Statistics - Research Track 2

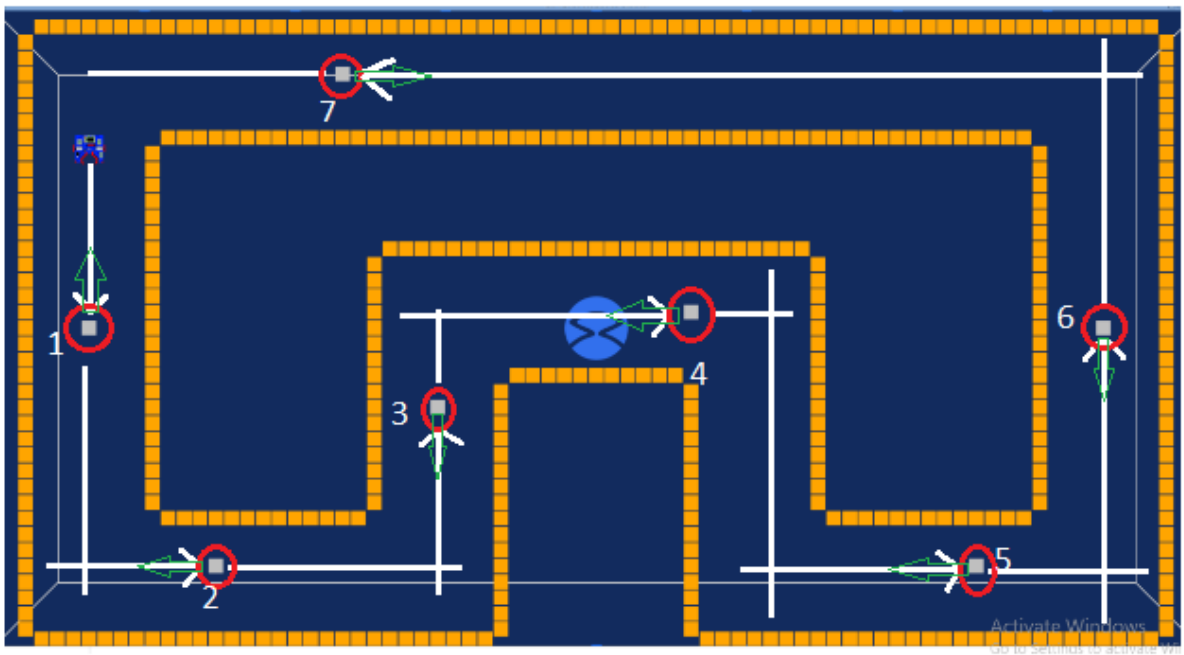
Introduction

This work is the assignment of the RT_2 course, here I have to collect the data of one of my pervious Assignments of RT_1, the project on which the analyses is performed is the first assignment of the course "Research Track 1" The project aims to make a homonymic robot move inside a maze without hitting walls made of golden boxes. Furthermore, inside the maze, there are many silver tokens that the robot have to grab, move them behind him and then start again with the search for the next tokens and I have to collect the data and comparison to Prof.'s robot, which is available in this repository: https://github.com/CarmineD8/python_simulator/tree/rt2

Data Collection

I analyzed the 5 laps of the project in each the location of the silver crates and the Robot's moving position is different from previous lap so does the behavior of the Robot and also time to complete the each will also be different . Both my robot and the professor's robot were run for five laps in each of the five configurations, acquiring run-time the following information:

- **DISTANCE:** the robot's distance from the obstacles on the
- **ELAPSED:** The time it took the robot to complete the laps.
- WRONG DIRECTON**:** The frequency with which the robot wrongly changes direct



As seen in the image there are total 7 silver boxes that robot have to grab and put behind, the robot moves in the counter clockwise direction and the time taken by the robot to complete one lap of the circuit is calculated as the time it takes the robot to process 7 silver boxes . As a result, in order to avoid distorting the time statistics, I prevented varying the position of the seventh box significantly so does not have a large excursion

Data Analysis

Data analysis in done on the Matlab, the data is organized into cells by diving the particular distance from obstacles for each lap performed by the robot.

