COMP0037

Report

Planning in Uncertain Worlds

Group AS

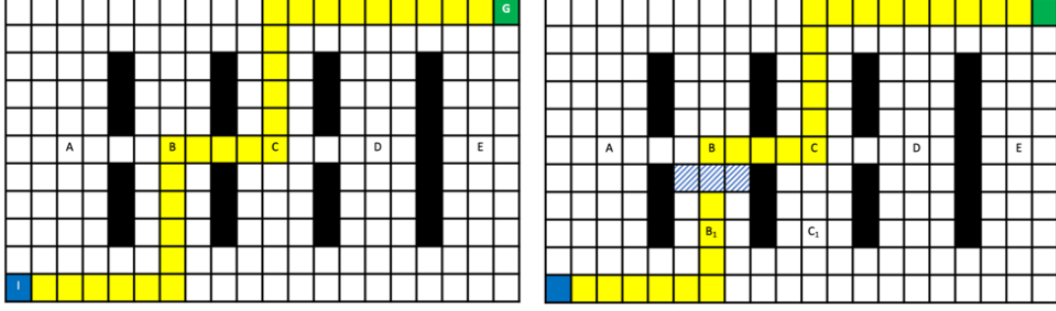| Student Name	Student number
Arundathi Shaji Shanthini	16018351
Dmitry Leyko	16021440
Tharmetharan Balendran	17011729

Department: Department of Electronic and Electrical Engineering
Submission Date: 28th of April 2020

Contents

1 Decision Re-Plan Policy

1.1 Policy Selection when Obstacle is Observed

(a) The original planned path form I to G going through Aisle B and C.

(b) An obstacle in aisle B obstructs the planned path of the robot.

Figure 1: Illustration of case where robot observes an obstruction to it's planned path.

The scenario that we will be analysing is the case shown in Fig. 1a. The robot is required to go from a cell I to a cell G. These cells are marked blue and green in Fig. 1a respectively. The figure also shows the original planned path that the robot computed going down aisle B. However, once the robot turns into aisle B it observes that the aisle is blocked. This observation is done at the point when the robot reaches the cell labelled B_1. At this point the robot can either decide to wait until the obstruction clears or it can re-plan a path. Once the robot observes the obstacle, the time the robot must wait for the obstacle to clear may be represented by the expression in Eq. 1.

$$T = \frac{0.5}{\lambda_B} + \widetilde{T} \tag{1}$$

The wait time is dependent on λ_B and a random variable $\widetilde{T}$. The random variable $\widetilde{T}$ is sampled from a exponential distribution with a rate parameter of $2\lambda_B$. The probability density function (PDF) for $\widetilde{T}$ is shown in Eq. 2.

$$f(t) = \begin{cases} 2\lambda_B e^{-2\lambda_B t} & t \geq 0 \\ 0 & t < 0 \end{cases} \tag{2}$$

where, $\lambda = \dfrac{\lambda_B}{0.5}$.

As previously mentioned, the robot has two options to choose from: to wait for the obstacle to clear, or to re-plan and execute the new path. The two are different policies the robot must choose from. We use the symbol π to denote a policy. A policy is a mapping from the world state to an action the robot can execute.

Let us assume that if the robot decides to wait the total path length (number of cells) will be K_1 while if the robot decides to re-plan and execute the total path length will be K_2. We let the quantity K equal to the larger value between K_1 and K_2. Now we may write the policy for the robot to wait as π_K^1 and the policy for re-planning as π_K^2. These policies are padded correspondingly to produce actions $\mathbf{u}_K^1$ and $\mathbf{u}_K^1$ that are padded with zero-cost state preserving actions.

To see which policy is better on average, we consider the expected value of the cost function for both cases. The case when policy π_K^1 is chosen is characterized by the inequality shown in Eq. 3.

$$\mathbb{E}\left[L\left(\pi_K^1\right)\right] \leq \mathbb{E}\left[L\left(\pi_K^2\right)\right] \tag{3}$$
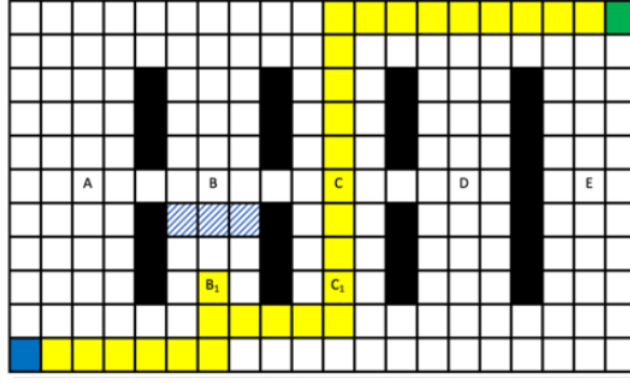
Figure 2: The path for the re-plan policy π_K^2 which bypasses aisle B and goes down aisle C.

