Modeling of the Problem

DPOC Programming Excercise

Nov. 21, 2019.

1 States, Actions and Disturbance

States

As the system is a time-invariant system, we denote the state space as $S = \{s_1, s_2, \cdots, s_K\}$, $where K = M \times N - num(TREE)$, $s_i = (m_i, n_i, \phi_i)$. There is only one terminal state of the system: $s_T = (m_T, n_T, 1)$, where $map(m_T, n_T) = DROP\_OFF$. At time k, the state is $x_k \in S$.

Actions

There are five different actions in total, for simplicity, we write actions as

$$U = \{(1,0), (-1,0), (0,1), (0,-1), (0,0)\}$$

corresponding to $\{East, West, North, South, Stay\}$ respectively. At time t, the action is $u_k \in U$

Disturbance

Directly let $\omega_k = x_{k+1}, x_{k+1} \in S$, so the transition matrix is $P_{x_{k+1}|x_k,u_k} = P_{w_k|x_k,u_k}$

2 State Transition Matrix

The State Transition can be divided into 4 substeps: take action, moved by wind, shoot by the resident and pick up. In order to figure out the exact transition of each substep, we introduce 3 internal states x_k^1, x_k^2, x_k^3, which denote the state of the drone after taking 1-3 substeps, as is illustrated in Figure 1.

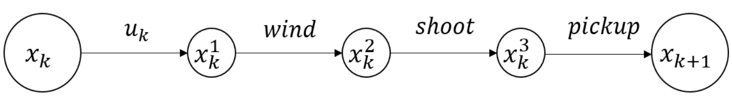

Figure 1: Substeps in one stage

Consider the influence of crash or teminal directly on the next states disregarding the internal states (e.g. if x_k^2 crash then x_{k+1} can only be in base; if x_k is terminal states, then x_{k+1} can only be in ternimal states), we introduce extra states C and T for internal statespace. C means already crash and T means x_k already in terminal in the last stage. Therefore, we have:

$$
\begin{aligned}
&x_k \in S \\
&x_k^1 \in S \cup \{T\} \\
&x_k^2 \in S \cup \{C\} \cup \{T\} \\
&x_k^3 \in S \cup \{C\} \cup \{T\} \\
&x_{k+1} \in S
\end{aligned}
$$

In this way, the transition probability canbe calculated as:

$$P(x_{k+1}|x_k, u_k) = P(x_k^1|x_k, u_k)P(x_k^2|x_k^1)P(x_k^3|x_k^2)P(x_{k+1}|x_k^3)$$

In the following subsections, we will discuss each part in order.

2.1 Substage 1

In this substage we will calculate $P(x_k^1|x_k, u_k)$, x_k^1 means the state after taking the action, without wind and shoot. We consider the after taking the action that is not allowed, i.e. will arrive at tree or out of bound, we set the state to stay. In the stage cost section, we set the cost of unallowable action to be larger than cost of stay, so it will not be considered in the optimizing process. It can be written as:

$$
\begin{aligned}
&P(x_k^1|x_k \in S/s_T, u_k) = 0, \quad if \quad crash \\
&P(x_k^1|x_k \in S/s_T, u_k) = \begin{cases} 1, & x_k^1 = x_k + u_k \\ 0, & x_k^1 = otherstates \end{cases} \\
&P(x_k^1|x_k = s_T, u_k) = \begin{cases} 1, & x_k^1 = T \\ 0, & x_k^1 = otherstates \end{cases}
\end{aligned}
$$

2.2 Substage 2

In this substate we will calculate $P(x_k^2|x_k^1)$, x_k^2 means the state after taking the action, after moved by the wind, not been shot yet. It can be written as:

$$
\begin{aligned}
&P(x_k^2|x_k^1 \in S) = \begin{cases} \frac{p_{wind}}{4}, & x_k^2 = crash?C : x_k^1 + (1,0,0)^T \\ \frac{p_{wind}}{4}, & x_k^2 = crash?C : x_k^1 + (0,1,0)^T \\ \frac{p_{wind}}{4}, & x_k^2 = crash?C : x_k^1 + (-1,0,0)^T \\ \frac{p_{wind}}{4}, & x_k^2 = crash?C : x_k^1 + (0,-1,0)^T \\ 1 - p_{wind}, & x_k^2 = x_k^1 0, \quad x_k^2 = otherstates \end{cases} \\
&P(x_k^2|x_k^1 = T) = \begin{cases} 1, & x_k^2 = T \\ 0, & x_k^2 = otherstates \end{cases}
\end{aligned}
$$

2.3 Substage 3

In this substate we will calculate $P(x_k^3|x_k^2)$, x_k^3 means the state after taking the action, after moved by the wind, and after being shot. It can be written as:

$$P(x_k^3|x_k^2 \in shotrange) = \begin{cases} \prod_{i=0}^{m}(1 - \frac{\gamma}{d_i+1}), & x_k^3 = x_k^2 \\ 1 - \prod_{i=0}^{m}(1 - \frac{\gamma}{d_i+1}), & x_k^3 = C \\ 0, & x_k^3 = otherstates \end{cases}$$

$$P(x_k^3|x_k^2 \in S/shotrange) = \begin{cases} 1, & x_k^3 = x_k^2 \\ 0, & x_k^3 = otherstates \end{cases}$$

$$P(x_k^3|x_k^2 = C) = \begin{cases} 1, & x_k^3 = C \\ 0, & x_k^3 = otherstates \end{cases}$$

$$P(x_k^3|x_k^2 = T) = \begin{cases} 1, & x_k^3 = T \\ 0, & x_k^3 = otherstates \end{cases}$$

2.4 Substage 4

In this substate we will calculate $P(x_{k+1}|x_k^3)$. We denote the special state to change ϕ as pickupstate, which is $(m_p, n_p, 0)$. The transition matrix can be written as:

$$P(x_{k+1}|x_k^3 \in S/pickupstate) = \begin{cases} 1, & x_{k+1} = x_k^3 \\ 0, & x_{k+1} = otherstates \end{cases}$$

$$P(x_{k+1}|x_k^3 = pickupstate) = \begin{cases} 1, & x_{k+1} = x_k^3 + (0,0,1)^T \\ 0, & x_{k+1} = otherstates \end{cases}$$

$$P(x_{k+1}|x_k^3 = C) = \begin{cases} 1, & x_{k+1} = basestate \\ 0, & x_{k+1} = otherstates \end{cases}$$

$$P(x_{k+1}|x_k^3 = T) = \begin{cases} 1, & x_{k+1} = s_T \\ 0, & x_{k+1} = otherstates \end{cases}$$

3 State Cost

We can use dynamic programming method to evaluate the cost in each stage by back-propagating among substates. Because the stage cost is irrelavant to the state x_{k+1}, given x_k^3. So we directly update cost from x_k^3:

$$g_3(x_k^3) = \begin{cases} N_c, & x_k^3 = C \\ 0, & x_k^3 = T \\ 1, & x_k^3 = otherstates \end{cases}$$

Then $g_2(x_k^2$ and $g_1(x_k^1)$ can be updtated in order:

$$g_2(x_k^2) = p(x_k^3|x_k^2)g_3(x_k^3)$$
$$g_1(x_k^1) = p(x_k^2|x_k^1)g_2(x_k^3)$$

Finally, $g_k(x_k, u_k)$ can be computed considering the not allowed action:

$$g_k(x_k, u_k) = \begin{cases} inf + p(x_k^1|x_k, u_k)g_1(x_k^1), & x_k + u_k \to crashed \\ p(x_k^1|x_k, u_k)g_1(x_k^1), & otheractions \end{cases}$$