Lab Report of Research Track 2 - Assignment 3

Emanuele Bua Odetti and Alessandro Trovatello

Abstract—This study investigates the impact of bot velocity on reliability, defined as the consistency in performing tasks across various environments. The **Null Hypothesis (H0)** posits that bot reliability is independent of velocity, while the **Alternative Hypothesis (H1)** suggests that reliability is influenced by velocity. Two bots were tested in three environments: 6 tokens in a circle, 8 tokens in a circle, and 6 tokens placed randomly, with each bot undergoing 15 tests per environment. Results indicate significant differences in bot performance, providing insights into how velocity affects reliability and guiding future bot design optimizations.

Keywords—*velocity, reliability, statistics, t-test*

Contents

1. Hypotheses made

Null hypothesis (H0): The reliability of the bots is not affected by their velocity.
Alternative Hypothesis (H1): The reliability of the bots is influenced by their velocity.

2. Description and motivation of the experimental setup

The motivation behind the experiment is to assess whether the velocity of the bots has any impact on their reliability. By testing them in three different environments with varying numbers and configurations of tokens, we aim to provide a comprehensive understanding of how the bots performance is affected by their speed.

In each environment:

- 6 tokens in a circle[1]: This environment represents a simple, structured layout.
- 8 tokens in a circle[2]: This environment introduces a higher token density, challenging the bots' navigation abilities.
- 6 tokens placed randomly[3]: This environment simulates unpredictable conditions, requiring adaptability from the bots.

Each bot will undergo 15 tests in each environment, allowing for a thorough evaluation of their reliability across different scenarios. The results of these tests will provide valuable insights into the relationship between bot velocity and reliability, aiding in the optimization of future bot designs and applications.

2.1. Figures

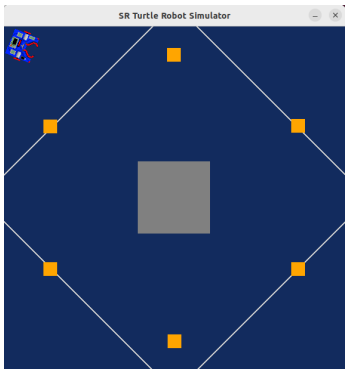


Figure 1. Environment 1: 6 tokens placed in circle.

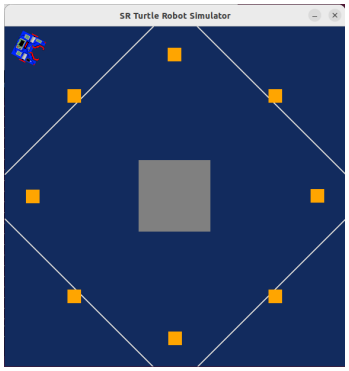


Figure 2. Environment 2: 8 tokens placed in circle.

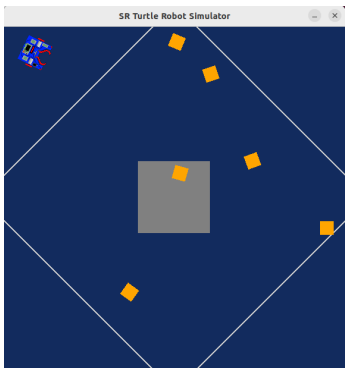


Figure 3. Environment 3: 6 tokens placed randomly. (Every time the code is executed, the placement is changed.)

2.2. Formulas Used

2.2.1. Normalization

The formula (1) shows how to normalize a value x:

$$x' = \frac{x - \min(x)}{\max(x) - \min(x)} \tag{1}$$

where:

- x is the original value,
- x' is the normalized value,
- $\min(x)$ is the minimum value in the dataset,
- $\max(x)$ is the maximum value in the dataset.

2.2.2. Mean

The mean (average) of a set of values $x_1, x_2, \ldots, x_n$ is the sum of the values divided by the number of values:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{2}$$

where $\bar{x}$ is the mean, n is the number of values, and x_i represents each value.

2.2.3. Standard Deviation

The standard deviation is a measure of the amount of variation or dispersion in a set of values:

$$\sigma = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2} \tag{3}$$

where σ is the standard deviation, n is the number of values, x_i represents each value, and $\bar{x}$ is the mean.

T-Test for Paired Samples

For paired samples $(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)$:

2.2.4. Mean Difference

The mean difference is the average of the differences between each pair of values:

$$\bar{d} = \frac{1}{n}\sum_{i=1}^{n}(x_i - y_i) \tag{4}$$

where $\bar{d}$ is the mean difference and $d_i = x_i - y_i$ is the difference for each pair.