We can see from Fig. 1a that the cost of the original planned path is given by the expression in Eq. 4. In this equation, the terms L_{XY} denote the cost of the shortest path between cell X and cell Y. Additionally, the term L_W represents the cost of waiting 1 unit of time.

$$L\left(\pi_K^1\right) = L_{IB_1} + TL_W + L_{B_1B} + L_{BC} + L_{CG} \tag{4}$$

From Fig. 2 which shows the re-planned path, we can also see that the cost of this path is equal to the expression in Eq. 5

$$L\left(\pi_K^2\right) = L_{IB_1} + L_{B_1C_1} + L_{C_1C} + L_{CG} \tag{5}$$

Substituting the expressions in Eq. 4 and Eq. 5 into Eq. 3. We obtain the inequality shown in Eq. 6

$$
\begin{aligned}
\mathbb{E}\left[L_{IB_1} + TL_W + L_{B_1B} + L_{BC} + L_{CG}\right] &\leq \mathbb{E}\left[L_{IB_1} + L_{B_1C_1} + L_{C_1C} + L_{CG}\right] \\
L_{IB_1} + \mathbb{E}\left[T\right]L_W + L_{B_1B} + L_{BC} + L_{CG} &\leq L_{IB_1} + L_{B_1C_1} + L_{C_1C} + L_{CG} \\
\mathbb{E}\left[T\right] &\leq \frac{L_{B_1C_1} + L_{C_1C} - L_{B_1B} - L_{BC}}{L_W}
\end{aligned}
\tag{6}
$$

The quantity $\mathbb{E}\left[T\right]$ is the expected value for the time the robot has to wait for the obstacle to clear. As we know the distribution that the variable is sampled from we can compute the expected value. The expected value for the time taken is given by the expression found in Eq. 7

$$
\begin{aligned}
\mathbb{E}\left[T\right] &= \mathbb{E}\left[\frac{0.5}{\lambda_B} + \widetilde{T}\right] \\
&= \frac{0.5}{\lambda_B} + \mathbb{E}\left[\widetilde{T}\right] \\
&= \frac{0.5}{\lambda_B} + \int_0^\infty t\frac{\lambda_B}{0.5}e^{-\frac{\lambda_B}{0.5}t}dt \\
&= \frac{0.5}{\lambda_B} + \frac{\lambda_B}{0.5}\left[-(t)\left(\frac{0.5}{\lambda_B}e^{-\frac{\lambda_B}{0.5}t}\right) - \int_0^\infty \frac{0.5}{\lambda_B}e^{-\frac{\lambda_B}{0.5}t}dt\right]_0^\infty \\
&= \frac{0.5}{\lambda_B} + \frac{\lambda_B}{0.5}\times\frac{0.5}{\lambda_B}\left[-te^{-\frac{\lambda_B}{0.5}t} + \left[-\frac{\lambda_B}{0.5}e^{-\frac{\lambda_B}{0.5}t}\right]\right]_0^\infty \\
&= \frac{0.5}{\lambda_B} + \left[-te^{-\frac{\lambda_B}{0.5}t} - \frac{\lambda_B}{0.5}e^{-\frac{\lambda_B}{0.5}t}\right]_0^\infty \\
&= \frac{0.5}{\lambda_B} + \left[\lim_{t\to\infty} -te^{-\frac{\lambda_B}{0.5}t} + \frac{0.5}{\lambda_B}e^{-\infty} + 0 + \frac{\lambda_B}{0.5}e^0\right] \\
&= \frac{0.5}{\lambda_B} + \frac{0.5}{\lambda_B} = \frac{1}{\lambda_B}
\end{aligned}
\tag{7}
$$

3

Substituting the expression from Eq. 7 into Eq. 6 we obtain the expression in Eq. 8. The right-hand side of the inequality in Eq. 8 represents the smallest possible value for λ_B for which the waiting policy π_K^1 is a better option than the re-plan policy π_K^2. The inequality also takes into consideration the constraint that $\lambda_B > 0$ and negates the solution when $\lambda_B < 0$.

$$\frac{1}{\lambda_B} \leq \frac{L_{B_1 C_1} + L_{C_1 C} - L_{B_1 B} - L_{BC}}{L_W}$$
$$\lambda_B \geq \frac{L_W}{L_{B_1 C_1} + L_{C_1 C} - L_{B_1 B} - L_{BC}} \tag{8}$$

