ADI
Planning, Learning and Decision Making

Erik Hamberg
Matej Rojec

Homework 3

Question 1

Write down the state space $\mathcal{X}$, the action space $\mathcal{A}$, and the observation space $\mathcal{Z}$ for the POMDP describing the decision process of the truck driver. Consider that a new time step occurs whenever the driver takes an action at one of the seven dotted locations (Recycling plant and stops A to F).

Solution: **Action Space**

The action space for the POMDP is:

$$\begin{aligned}\mathcal{A} &= \{\text{Collect Garbage}, \text{Drop Garbage}, \text{Move Up}, \text{Move Down}, \text{Move Left}, \text{Move Right}\} \\ &= \{\text{C}, \text{Dr}, \text{U}, \text{D}, \text{L}, \text{R}\}\end{aligned}$$

Solution: **State Space**

The state space $\mathcal{X}$ consists of all tuples $(\mathcal{P}, l, t_B, t_C, t_D)$, where $\mathcal{P} \in \{A, B, C, D, E, F, R\}$ indicates the position of the truck, the component $l \in \{\text{loaded}, \text{not loaded}\}$ indicates whether the truck has collected the trash at any of the locations or not and t_B, t_C, t_D are binary indicators of whether there is garbage at stops B, C, and D. So if $t_B = 1$ then there is trash at stop B and if it is 0 there is no trash in stop B.

Solution: **Observation Space**

The state space $\mathcal{Z}$ consists of all tuples $(\mathcal{P}, l, t)$, where $\mathcal{P} \in \{A, B, C, D, E, F, R\}$ indicates the position of the truck, the component $l \in \{\text{loaded}, \text{not loaded}\}$ indicates whether the truck has collected the trash at any of the locations or not, the component t is a binary indicator of whether there is garbage at stops $p \in \mathcal{P}$ (the current position).

Question 2

Write down the cost function for the POMDP.

Solution:

We can write the cost function c as:

$$c(s,a) = \begin{cases} 0 & \text{if } a = \text{Dr and } p = \text{R and } l = \text{loaded} \\ 1/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } a = \text{C and } (\{p = \text{B and } l = \text{not loaded and } t_B = 1\} \\ & \text{or } \{p = \text{C and } l = \text{not loaded and } t_C = 1)\} \\ & \text{or } \{p = \text{D and } l = \text{not loaded and } t_D = 1\}) \\ 2/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{R and p} = \text{E }\} \\ & \text{or } \{a = \text{D and p} = \text{F}\} \\ 3/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{R and p} = \text{R }\} \\ & \text{or } \{a = \text{L and p} = \text{A}\} \\ 4/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{R and p} = \text{A }\} \\ & \text{or } \{a = \text{L and p} = \text{B}\} \\ 55/300 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{R and p} = \text{C }\} \\ & \text{or } \{a = \text{L and } p \in \{\text{C, E}\}\} \\ & \text{or } \{a = \text{D and p} = \text{A}\} \\ 7/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{U and } p \in \{\text{A, F}\}\} \\ & \text{or } \{a \in \{\text{L, R}\} \text{ and p} = \text{D}\} \\ 8/30 \cdot (1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D)) & \text{if } \{a = \text{R and p} = \text{B}\} \\ & \text{or } \{a \in \text{L and p} = \text{F}\} \\ 1 & \text{otherwise} \end{cases}$$

where $(1 - \mathbb{1}_{\{0\}}(t_B + t_C + t_D))$ is 0 if $t_B + t_C + t_D = 0$ and 1 otherwise.

Question 3

Suppose that, at time step t, the driver is in stop A and decides to move up, towards stop D. At that moment, the driver has a belief of 0.3 that there is garbage in location D. Indicate the possible beliefs at time step $t + 1$ regarding whether there is garbage in location D, explaining your reasoning.

Solution:

At time $t + 1$ the driver makes one of the 4 possible observations either (D, loaded , 0), (D, not loaded , 0), (D, loaded , 1) or (D, not loaded , 1). In the first two cases the belief that there is garbage at stop D is 0 since he observed no trash there, in the second case the belief is 1 since he observed the trash there. Here we used the Markov property of the belief.