Surname, Name, ID number: ____________________

Autonomous and Mobile Robotics M

22 December 2022 - Theory

Some questions may have more than one correct answers: for each question, indicate all the correct answers.

1. The configuration space of a unicycle mobile robot is:
   - ○ $[x\ y\ \theta]^T \in \mathbb{R}^3$
   - ⊗ $[x\ y\ \theta]^T \in \mathbb{R}^2 \times \mathbb{S}$
   - ○ $[x\ y\ \theta\ \gamma]^T \in \mathbb{R}^2 \times \mathbb{S}^2$
2. A constraint is said *non-holonomic* if:
   - ⊗ the differential relation between the coordinates is not reducible to finite form
   - ○ finite relations between the coordinates of the system are present
   - ○ if differentiable/integrable relations between the coordinates of the system are present
3. Given the constraints matrix equation in Pfaffian form $A(q)\dot{q} = 0$, the admissible robot speed:
   - ○ is generated by a matrix $G(q)$ such that $\mathrm{Ker}(G(q)) = \mathrm{Im}(A(q)), \forall q$
   - ⊗ is generated by a matrix $G(q)$ such that $\mathrm{Im}(G(q)) = \mathrm{Ker}(A(q)), \forall q$
   - ○ is generated by a matrix $G(q)$ such that $G(q) = A(q)^{-1}, \forall q$
4. Consider an obstacle avoidance algorithm based on potential fields
   - ○ a concave shape of the obstacle can in many cases avoid the problem of local minima
   - ○ the control consists in setting the velocity of the robot equal to the gradient of the potential
   - ⊗ the overall potential is the sum of an attractive potential generated by the goal and repulsive potentials generated by the obstacles
5. The robot configuration space is randomly sampled taking into account feasible trajectories:
   - ○ in the Probabilistic Roadmap algorithm;
   - ⊗ in the Rapidly-Exploring Random Tree;
   - ○ in the Voronoi Roadmap algorithm.
6. A process satisfies the Markov property if:
   - ○ the agent state is the same as the environment state
   - ⊗ one can make predictions for the future of the process based solely on its present state
   - ○ one can make predictions for the future of the process only based on the process full history
7. The *action value function* is defined as:
   - ○ $q_\pi(s, a) = \mathbb{E}_\pi[R_{t+1}|S_t = s, A_t = a]$
   - ○ $q_\pi(s, a) = \mathbb{E}_\pi[G_t|S_t = s]$
   - ⊗ $q_\pi(s, a) = \mathbb{E}_\pi[\sum_{k=0}^{\infty} \gamma^k R_{t+k+1}|S_t = s, A_t = a]$
8. The Bellman optimality equation for the state value function can be written as
   - ○ $v_*(s) = \max v_\pi(s)$
   - ⊗ $v_*(s) = \max_{a\in\mathcal{A}} q_*(s, a)$
   - ⊗ $v_*(s) = \max_{a\in\mathcal{A}} \mathbb{E}_*[G_t|S_t = s, A_t = a]$
9. The λ-return is defined as:
   - ○ $G_t^\lambda = R_{t+1} + \lambda V(S_{t+1})$
   - ○ $G_t^\lambda = \sum_{k=0}^{\infty} \lambda^k R_{t+k+1}$
   - ⊗ $G_t^\lambda = (1-\lambda) \sum_{n=1}^{\infty} \lambda^{n-1} G_t^{(n)}$
10. In value function approximation by stochastic gradient descent, the parameter vector update is defined as:
    - ⊗ $\Delta\mathbf{w} = \alpha\,(v_\pi(S) - \hat{v}(S, \mathbf{w}))\nabla_{\mathbf{w}}\hat{v}(S, \mathbf{w})$
    - ○ $\Delta\mathbf{w} = \alpha\,\mathbb{E}_\pi\left[(v_\pi(S) - \hat{v}(S, \mathbf{w}))\nabla_{\mathbf{w}}\hat{v}(S, \mathbf{w})\right]$
    - ○ $\Delta\mathbf{w} = -\frac{1}{2}\alpha\nabla_{\mathbf{w}} J(\mathbf{w})$

Autonomous and Mobile Robotics M

22 December 2022 - Exercise

The student is asked to solve the following problem.

Let us consider a fully observable and deterministic environment with 5 states providing the following rewards

$R(s_1)$	$R(s_2)$	$R(s_3)$	$R(s_4)$	$R(s_5)$
1	0	-1	0	2

The set of possible actions is {MoveLeft, MoveRight}, proving with probability 1 the transition of the state to the left one or to the right one respectively. The environment is initially at state s_1, and first the policy $\pi_1(\cdot) = \text{MoveRight}$ is applied for 4 time steps. Then, the policy $\pi_2(\cdot) = \text{MoveLeft}$ is applied for additional 4 time steps.

Starting from an arbitrary initialisation of the state value function and assuming a discount factor $\gamma = 1$ and a weight $\alpha = 0.5$, compute the state value function provided by a TD algorithm after the execution of π_1 and π_2 in the following two tables.

$v_{\pi_1}(s_1)$	$v_{\pi_1}(s_2)$	$v_{\pi_1}(s_3)$	$v_{\pi_1}(s_4)$	$v_{\pi_1}(s_5)$

$v_{\pi_2}(s_1)$	$v_{\pi_2}(s_2)$	$v_{\pi_2}(s_3)$	$v_{\pi_2}(s_4)$	$v_{\pi_2}(s_5)$

Solution:

The state value function is initialized to 0 for all the states.

$v_{\pi_1}(s_1)$	$v_{\pi_1}(s_2)$	$v_{\pi_1}(s_3)$	$v_{\pi_1}(s_4)$	$v_{\pi_1}(s_5)$
0	-0.5	0	1	0

$v_{\pi_2}(s_1)$	$v_{\pi_2}(s_2)$	$v_{\pi_2}(s_3)$	$v_{\pi_2}(s_4)$	$v_{\pi_2}(s_5)$
0	0.25	-0.25	0	0.5