2.2.5. Standard Deviation of Differences

The standard deviation of the differences is calculated as:

$$s_d = \sqrt{\frac{1}{n-1}\sum_{i=1}^{n}(d_i - \bar{d})^2} \tag{5}$$

where s_d is the standard deviation of the differences, d_i is each difference, and $\bar{d}$ is the mean difference.

2.2.6. T-Statistic

The t-statistic for the paired samples t-test is calculated as:

$$t = \frac{\bar{d}}{s_d/\sqrt{n}} \tag{6}$$

where t is the t-statistic, $\bar{d}$ is the mean difference, s_d is the standard deviation of the differences, and n is the number of pairs.

2.2.7. Degrees of Freedom

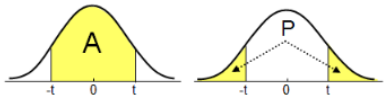
The degrees of freedom for the paired samples t-test is given by:

$$\text{df} = n - 1 \tag{7}$$

where df is the degrees of freedom and n is the number of pairs.

2.2.8. Critical Value and P-Value

The critical value for a given confidence level and degrees of freedom can be found using the t-distribution table. The p-value indicates the probability of obtaining a result at least as extreme as the one observed, under the assumption that the null hypothesis is true.



DF	A P	0.80 0.20	0.90 0.10	0.95 0.05	0.98 0.02	0.99 0.01	0.995 0.005	0.998 0.002	0.999 0.001
1		3.078	6.314	12.706	31.820	63.657	127.321	318.309	636.619
2		1.886	2.920	4.303	6.965	9.925	14.089	22.327	31.599
3		1.638	2.353	3.182	4.541	5.841	7.453	10.215	12.924
4		1.533	2.132	2.776	3.747	4.604	5.598	7.173	8.610
5		1.476	2.015	2.571	3.365	4.032	4.773	5.893	6.869
6		1.440	1.943	2.447	3.143	3.707	4.317	5.208	5.959
7		1.415	1.895	2.365	2.998	3.499	4.029	4.785	5.408
8		1.397	1.860	2.306	2.897	3.355	3.833	4.501	5.041
9		1.383	1.833	2.262	2.821	3.250	3.690	4.297	4.781
10		1.372	1.812	2.228	2.764	3.169	3.581	4.144	4.587
11		1.363	1.796	2.201	2.718	3.106	3.497	4.025	4.437
12		1.356	1.782	2.179	2.681	3.055	3.428	3.930	4.318
13		1.350	1.771	2.160	2.650	3.012	3.372	3.852	4.221
14		1.345	1.761	2.145	2.625	2.977	3.326	3.787	4.140
15		1.341	1.753	2.131	2.602	2.947	3.286	3.733	4.073

Figure 4. T-Distribution Table.

2.2.9. Pearson Correlation Coefficient

The Pearson correlation coefficient r measures the strength and direction of the linear relationship between two variables:

$$r = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}\sqrt{\sum_{i=1}^{n}(y_i - \bar{y})^2}} \tag{8}$$

where r is the Pearson correlation coefficient, x_i and y_i are the individual sample points, and $\bar{x}$ and $\bar{y}$ are the means of the x and y values, respectively.

3. Results

3.1. Environment 1

Execution Time				
n	Bot 1 Time	Bot 1 Norm	Bot 2 Time	Bot 2 Norm
1	54,95987	0,37264	84,65280	0,42076
2	55,27446	0,41775	85,44914	1,00000
3	52,36114	0,00000	84,46955	0,28746
4	54,49469	0,30594	85,05623	0,71420
5	53,37611	0,14554	85,26286	0,86450
6	54,61364	0,32300	85,14409	0,77811
7	58,05734	0,81680	84,78617	0,51777
8	59,33492	1,00000	84,36953	0,21471
9	55,13069	0,39714	85,04686	0,70739
10	59,19231	0,97955	84,77364	0,50865
11	55,94048	0,51326	84,87707	0,58389
12	55,06786	0,38813	84,82687	0,54737
13	54,45313	0,29998	84,47823	0,29378
14	55,04985	0,38555	84,07435	0,00000
15	54,39741	0,29199	84,99850	0,67221

