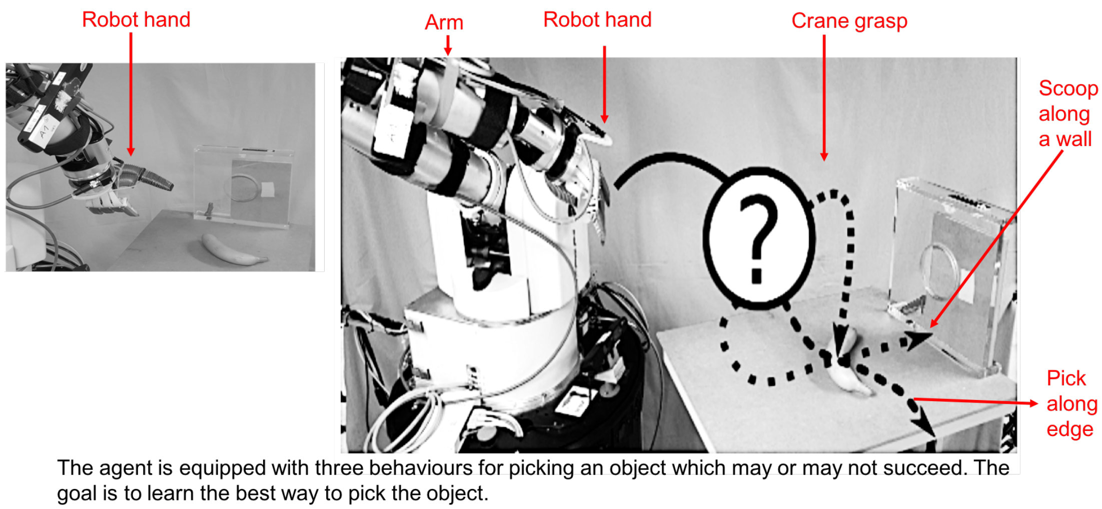

Figure 3: Problem setup where an agent is learning to grasp objects.

(a) Please provide a (data-efficient) approach for the agent to learn to grasp a given object (e.g., a banana on the table).

(b) How can you extend your learning algorithm to picking objects from a set of n possible objects (e.g., a medicine box, a can, a screw-driver etc.). If the robot has a prior belief over which grasping behaviors are likely to succeed for an object, how can this information be incorporated in your approach? Can your approach generalize to learning to pick *novel* objects not seen during training? Describe briefly.

Name and Entry No.: ____________________

6. (10 points) This is a design exercise that involves developing a language understanding system for an unmanned aerial vehicle (or a drone) intended for visual inspection of a building. Assume that the vehicle is capable of flying to a position in 3D space and can hover at a pre-specified height above the ground. The human operator requires a language understanding system that can interpret instructions such as "*robot, fly through the window*", "*robot, hover near the door of the building*" etc. Please describe how you can formulate and train a model (such as Distributed Correspondence Graphs or Generalized Grounding Graphs) for this requirement. Describe the key steps.

Name and Entry No.: ____________________

7. (5 points) An agent (e.g., a car) is mounted with a visual sensor that captures an image of surrounding 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 a simple algorithm of taking an image, vectorizing it and computing the pairwise sum of square difference between the current and each of the previously recorded images. In which scenarios will this localization technique succeed and in which scenarios will it fail?

8. (5 points) Show that a symbolic planning domain with l literals and a action symbols admits a planning graph (with n levels) that is of size polynomial in the size of the planning domain.

Name and Entry No.: ______________________

9. (10 points) An agent is situated in an environment which can be modeled as an MDP ($< S, A, R, T, \gamma >$). Let $s \in S$ and $a \in A$ denote the states and actions. Let the function $R(s_t, a_t)$ denote the reward obtained when an action a_t is executed from a state s_t. Assume a discounted reward with $\gamma < 1$. The transition function T is represented as a successor state distribution $p(s_{t+1}|s_t, a_t; \theta)$. The transition function is unknown and the uncertainty in the transition function is modeled by the parameter θ. The agent is initialized with a prior belief over the unknown parameters as $p(\theta)$. Let $h_t = \{s_1, a_1, \ldots s_t\}$ denote the state action history and let $p(\theta|h_t)$ denote the belief over the latent parameters θ given the action history.

   (a) The agent's task is to make progress towards the goal (by maximizing cumulative reward) as well as learn a transition model of the environment (since the transition model is unknown). Suggest how the reward function can be formulated in order to balance both objectives?

   (b) Please explain how the degree of uncertainty in the transition function parameters affects the agent's behavior based on the agent design in part (a). Further, explain your agent's behaviour in states where the transitions are totally random.

10. (20 points) Please provide brief responses (1-3 sentences) for the following questions.

(a) In a Kalman Filter, the uncertainty over the state variables increases with the application of the transition model. Yes/No. Justify.

(b) A car traveling on Mehrauli road is trying to estimate its position by observing the Qutab Minar. Assume that the vehicle is equipped with a distance sensor that can measure the distance to five salient visual features on the Qutab Minar. The car is using the standard Kalman Filter to estimate its position. Will the position estimates converge to the true position? Yes/No. Justify.

(c) Why are conjugate priors useful in Bayesian models?

(d) Inverse Reinforcement Learning (IRL) is concerned with learning an unknown transition function. Yes/No. Justify.

(e) What is the intuition behind the FF heuristic for symbolic planning?

(f) How is DAGGER different from Behaviour Cloning?

(g) The Distributed Correspondence Graph model does not perform sampling over continuous quantities as in Generalized Grounding Graphs. Yes/No. Justify.

(h) How does the Max-Ent MDP formulation differ from the standard MDP formulation?

(i) Consider two places on a map possessing the same visual features but with different pair-wise distances. The FAB-MAP algorithm is likely to declare the two places as a loop closure. Yes/No. Justify.

(j) The global positioning system guiding an autonomous vehicle on the road is likely to be affected by tall buildings. Yes/No. Justify.

Name and Entry No.: ______________________