UNIVERSITÀ DEGLI STUDI DI GENOVA

DIBRIS

DEPARTMENT OF COMPUTER SCIENCE AND TECHNOLOGY,
BIOENGINEERING, ROBOTICS AND SYSTEM ENGINEERING

RESEARCH TRACK 2

Second Assignment

Statistic

Author:

Pumpo Raffaele

Student ID:

s5457102

Professors:

Carmine Recchiuto

May 26, 2023

Contents

1 Hypotheses made

In this experiment have been analysed two type of algorithms, which have been implemented from two different developers.
The algorithms have the goal of controlling a mobile robot, equipped of grab, that moves in an arena with 12 token, 6 gold and the same number silver.
The task is to bring each silver token close to a golden one, and when all those are paired, the simulation finishes.
The alternative hypothesis made for this statistic is that the algorithm A, implemented by Raffaele Pumpo, is faster, on average, than the algorithm B, implemented by Emanuele Buzzurro, to complete the task.
Therefore, as in any statistical test, null hypothesis, states that there is no significant difference between the two algorithms, and the goal is to reject that hypothesis in favor of the alternative one, thus demonstrating the existence of a significant difference.

2 Description and motivation of the experimental setup

The experiment focused on modifying the positions of obstacles within the arena. The objective was to investigate the algorithms' performance in the presence of obstacles and assess their ability to navigate around or through them. By strategically placing obstacles at different locations, the algorithms' adaptability has be tested.
To ensure reliable and statistically significant results, the experiment was repeated 30 times. This number of repetitions was chosen to minimize the influence of random variations and provide a robust dataset for subsequent statistical analysis. By conducting multiple repetitions, any inconsistencies or outliers could be identified, and the overall performance trends of the algorithms could be established.
The setup enabled a direct comparison between the two algorithms. By subjecting both algorithms to the same experimental conditions, it became possible to evaluate their relative strengths and weaknesses. Understanding the contrasting performance characteristics of the algorithms in different configurations could provide valuable insights for selecting the most suitable algorithm for specific robotic tasks.

3 Results

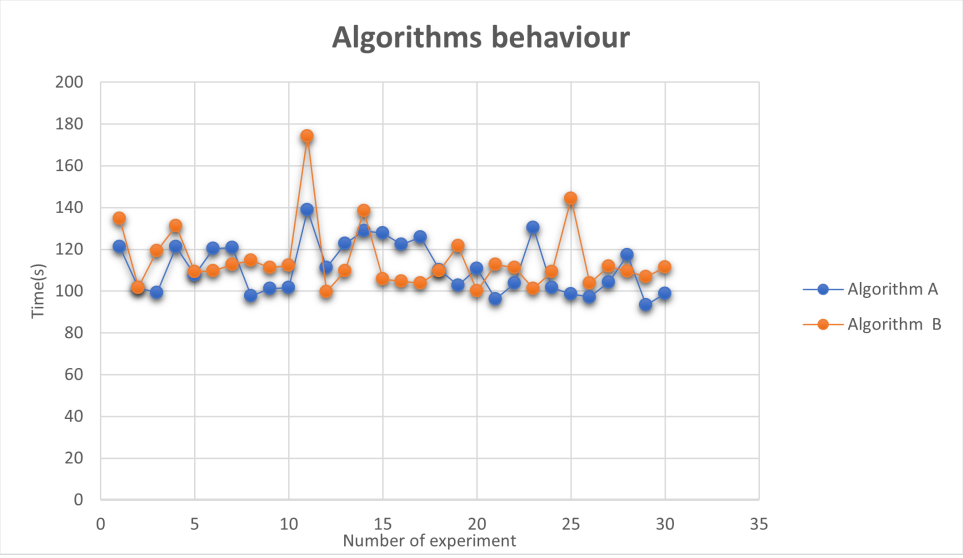
The table in *Figure 1* presents the values of time statistics obtained from the two robot control algorithms in the conducted analysis.

The table provides a clear and concise overview of the time measurements for each algorithm, allowing for easy comparison between the two. The values in the table were obtained through the function ***random.seed()*** in such way putting the same value inside that, it generates the same random value for both algorithms, used as ***angle offset*** parameter, which change the obstacles positions.
This value is written in the table in the third column *(Seed)*.