Table 1. Environment 1 - 6 tokens circle

3.2. Environment 2

Execution Time				
n	Bot 1 Time	Bot 1 Norm	Bot 2 Time	Bot 2 Norm
1	78,26205	0,88435	116,657028	0,47663
2	75,80831	0,61791	116,31924	0,42759
3	75,55868	0,59081	116,99526	0,52573
4	78,65301	0,92680	116,06432	0,39058
5	72,04576	0,20935	118,91829	0,80490
6	77,80490	0,83471	115,75485	0,34566
7	76,69070	0,71373	116,05548	0,38930
8	74,82334	0,51096	114,75795	0,20094
9	74,27251	0,45115	113,37381	0,00000
10	70,11775	0,00000	116,49312	0,45283
11	74,89910	0,51918	116,00306	0,38169
12	76,60168	0,70406	119,4113	0,87647
13	79,25859	0,99256	116,68213	0,48027
14	79,32709	1,00000	120,26223	1,00000
15	78,25032	0,88308	116,50902	0,45514

Table 2. Environment 2 - 8 tokens circle

3.3. Environment 3

Execution Time				
n	Bot 1 Time	Bot 1 Norm	Bot 2 Time	Bot 2 Norm
1	40,401829	0,01667	84,97362	0,46923
2	135,85114	0,96113	85,44431	0,47718
3	45,36058	0,06574	85,65030	0,48065
4	42,24703	0,03493	57,17782	0,00000
5	48,69830	0,09876	87,33356	0,50907
6	52,72143	0,13857	86,11785	0,48855
7	49,40143	0,10572	85,95040	0,48572
8	60,81553	0,21866	116,41468	1,00000
9	113,40459	0,73903	107,32852	0,84661
10	42,96383	0,04202	86,54248	0,49572
11	46,06564	0,07271	87,01203	0,50364
12	53,95710	0,15080	96,37539	0,66171
13	139,77931	1,00000	73,42112	0,27421
14	55,21156	0,16321	61,83233	0,07857
15	38,71700	0,00000	90,46885	0,56200

Table 3. Environment 3 - 6 tokens random

4. Discussion of the results with statistical analysis

4.1. Environment 1: 6 Tokens in a Circle

Metric	Bot 1	Bot 2	t-Test
Mean	55.447	84.818	
Normalized Mean	0.442	0.541	
Standard Deviation	1.904	0.353	
Normalized Standard Deviation	0.273	0.257	
Mean Difference			0.447
Variance Difference			0.086
Observations			15
Pearson Correlation			-0.206
Degrees of Freedom			14
t Stat			-0.866
P(T<=t) two-tail			0.401
t Critical two-tail			2.145

4.2. Environment 2: 8 Tokens in a Circle

Metric	Bot 1	Bot 2	t-Test
Mean	76.158	116.684	
Normalized Mean	0.656	0.481	
Standard Deviation	2.565	1.679	
Normalized Standard Deviation	0.279	0.244	
Mean Difference			0.640
Variance Difference			0.085
Observations			15
Pearson Correlation			0.164
Degrees of Freedom			14
t Stat			1.656
P(T<=t) two-tail			0.120
t Critical two-tail			2.145

4.3. Environment 3: 6 Tokens Placed Randomly

Metric	Bot 1	Bot 2	t-Test
Mean	64.373	86.136	
Normalized Mean	0.254	0.489	
Standard Deviation	33.553	14.341	
Normalized Standard Deviation	0.332	0.242	
Mean Difference			0.271
Variance Difference			0.123
Observations			15
Pearson Correlation			0.076
Degrees of Freedom			14
t Stat			-1.955
P(T<=t) two-tail			0.071
t Critical two-tail			2.145

4.4. Statistical Analysis of Bot Performance Across Environments

The *paired t-tests* conducted across the three environments reveal the following: in all environments, the p-values exceed 0.05, indicating no significant difference in reliability between the bots. **These results suggest that bot reliability is not influenced by velocity, thus failing to reject the *Null Hypothesis [H0]*.**

Across all environments, Bot 1 consistently exhibits lower mean execution times, indicating generally better performance compared to Bot 2. However, statistical significance is not attained in any of the environments, suggesting that the observed differences may stem from random variation rather than true disparities in bot performance. Notably, the results from Environment 3 come closest to significance, hinting that Bot 2 may demonstrate enhanced reliability with randomly placed tokens, as indicated by its minor standard deviation.

This analysis underscores the importance of considering both mean performance and variability, alongside conducting comprehensive statistical tests to ascertain the significance of observed differences.

5. References

For additional details and access to the source code, please visit the GitHub repository: RT2 assignment 3

Contact us

emanuelebuaodetti@gmail.com
alessandro.trovatello00@outlook.it