

# COMP0037

## Report

### Planning in Uncertain Worlds

#### Group AS

<u>Student Name</u>	<u>Student number</u>
Arundathi Shaji Shanthini	16018351
Dmitry Leyko	16021440
Tharmetharan Balendran	17011729

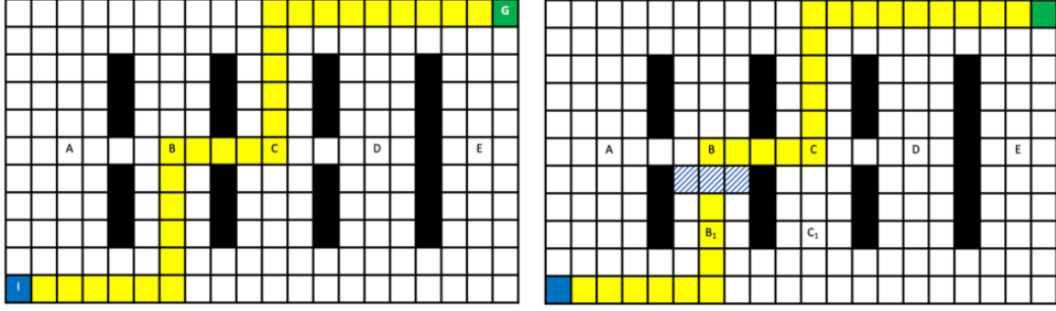
**Department:** Department of Electronic and Electrical Engineering  
**Submission Date:** 28<sup>th</sup> of April 2020

# Contents

<b>1</b>	<b>Decision Re-Plan Policy</b>	<b>2</b>
1.1	Policy Selection when Obstacle is Observed . . . . .	2
1.2	Policy Selection at Start . . . . .	4
1.3	Considering the Probability of the Obstacle Being Present . . . . .	5
1.4	Considering Multiple Obstacles . . . . .	6
<b>2</b>	<b>ROS Implementation</b>	<b>6</b>
	<b>Appendices</b>	<b>7</b>

# 1 Decision Re-Plan Policy

## 1.1 Policy Selection when Obstacle is Observed



(a) The original planned path from I to G (b) An obstacle in aisle B obstructs the going through Aisle B and C. planned path of the robot.

Figure 1: Illustration of case where robot observes an obstruction to its 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).

$$T = \frac{0.5}{\lambda_B} + \tilde{T} \quad (1)$$

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

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

where, the rate parameter  $\lambda = 2\lambda_B$ .

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  $\pi$  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  $\pi_K^1$  and the policy for re-planning as  $\pi_K^2$ . 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  $\pi_K^1$  is chosen is characterized by the inequality shown in Eq. (3).

$$\mathbb{E} [L (\pi_K^1)] \leq \mathbb{E} [L (\pi_K^2)] \quad (3)$$

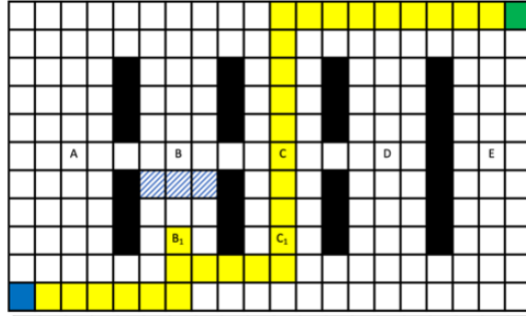


Figure 2: The path for the re-plan policy  $\pi_K^2$  which bypasses aisle B and goes down aisle C.

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 non-zero cost of the action  $u_w$ . This action is state preserving (i.e.  $\mathbf{x}_{k+1} = f(\mathbf{x}_k, \mathbf{u}_w) = \mathbf{x}_k$ ) and is used to represent the action of the robot waiting for one unit of time.

