Surname, Name, ID number: ______________________

Autonomous and Mobile Robotics M

14 June 2022 - Theory

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

1. 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
2. Consider a WMR with constraint matrix equation $A(q)\dot{q} = 0$, with $\dot{q} \in \mathbb{R}^N$
   - ◯ the equation can be fully integrated if the constraints represent N pure rolling wheels (no slipping)
   - ⊗ the allowable speeds can be generated by a matrix $G(q)$ such that $Im(G(q)) = Ker(A(q))$
   - ◯ the equation represents N non-holonomic constraints if slipping of the wheels is not allowed
3. The constraint introduced by a single wheel can be expressed as:
   - ◯ $x \sin\theta - y \cos\theta = 0$
   - ⊗ $\dot{x} \sin\theta - \dot{y} \cos\theta = 0$
   - ◯ $x \sin\dot{\theta} - y \cos\dot{\theta} = 0$
4. 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 $\mathrm{Im}(G(q)) = \mathrm{Ker}(A(q)), \forall q$
   - ◯ 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 $G(q) = A(q)^{-1}, \forall q$
5. In reactive navigation, the robot:
   - ◯ plans the trajectory using a map of the environment
   - ◯ updates the planned path on the based of sensor information
   - ⊗ navigates the environment on the base of the sensor information only
6. Examples of map-based navigation algorithms are:
   - ⊗ distance transform planning;
   - ⊗ A* and D*;
   - ◯ bug algorithms.
7. In map-based navigation, the robot:
   - ⊗ plans the trajectory using a map of the environment
   - ◯ updates the planned path on the based of sensor information
   - ◯ navigates the environment on the base of the sensor information only
8. The *state value function* is defined as:
   - ◯ $v_\pi(s) = \mathbb{E}_\pi[R_{t+1}|S_t = s]$
   - ⊗ $v_\pi(s) = \mathbb{E}_\pi[G_t|S_t = s]$
   - ◯ $v_\pi(s) = \mathbb{E}_\pi[R_{t+1} + \gamma v_\pi(s')|S_t = s]$
9. The future discounted reward is defined as:
   - ⊗ $G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{i=0}^{\infty} \gamma^i R_{t+i+1}$
   - ◯ $G_t = R_{t+1} + R_{t+2} + R_{t+3} + \cdots = \sum_{i=0}^{\infty} R_{t+i+1}$
   - ◯ $G_t = R_{t-1} + \gamma R_{t-2} + \gamma^2 R_{t-3} + \cdots = \sum_{i=0}^{\infty} \gamma^i R_{t-i-1}$
10. In Reinforcement Learning algorithms, the reward:
    - ◯ must be a function of the agent state
    - ⊗ can be a function of the environment state
    - ◯ depends on time

Autonomous and Mobile Robotics M

14 June 2022 - Exercise

The student is asked to solve the following problem.

Let us consider a fully observable and deterministic environment with 5 states $s_{\{1,\dots,5\}}$.

s_1	s_2	s_3	s_4	s_5

- Action set : {TryLeft, TryRight}
- Rewards:
  - +2 in state s_1
  - 0 in state s_2
  - -1 in state s_4
  - +1 in all other states
- Transition model:
  - $p(s_1|s_1, \text{TryLeft}) = p(s_5|s_5, \text{TryRight}) = 1$
  - $p(s_1|s_1, \text{TryRight}) = p(s_2|s_1, \text{TryRight}) = 0.5$
  - $p(s_1|s_2, \text{TryLeft}) = p(s_2|s_2, \text{TryLeft}) = 0.5$
  - $p(s_2|s_2, \text{TryRight}) = p(s_3|s_2, \text{TryRight}) = 0.5$
  - ...
- Policy: $\pi(\text{TryLeft}|s_{\{1,\dots,5\}}) = \pi(\text{TryRight}|s_{\{1,\dots,5\}}) = 0.5$
- Discount factor $\gamma = 0.8$

Starting from an arbitrary initialisation of the state value function, compute the first iteration of the state value function evaluation provided by a Dynamic Programming algorithm assuming the random policy π.

$v_\pi(s_1)$	$v_\pi(s_2)$	$v_\pi(s_3)$	$v_\pi(s_4)$	$v_\pi(s_5)$

Solution:

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

$v_\pi(s_1)$	$v_\pi(s_2)$	$v_\pi(s_3)$	$v_\pi(s_4)$	$v_\pi(s_5)$
1.5	0.75	0.25	0	0.5