Autonomous Navigation and Perception (086762)
Homework #5

Submission guidelines:

- Submission in pairs by: 25 May 2022, 24:00.
- Email submission as a *single zip or pdf file* to `anpsubmission@gmail.com`.
- Make sure to name the submitted file ID1-ID2.zip (or .pdf) with ID1,ID2 - the students' ids.
- For hands-on parts, your source code should be submitted as well (include it in the above zip file).

Theoretical questions:

1. Explain why out of the box Monte Carlo Tree Search will perform poorly in continuous state spaces.

Hands-on tasks:

1. Consider a mobile robot navigating in a 2D environment from Homework 2. Several beacons are scatted in the environment. The locations of these beacons are known to the robot (denote it by X_j^b for the jth beacon), and represented by a matrix $X^b \in \mathbb{R}^{2\times n}$, i.e. $X^b = [X_1^b, \ldots, X_n^b]$, where n is the number of beacons.

   The robot state at a time instant i is the x-y robot position $X_i \in \mathbb{R}^2$. The motion and observation models, $\mathbb{P}(X_{k+1}|X_k, a_k)$ and $\mathbb{P}(z_k|X_k)$, are given by

   $$X_{k+1} = f(X_k, a_k) + w \quad , \quad z_k = h(X_k) + v \tag{1}$$

   where w and v are process and measurement noise terms, with $w \sim N(0, \Sigma_w)$ and $v \sim N(0, \Sigma_v(X_k))$.

   In this exercise, we shall consider simple linear models, i.e. $f(X_k, a_k) = F \cdot X_k + a_k$ with $F = I_{2\times 2}$ and $a_k \in \mathbb{R}^2$ is the commanded position displacement. We also consider position measurements such that $z_k \in \mathbb{R}^2$ and $h(X_k) \equiv X_k$.

   Consider a prior over robot position at time instant 0 is available: $b(X_0) \doteq \mathbb{P}(X_0) = N(\mu_0, \Sigma_0)$.

   Assume the robot gets a position observation $z \in \mathbb{R}^2$ from a beacon if it is sufficiently close to it (distance less than d). For simplicity, consider the beacons are positioned sufficiently apart such that an observation from only one of the beacons can be obtained for every possible robot location. Moreover, assume the accuracy of these observations deteriorates with the distance r to the beacon, starting from a certain given distance $r_{min} \leq d$, such that for $r \leq d$, the measurement covariance is $\Sigma_v = (0.01 \cdot \max(r, r_{min}))^2 \cdot I_{2\times 2}$.

   Convert this POMDP problem to Belief-MDP problem in the following way. Implement a new transition function `TransitBeliefMDP`.

   - Input: Robot belief over the state parameterized by mean and covariance $\mathcal{N}(\mu_k, \Sigma_k)$, action a_k
   - Output: Posterior belief $\mathcal{N}(\mu_{k+1}, \Sigma_{k+1})$

   The function `TransitBeliefMDP` propagates the belief using the action and transition model, samples a state from propagated belief, uses the state to create an observation, if the observation is not null, updates the propagated belief with observation.

2. Consider $n = 9$ beacons equally scattered in a 9×9 grid, and let the coordinate of its bottom left corner be $(0, 0)$. To clarify, one of the beacons is located at $(0, 0)$. Assume $b(X_0) = N(\mu_0, \Sigma_0)$ with $\mu_0 = (0, 0)^T$ and $\Sigma_0 = I_{2\times 2}$, $\Sigma_w = 0.1^2 \cdot I_{2\times 2}$. Let the actual initial robot location (ground truth) be $X_0 = (-0.5, -0.2)^T$. Further, let $d = 1$ and $r_{min} = 0.1$. In this question we consider the following cost over the belief

$$c(b, a) = \|\mu - x^g\|_2 + \lambda \det(\Sigma), \tag{2}$$

where x^g is the robot goal, μ is the mean value of the belief and Σ it's covariance. Let us define the following action space

$$\mathcal{A} = \left\{(1,0)^T, (-1,0)^T, (0,1)^T, (0,-1)^T, \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T, \left(-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right)^T, \left(\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T, \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right)^T, (0,0)^T\right\}$$

(a) Implement sparse sampling algorithm.

(b) Adjust the parameter λ and number of observations in sparse sampling to receive best results. Report values you chose for plots.

(c) Run simulation of the 15 consecutive planning sessions with planning horizon of 15. Namely, engage planning session, apply the received action to the ground truth state, generate observation, update the belief with action and observation, repeat planning session with obtained belief.

(d) Plot the ground truth robot trajectory yielded by action sequence generated by 10 consecutive planning sessions. Indicate in all plots also the n beacons.

(e) Plot the observations received alongside the ground truth trajectory.

(f) Plot the beliefs alongside the trajectory.