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

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)

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

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

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

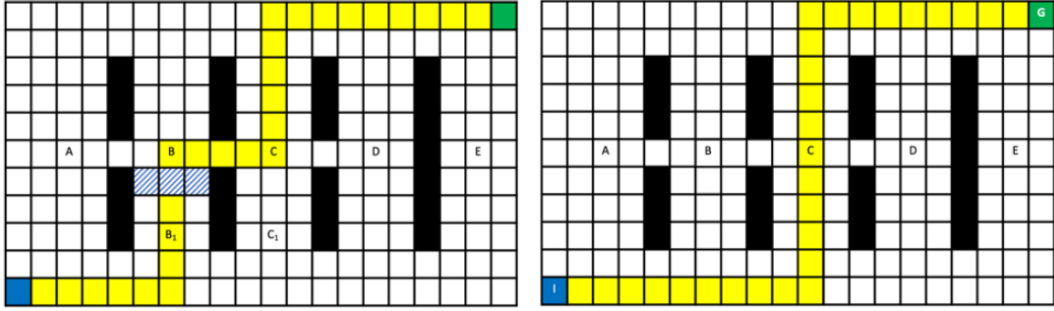
The quantity  $\mathbb{E}[T]$  is the expected value for the time the robot has to wait for the obstacle to clear. As we know the distribution that the variable is sampled from we can compute the expected value. The expected value for the time taken is given by the expression found in Eq. (7)

$$\begin{aligned} \mathbb{E}[T] &= \mathbb{E}\left[\frac{0.5}{\lambda_B} + \tilde{T}\right] \\ &= \frac{0.5}{\lambda_B} + \mathbb{E}[\tilde{T}] \\ &= \frac{0.5}{\lambda_B} + \int_0^\infty 2\lambda_B t e^{-2\lambda_B t} dt \\ &= \frac{0.5}{\lambda_B} + 2\lambda_B \left[ (t) \left( -\frac{1}{2\lambda_B} e^{-2\lambda_B t} \right) + \frac{1}{2\lambda_B} \int_0^\infty e^{-2\lambda_B t} dt \right]_0^\infty \\ &= \frac{0.5}{\lambda_B} + 2\lambda_B \left[ (t) \left( -\frac{1}{2\lambda_B} e^{-2\lambda_B t} \right) - \left( \frac{1}{2\lambda_B} \right)^2 e^{-2\lambda_B t} \right]_0^\infty \\ &= \frac{0.5}{\lambda_B} - \left[ t e^{-2\lambda_B t} + \frac{1}{2\lambda_B} e^{-2\lambda_B t} \right]_0^\infty \\ &= \frac{0.5}{\lambda_B} - (0 + 0 - 0 - \frac{1}{2\lambda_B}) = \frac{0.5}{\lambda_B} + \frac{1}{2\lambda_B} = \frac{1}{\lambda_B} \end{aligned} \quad (7)$$

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

$$\begin{aligned} \frac{1}{\lambda_B} &\leq \frac{L_{B_1C_1} + L_{C_1C} - L_{B_1B} - L_{BC}}{L_W} \\ \lambda_B &\geq \frac{L_W}{L_{B_1C_1} + L_{C_1C} - L_{B_1B} - L_{BC}} \end{aligned} \quad (8)$$

## 1.2 Policy Selection at Start



(a) The scenario where the robot decides to go down Aisle B, encounters an obstacle 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 starts 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). By avoiding travelling down aisle B before replanning the robot avoids traversing the extra distance to the cell marked  $B_1$  which makes this approach favorable.

Similar to the approach in §1.1, we may denote the policy that goes down aisle B as  $\pi_K^1$  and the policy that goes down aisle C as  $\pi_K^2$ . Once again these are padded with zero-cost state preserving actions so as to be of the same dimension.

Fig. 3 depicts the two possible paths that represent the two policies. The two different policies and their costs are as follows:

- Drive down aisle B and wait for the obstacle to clear. The cost of this scenario is given by the expression in Eq. (9).

$$L(\pi_K^1) = L_{IB_1} + TL_W + L_{B_1B} + L_{BC} + L_{CG} \quad (9)$$

- Completely avoid aisle B and traverse directly to the goal via aisle C. It is known that no obstacle exists on this aisle. The cost for this policy is given in Eq. (10).

$$L(\pi_K^2) = L_{IC} + L_{CG} \quad (10)$$

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. Policy  $\pi_K^2$  will be chosen over policy  $\pi_K^1$  if the expected value of it's cost is lower as shown by Eq. (11).

$$\mathbb{E}[L(\pi_K^2)] \leq \mathbb{E}[L(\pi_K^1)] \quad (11)$$

Then, substituting the loss functions from Eq. (9) and (10) into the inequality in Eq. (11), we obtain the result shown in Eq. (12).

