Surname, Name, ID number: ______________________________

Autonomous and Mobile Robotics M

WRITTEN EXAM SAMPLE - Theory

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

1. A constraint is said *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. The constraint introduced by a single wheel can be expressed as:
   - ○ $x \sin\theta - y \cos\theta = 0$
   - ⊗ $\dot{x} \sin\theta - \dot{y} \cos\theta = 0$
   - ○ $x \sin\dot{\theta} - y \cos\dot{\theta} = 0$
3. A single Swedish wheel:
   - ○ enables to control the motion along both the rolling and the driven directions
   - ⊗ does not allow any control in the rolling direction
   - ○ does not allow to control the motion neither along the rolling nor the driven directions
4. 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
5. The distance transform of a map
   - ⊗ has the same size of the original map
   - ⊗ has elements which values is the distance to the target position
   - ○ can be computed only usign Euclidean distance
6. Sequential Monte-Carlo Localization:
   - ○ resamples the particles on the based of their spacial distribution
   - ○ performs better than the EKF in case the robot configuration is described by a Gaussian distribution
   - ⊗ works also in case the probability distribution function of the robot configuration is not known
7. In Reinforcement Learning algorithms, the reward:
   - ⊗ is a scalar feedback signal
   - ○ is minimized by the agent
   - ⊗ indicates how well agent is doing at step t
8. The future discounted reward is defined as:
   - ⊗ $G_t = R_{t+1} + \gamma R_{t+2} + \gamma^2 R_{t+3} + \cdots = \sum_{i=0}^{\infty} \gamma^i R_{t+i+1}$
   - ○ $G_t = R_{t+1} + R_{t+2} + R_{t+3} + \cdots = \sum_{i=0}^{\infty} R_{t+i+1}$
   - ○ $G_t = R_{t-1} + \gamma R_{t-2} + \gamma^2 R_{t-3} + \cdots = \sum_{i=0}^{\infty} \gamma^i R_{t-i-1}$
9. The agent state:
   - ○ is always the same as the environment state
   - ⊗ is the same as the environment state in case of fully observable environments
   - ⊗ is the same as the environment state in Markov decision processes
10. The *state value function* is defined as:
    - ○ $v_\pi(s) = \mathbb{E}_\pi[R_{t+1} | S_t = s]$
    - ⊗ $v_\pi(s) = \mathbb{E}_\pi[G_t | S_t = s]$
    - ⊗ $v_\pi(s) = \mathbb{E}_\pi[R_{t+1} + \gamma v_\pi(s') | S_t = s]$

WRITTEN EXAM SAMPLE - Open Questions

The student is asked to answer the following questions.

1. Describe synthetically with words and formulas the main steps of the EKF locatization:

$$
\begin{aligned}
\hat{\mathbf{q}}^{+}\langle k+1\rangle &= \mathbf{f}(\hat{\mathbf{q}}\langle k\rangle, \hat{\mathbf{u}}\langle k\rangle) && \text{State prediction}\\
\hat{\mathbf{P}}^{+}\langle k+1\rangle &= \mathbf{F}_q\hat{\mathbf{P}}\langle k\rangle\mathbf{F}_q^T + \mathbf{F}_v\hat{\mathbf{V}}\mathbf{F}_v^T && \text{Covariance projection}\\
\mathbf{v} &= \mathbf{z}^{\#}\langle k+1\rangle - \mathbf{h}(\hat{\mathbf{q}}^{+}\langle k+1\rangle, \mathbf{p}_i) && \text{Innovation}\\
\mathbf{K} &= \hat{\mathbf{P}}^{+}\langle k+1\rangle\mathbf{H}_q^T(\mathbf{H}_q\hat{\mathbf{P}}^{+}\langle k+1\rangle\mathbf{H}_q^T + \mathbf{H}_w\mathbf{W}\mathbf{H}_w^T)^{-1} && \text{Kalman gain}\\
\hat{\mathbf{q}}\langle k+1\rangle &= \hat{\mathbf{q}}^{+}\langle k+1\rangle + \mathbf{K}\mathbf{v} && \text{State update with innovation}\\
\hat{\mathbf{P}}\langle k+1\rangle &= \hat{\mathbf{P}}^{+}\langle k+1\rangle - \mathbf{K}\mathbf{H}_q\hat{\mathbf{P}}^{+}\langle k+1\rangle && \text{Covariance update}
\end{aligned}
$$

2. The λ-return G_t^λ combining all n-step returns $G_t^{(n)}$ can be expressed as:

$$G_t^\lambda = (1-\lambda)\sum_{n=1}^{\infty}\lambda^{n-1}G_t^{(n)}$$

WRITTEN EXAM SAMPLE - Exercise

The student is asked to solve the following problem.

Let us consider a fully observable and deterministic environment with 7 states $s_{\{1,\ldots,7\}}$.

s_1	s_2	s_3	s_4	s_5	s_6	s_7

- Action set : {TryLeft, TryRight}
- Rewards:
  - +1 in state s_1
  - +10 in state s_7
  - 0 in all other states
- Transition model:
  - $p(s_1|s_1, \text{TryRight}) = p(s_2|s_1, \text{TryRight}) = 0.5$
  - $p(s_1|s_2, \text{TryLeft}) = p(s_2|s_2, \text{TryLeft}) = 0.5$
  - $p(s_2|s_2, \text{TryRight}) = p(s_3|s_2, \text{TryRight}) = 0.5$
  - ...
- Policy: $\pi(\text{TryLeft}|s_{\{1,\ldots,7\}}) = \pi(\text{TryRight}|s_{\{1,\ldots,7\}}) = 0.5$
- Discount factor $\gamma = 0.5$

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.

$v_\pi(s_1)$	$v_\pi(s_2)$	$v_\pi(s_3)$	$v_\pi(s_4)$	$v_\pi(s_5)$	$v_\pi(s_6)$	$v_\pi(s_7)$

Solution:

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

$v_\pi(s_1)$	$v_\pi(s_2)$	$v_\pi(s_3)$	$v_\pi(s_4)$	$v_\pi(s_5)$	$v_\pi(s_6)$	$v_\pi(s_7)$
0.75	0.25	0	0	0	2.5	7.5