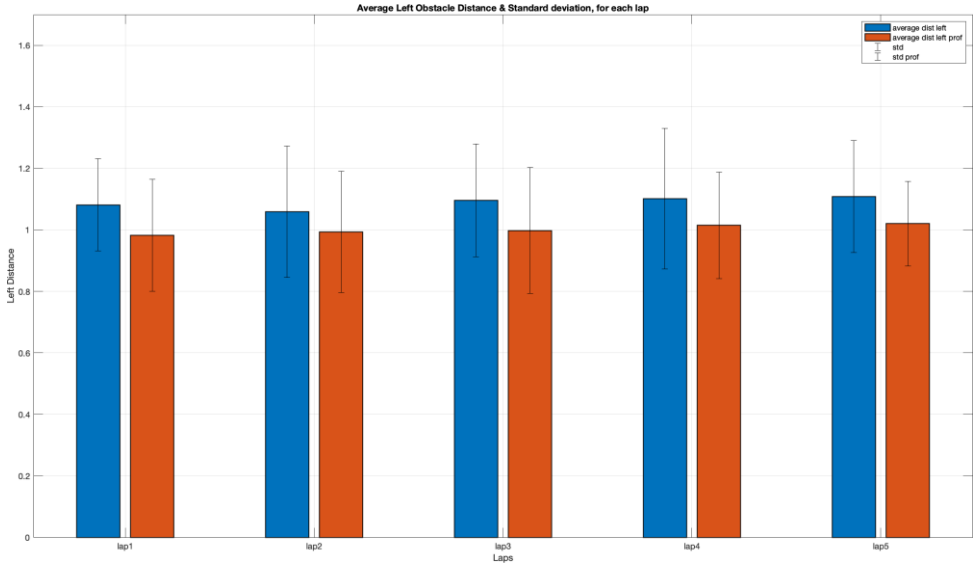
Obstacles Distance

Concerning the distance to obstacles, the differences between the average distance held by my robon both sides and that held by the professor's robot.

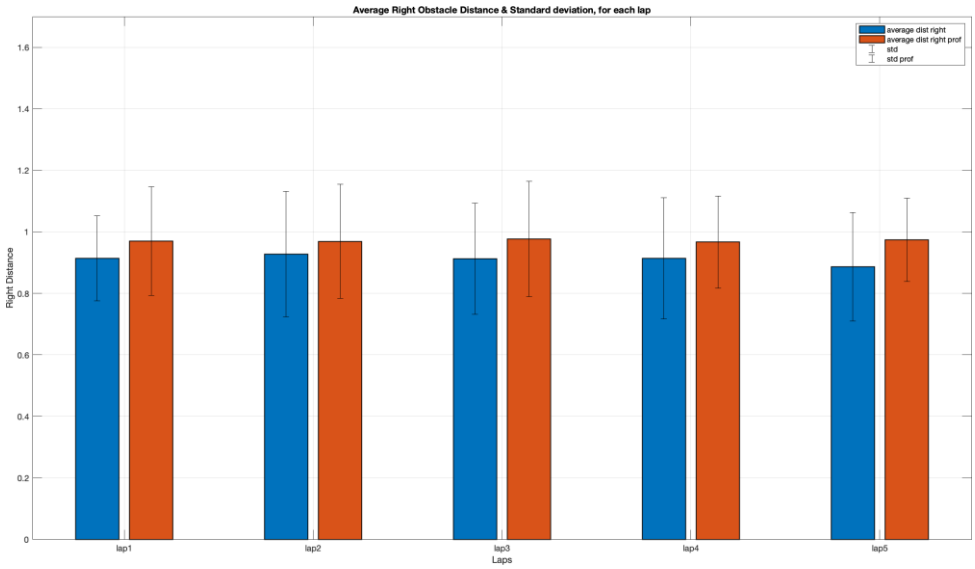
The graph below reflects the average trend of the distance held by the robots to their left during the first and second laps of the circuit. Distances greater than 2 that I considered insignificant were eliminated because I was interested in proximity to obstacles. These discarded values were interpolated into the vector's rest. Moreover, to improve the visibility of the average trend in the graph, the data were smoothed by graphing the average envelope of the data. The shading represents the standard deviation.

It is also noticed that my robot does not just finish the first and second lap faster than the professor's robot, the time gradually increases from one lap to another.Thus I representing only two laps here; the time difference would have made the graph indecipherable.

This bar plot shows that on average my robot stands fartherfrom the left-hand obstacles per lap. The standard deviation is also clearly visible.



Similarly, the below graphs that represent the comparison of distance between robots from the obstacles to right, the average distance to obstacles on the right turnsout to be less for my robot than for the professor's robot



BASIC CONCEPTS

Null hypothesis and alternative hypothesis, are Explained in these context of testing, If we are to compare method A with method B about its superiority and if we proceed on the assumption that both methods are equally good, then this assumption is termed as the null hypothesis. Similarly, if the method A is superior or the method B is inferior, then this known as alternative hypothesis. The null hypothesis is generally symbolized as "**Ho**" and the alternative hypothesis as "**Ha**". If our sample results do not support this null hypothesis, means rejecting the null hypothesis so other is alternative hypothesis or vice versa.