$$\begin{aligned} \mathbb{E}[L_{IC} + L_{CG}] &\leq \mathbb{E}[L_{IB_1} + TL_W + L_{B_1B} + L_{BC} + L_{CG}] \\ \mathbb{E}[T]L_W &\geq L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC} \\ \mathbb{E}[T] &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W} \end{aligned} \quad (12)$$

By substituting the expected value for  $T$  found in Eq. (7) we obtain a constraint for  $\lambda_B$  as shown in Eq. (13). The right hand side of the final inequality in Eq. (13) represents the largest value for  $\lambda_B$  for which the robot will decide to go directly down aisle C avoiding aisle B. Once again the sultions for  $\lambda_B < 0$  have been ignored.

$$\begin{aligned} \frac{1}{\lambda_B} &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W} \\ \lambda_B &\leq \frac{L_W}{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}} \end{aligned} \quad (13)$$

### 1.3 Considering the Probability of the Obstacle Being Present

In reality, the obstacle would not be present there all the time. This can be taken into account using a probability, say  $p_B$  associated with the scenario that the obstacle is present<sup>1</sup>. The mean wait time in the case the obstacle is present has been derived in Eq. (7). In the case the obstacle isn't present, the robot does not have wait and therefore the mean wait time is 0. By taking a sum of these wait times weighted by their probabilities of occurring we obtain the expected wait time that takes into consideration the probability of the obstacle being present as shown below in Eq. (14):

$$\begin{aligned} \mathbb{E}_B[T] &= p_B \cdot \mathbb{E}[T] + (1 - p_B) \cdot (0) \\ &= p_B \cdot \frac{1}{\lambda_B} = \frac{p_B}{\lambda_B} \end{aligned} \quad (14)$$

where,  $\mathbb{E}[T]$  is expected time to wait given the obstacle in aisle B exist and it is given by Eq. 6 Given Eq. (14) we may reconsider Eq. (12) and substitute this new value for the expected wait time:

$$\begin{aligned} \mathbb{E}_B[T] &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W} \\ \frac{p_B}{\lambda_B} &\geq \frac{L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC}}{L_W} \end{aligned} \quad (15)$$

Assuming that the rate parameter  $\lambda_B$  is constant, we may arrive at the equation shown in Eq. (16).

$$p_B \geq \frac{\lambda_B (L_{IC} - L_{IB_1} - L_{B_1B} - L_{BC})}{L_W} \quad (16)$$

It should be noted however that this was the inequality to choose the policy to plan straight through aisle C. Therefore the R.H.S of the inequality in Eq. (16) represents the value for  $p_B$  below which the robot will attempt to drive down aisle B first.

---

<sup>1</sup>This means that the probability that the obstacle is absent can be given by  $(1 - p_B)$

## 1.4 Considering Multiple Obstacles

There can be 4 cases:

→ Driving down to aisle B and waiting (17)

$$L_{IB_1} + V_W T_B + L_{B_1V} + L_{BC} + L_{CG} \quad (17)$$

→ Driving down to aisle B encountering an obstacle, driving down aisle C (18)

$$L_{IB_1} + L_{B_1C} + L_{CG} \quad (18)$$

→ Driving down B, encountering an obstacle, waiting at C (19)

$$L_{IB_1} + L_{B_1C_1} + L_W T_C + L_{C_1C} + L_{CG} \quad (19)$$

→ Drivign down B, encountering an obstacle, driving down C, encountering and obstacle, drivign down D (20)

$$L_{IB_1} + L_{B_1C_1} + L_{C_1D} + L_{DG} \quad (20)$$

Ignring A because it is going to have the highest cost as the planner attempts moving to the goal through the shortest path

Average of the 4 costs are:

$$L_D + L_{DG} \leq \frac{\eta}{4}$$

where:

$$\eta = 4L_{IB_1} + 2L_{B_1C_1} + L_W T_B + L_{B_1B} + L_{BC} + 3L_{CG} + L_{B_1C} + L_{CG} + L_W T_C + L_{C_1G} + L_{DG} + L_{GD}$$

$$\mathbb{E}[T_B] = \frac{P_B}{\lambda_B}; \mathbb{E}[T_C] = \frac{P_C}{\lambda_C} \quad (21)$$

## 2 ROS Implementation

## References

# Appendices