Attempts	Algorithm A	Algorithm B	Seed
1	121.2214799	134.8373151	10
2	101.702688	101.733258	11
3	99.20445395	119.2783868	12
4	121.206516	131.273895	13
5	107.719151	109.2243431	14
6	120.280859	109.732805	15
7	120.7751992	112.7566001	16
8	97.69369698	114.732435	17
9	101.2019579	111.214663	18
10	101.7090321	112.236244	19
11	138.9172201	174.128315	20
12	111.2538671	99.72087002	21
13	122.732522	109.7373991	22
14	128.7574501	138.450841	23
15	127.7774251	105.6971152	24
16	122.240169	104.7319191	25
17	125.741663	103.6987329	26
18	110.2269199	109.7254181	27
19	102.682215	121.7553599	28
20	110.6832299	100.200933	29
21	96.15349913	112.7472529	30
22	103.69818	111.2472432	31
23	130.285403	101.1803241	32
24	101.7081621	109.271261	33
25	98.65636897	144.263567	34
26	97.15530705	103.7153721	35
27	104.3388999	111.8118179	36
28	117.2997921	109.4099331	37
29	93.36212611	106.8727081	38
30	98.79584193	111.5123539	39

4 Discussion of the results with statistical analysis

Analyzing the execution times of the algorithms is crucial to evaluate their efficiency and adaptability to different arena conditions. The recorded times reflect the time taken by the robots controlled by the algorithms to finish the task in each attempt.
Through the table, it will be possible to identify any trends or significant differences in the performance of the two algorithms in different arena configurations. These data will provide a solid foundation for discussing the results and implications about the algorithm. We can observe the trends of the two algorithms through the following figure:



The statistic method chosen has been the T-test. Through different parameters. the aim of this test is to estimate the t-value and the associated p-value. If the p-value is lower than a predetermined level of significance (usually 0.05), the null hypothesis is rejected, and it is concluded that there are significant differences between the groups or conditions under consideration.

Firstly, It has been computed the mean and standard deviation of the two algorithms respectively $\mu_1, \mu_2, \sigma_1, \sigma_2$, as following:

$\mu_1 = \frac{\sum_{n=1}^{N1} X_i}{N1}$ $\mu_2 = \frac{\sum_{n=1}^{N2} X_i}{N2}$

$\sigma_1 = \sqrt{\frac{\sum_{n=1}^{N1} (X_i - \mu_1)^2}{N1}}$ $\sigma_2 = \sqrt{\frac{\sum_{n=1}^{N1} (X_i - \mu_2)^2}{N1}}$

Where *N1=N2=30* are the number of experiments, X_i is the value, in our case the time, for the specific attempt.
In the study the values computed are the following:

$\mu_1 = 111,1727098$ $\mu_2 = 114,8966227$

$\sigma_1 = 12,50205343$ $\sigma_2 = 15,76357864$

To compute the t-value, goal of the study, other parameters are needed:

- Pooled SD : It represents an estimation of the overall standard deviation that takes into account the different standard deviations of the two groups and the sample sizes.

$$\sigma^2 pooled = \sqrt{\frac{(N1-1)\cdot\sigma 1^2 + (N2-1)\cdot\sigma 2^2}{N1+N2-2}}$$

- The standard error *SE* indicates how much the estimate of the difference between the means of the groups can vary from sample to sample.

$$SE = \sqrt{\sigma^2 pooled(\frac{1}{N1} + \frac{1}{N1})}$$

The values computed of these parameters are the following:

$\sigma^2 pooled = 202,3958758$ $SE = 3,673289858$

Once obtained these parameters is possible to compute the parameter useful to conclude the t-test, the **t-value**.
It measures the difference between the means of the two groups in terms of standard deviation units and provides an indication of the statistical significance of that difference, as it is possible to understand from its formula:

$$t_{\mu 1-\mu 2} = \frac{\mu 1 - \mu 2}{SE}$$

In the statistical study done, it's value is:

$$t_{\mu 1-\mu 2} = 1,013781381$$

5 Conclusion

Based on the t-test conducted on the two control algorithm of the same robot in the same arena, the t-value of 1.013781381, and the alternative hypotesis that the first algorithm, made by Raffaele Pumpo, is faster than the second, made by Emanuele Buzzurro. In this study, a total of 30 experiments were performed, with the arrangement of objects in the arena being varied.

The obtained t-value of 1.013781381 indicates a moderate difference between the means of the two algorithms.
However, to assess the statistical significance of this difference, it needs to be compared against the critical value at the desired significance level, considering 0.05, and degrees of freedom, in our case 58.
Using a t-distribution table or software, you can find the corresponding critical value with these parameters, it is $t_c = 1.672$.
Since that the t value found in the study is $t_{\mu 1-\mu 2} = 1,013781381$ and it is less of t_c, then we fail to reject the null hypothesis.
This implies that there is **not** enough evidence to support the claim that the first algorithm is significantly faster than the second.

In conclusion, based on the t-test results, the difference in speed between the two algorithms is not statistically significant. However, further studies and experiments could change this conclusion.