1.2 Policy Selection at Start
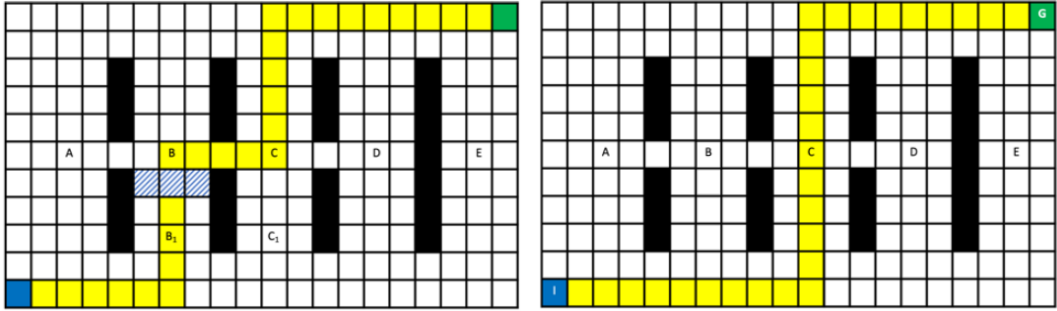
(a) The scenario where the robot decides to go down Aisle B, encounters an obstacles and waits for it to clear.

(b) The scenario where the robot decides to avoid Aisle B completely due to the obstacle.

Figure 3: The different policies the robot can pick from at the beginning.

Instead of reacting to the obstacle as the robot observes it, the robot may also make a decision before it start to move. The robot has knowledge of where the obstacles will be and the probability distribution for the wait time. Depending on the probabilities involved, the robot may decide to avoid the obstacle altogether instead of going the route with the obstacle and wait. In the case of the warehouse example with 5 aisles, and one obstacle in aisle B, the two policies are to choose to travel down aisle B and wait if an obstacle is present (illustrated in Fig. 3a) or to avoid aisle B and plan directly along aisle C (illustrated in Fig. 3b). Similar to the approach in §1.1, we may denote the policy that goes down aisle B as π_K^1 and the policy that goes down aisle C as π_K^2. Once again these are padded with zero-cost state preserving actions so as to be of the same dimension.

We can see from Fig. 3a that the cost of policy that goes down aisle B, π_K^1, can cause 2 different scenarios:

- Drive down aisle B and wait for the obstacle to move. The cost of this scenario is given by the expression:
$$L(\pi_k^1) = L_{IB_1} + TL_W + L_{B_1 B} + L_{BC} + L_{CG} \tag{9}$$

- Drive down aisle B, encounter an obstacle and reroute via aisle C.The cost of this scenario is given by the expression:
$$L(\pi_k^1) = L_{IB_1} + TL_W + L_{B_1 B} + L_{BC} + L_{CG} \tag{10}$$

The cost of the policy π_K^2, which associates to the robot deciding to drive down directly via aisle C is given by the expression in Eq. 11.

$$L(\pi_k^2) = L_{IC} + L_{CG} \tag{11}$$

As we are concerned with the average case scenario we compare the expected value of the loss function just as we did in section 1.1. π_K^2 will be chosen over π_K^1 if the expected value of it's cost is lower as shown by Eq. 12.

$$\mathbb{E}\left[L\left(\pi_K^2\right)\right] \leq \mathbb{E}\left[L\left(\pi_K^1\right)\right] \tag{12}$$

Given that we can intuitively think of the cost associated to the second scenario of driving down aisle B given by eq. (??) is a much longer path than directly driving down aisle C we can eliminate that from the comparison. Then, substituting the loss functions from eq. (??) and (11) into the inequality in eq. (12), we obtain the result shown in eq. (13).

$$\begin{aligned}
\mathbb{E}[L_{IC} + L_{CG}] &\leq \mathbb{E}[L_{IB_1} + TL_W + L_{B_1B} + L_{BC} + L_{CG}] \\
\mathbb{E}[T]L_W &\geq L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC} \\
\mathbb{E}[T] &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W}
\end{aligned} \tag{13}$$

By substituting the expected value for T found in eq. (7) we obtain a constraint for λ_B as shown in eq. (14). The right hand side of the final inequality in eq. (14) represents the largest value for λ_B for which the robot will decide to go directly down aisle C avoiding aisle B.

$$\begin{aligned}
\frac{1}{\lambda_B} &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W} \\
\lambda_B &\leq \frac{L_W}{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}
\end{aligned} \tag{14}$$

1.3 Considering the Probability of the Obstacle Being Present

1.4 Considering Multiple Obstacles

2 ROS Implementation

References

Appendices