T-Test

The t-test is a tool for assessing the means of one or two populations using a hypothesis test. A t-test is needed to evaluate whether a single group deviates from known values (1 sample t-test), or the two groups are different from each other (2 independent sample t-tests), or exists a significant difference between the paired measurements (paired or dependent sample tests), which is our case**.** As mentioned earlier, we chose the paired sample test because we need to compare the same experiment with two different implementations.
At First I checked for the paired t-test precondition: the samples should belong to a normal density population. The made use of a Lilliefors test, It is used to test the null hypothesis that data come from a normally distributed population.

The null hypothesis has been accepted (h = 0) at 5% significance level, so the right distance comes from a normally distributed population with a probability Report Statistics - Research Track 2 5 given by the p-value of ~ 70 % , I considered this result acceptable to continue the analysis with the paired t-test

Lilliefors Test - Left Distance	**h=0**	**P=0.7383**

The null hypothesis has been accepted (h = 0) at 5% significance level, so the right distance comes from a normally distributed population with a probability given by the p-value of ~ 30 % ,Although this outcome is not-significantly , I considered this result acceptable to continue the analysis with the paired t-test

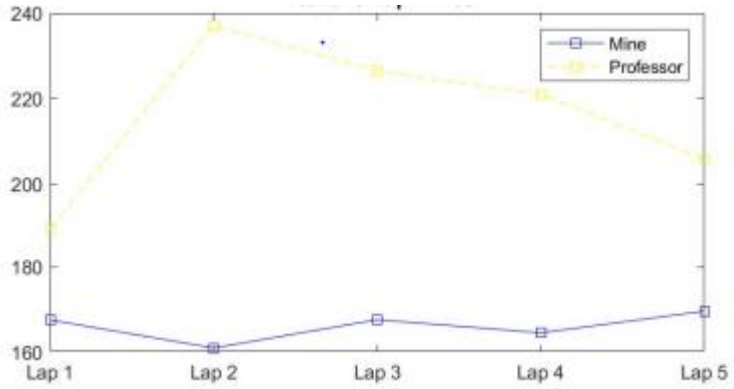
Lilliefors Test - Left Distance	**h = 0**	**p = 0.321**

For both the Test the p-value has been determined using a Monte Carlo approximation with a maximum Monte Carlo standard error of 0.05. The Monte Carlo standard error is the error due to simulating the p-value**.**
The Paired T-test has been performed between the mean minor distance between the robot ad obstacles in each lap kept by my robot and the one kept by the professor's robot.
Both the Tests for left and right distances reject the null hypothesis (h=1) at 5% significance level, that means that there's a difference between the two compared robots, The low amplitude of the p-value means that I'm sure with rejection of the null hypothesis, Therefore the outcome is significantly. The difference between my value and the professor's was also visible by the graphs mainly due to the different behavior in the maze.

Paired T-test - Left Distance	**Paired T-test - Right Distance**
h = 1	**h = 1**
p = 0.0014	**p = 1.4639e-04**

Lap Time

This Graph represent the medium time to conclude a lap of the circuit. The evidence is that my robot is faster than the professor's one, we can also see that the time increase from one lap to another. Moreover you can notice a decreasing of the time spent to conclude the fifth lap, I noticed that this is mainly due to the fact that the silver boxes are, in average, well organized after several laps and that results in a decrease in time."



Wrong Direction

As mentioned in the data collection, I have been able to associate the silver boxes with a code ranging from one to 7.
If the robot picks up a crate with the same row code twice, it has misdirected.
The one given below is not a true statistic because the misdirecting event is extremely rare and I have too few samples to generate any significance. However, it may be useful to note the percentage of times my robot and the professor's robot went in the wrong direction, and to correlate this data with circuit laps.

Only for this part, I decided to increase the number of trials coming to a total of 10 running's, 5 laps each.
Theconfiguration of the crates in the circuit have always been randomized.

My robot	
No: of laps	**Chances of wrong direction**
Lap:1	**0%**
Lap:2	**7%**
Lap:3	**19%**
Lap:4	**0%**
Lap:5	**7%**

Professor's robot	
No: of laps	**Chances of wrong direction**
Lap:1	**3%**
Lap:2	**21%**
Lap:3	**14%**
Lap:4	**0%**
Lap:5	**4%**

Conclusion

The main objective of this statistical part of the project is to find particular differences between my algorithm and algorithm's prof that "guide" the robot inside the environment. I have taken lots of data, since I know results are non-deterministic, in the end I found that there are substantial differences about lap times and distances from obstacles. In fact my robot maintain a longer distances from the walls, this is very important since many wall-crush can be avoided. My robot is faster than the robot's professor in concluding laps.