Università degli studi di Genova

DIBRIS

Department of Computer Science and Technology,
Bioengineering, Robotics and System Engineering

Research Track II

Research Track II

First Assignment

Author:

Giordano Emanuele

Student ID:

s4479444

Professors:

Carmine Tommaso Recchiuto

7

Contents

0 Abstract

The purpose of the following assignment is to perform a statistical analysis between the results of two robot control algorithms, which drive a simple differential two-wheeled mobile robot in a pre-defined environment. The two algorithms are evaluated over two different experimental setups, and various observables are taken into consideration during the statistical analysis, namely the average execution time and the accuracy of the robot's performance.

1 Experimental Setups

1.1 Environment & Robot's task

The two robot control algorithms were run in two different experimental setups, which has been devised to test the algorithms' performance and see whether one of them might be better suited for a particular task with respect to the other. The robot is placed in a flat environment, whose coordinates are represented by a pair of points (x, y), and three silver and three golden tokens are placed in the environment. The purpose of the control algorithm is to drive the robot towards the silver tokens, grab them and bring them close to the golden tokens, such as they are coupled as silver-golden.

In particular, the two experimental setups devised in this paper are:

- **Circle**: tokens are placed in a circular fashion around the center of the arena;
- **Random**: randomly scattered in the environment;

The experiments were performed for 50 times, with 3 tokens for each color, for a total of 6 tokens in the environment.

1.2 Data Measurements

Each experiment trial was run for a maximum time of 300 seconds, killing the process at such time, and considering the robot's work done up until that moment. In order to keep track of the quality and robustness of the control algorithms, the following observable quantities has been measured for each trial:

- **Accuracy**: the rate of successful trials - considered such only when the robot has coupled all tokens in the environment, over the total number of trials (50).

  $$Acc = n_s/N \in [0, 1]$$

  Also the failure rate is taken into account, as: $Fail = 1 - Acc$
- **Estimated Arrival Time**: the average time necessary for the robot to complete the task during each trial.

2 Statistical Tests

2.1 Hypothesis Formulation

The following hypothesis formulation is proposed for the statistical tests:

- **Null Hypothesis $\mathbf{H_0}$**: no statistical differences between the two control algorithms A_1 and A_2;
- **Alternative Hypothesis $\mathbf{H_1}$**: there is some statistical difference between the algorithms.

2.2 T-Test and analysis methodology

In order to determine whether the null or alternative hypothesis is true, a paired comparison t-Test: given two random variables X_1 and X_2, each of which has been observed for n_1 and n_2 realizations $[x_1^1, x_2^1, ..., x_{n_1}^1]$ and $[x_1^2, x_2^2, ..., x_{n_2}^2]$, then, the following statistical estimators can be computed:

$$\mu_i = \mathbb{E}[X_i] = \frac{1}{n_i} \sum_{j=1}^{n_i} x_j^i \qquad \textbf{Mean value}$$

$$\sigma_i^2 = \mathbb{E}[X_i^2] - \mathbb{E}[X_i]^2 \qquad \textbf{Variance} \text{ (biased)}$$

$$S_i^2 = \frac{n_i}{n_i - 1} \sigma_i^2 \qquad \textbf{Variance} \text{ (unbiased)}$$

With the expected value μ_i and the unbiased variance S_i^2 for each of the two groups of observations $[x_j^i]$, the paired t-value can be computed as:

$$t = \frac{\mu_1 - \mu_2}{\sqrt{\frac{S_1^2}{n_1} + \frac{S_2^2}{n_2}}}$$

Then, after choosing some level of confidence α for the statistical analysis, and computing the degrees of freedom of the paired t-distribution as:

$$df = n_1 + n_2 - 2$$

the computed t-value is compared with the corresponing value of the table (Fig. 1): if the computed t-value

t Table

cum. prob	$t_{.50}$	$t_{.75}$	$t_{.80}$	$t_{.85}$	$t_{.90}$	$t_{.95}$	$t_{.975}$	$t_{.99}$	$t_{.995}$	$t_{.999}$	$t_{.9995}$
one-tail	0.50	0.25	0.20	0.15	0.10	0.05	0.025	0.01	0.005	0.001	0.0005
two-tails	1.00	0.50	0.40	0.30	0.20	0.10	0.05	0.02	0.01	0.002	0.001
df											
1	0.000	1.000	1.376	1.963	3.078	6.314	12.71	31.82	63.66	318.31	636.62
2	0.000	0.816	1.061	1.386	1.886	2.920	4.303	6.965	9.925	22.327	31.599
3	0.000	0.765	0.978	1.250	1.638	2.353	3.182	4.541	5.841	10.215	12.924
4	0.000	0.741	0.941	1.190	1.533	2.132	2.776	3.747	4.604	7.173	8.610
5	0.000	0.727	0.920	1.156	1.476	2.015	2.571	3.365	4.032	5.893	6.869
6	0.000	0.718	0.906	1.134	1.440	1.943	2.447	3.143	3.707	5.208	5.959
7	0.000	0.711	0.896	1.119	1.415	1.895	2.365	2.998	3.499	4.785	5.408
8	0.000	0.706	0.889	1.108	1.397	1.860	2.306	2.896	3.355	4.501	5.041
9	0.000	0.703	0.883	1.100	1.383	1.833	2.262	2.821	3.250	4.297	4.781
10	0.000	0.700	0.879	1.093	1.372	1.812	2.228	2.764	3.169	4.144	4.587
11	0.000	0.697	0.876	1.088	1.363	1.796	2.201	2.718	3.106	4.025	4.437
12	0.000	0.695	0.873	1.083	1.356	1.782	2.179	2.681	3.055	3.930	4.318
13	0.000	0.694	0.870	1.079	1.350	1.771	2.160	2.650	3.012	3.852	4.221
14	0.000	0.692	0.868	1.076	1.345	1.761	2.145	2.624	2.977	3.787	4.140
15	0.000	0.691	0.866	1.074	1.341	1.753	2.131	2.602	2.947	3.733	4.073
16	0.000	0.690	0.865	1.071	1.337	1.746	2.120	2.583	2.921	3.686	4.015
17	0.000	0.689	0.863	1.069	1.333	1.740	2.110	2.567	2.898	3.646	3.965
18	0.000	0.688	0.862	1.067	1.330	1.734	2.101	2.552	2.878	3.610	3.922
19	0.000	0.688	0.861	1.066	1.328	1.729	2.093	2.539	2.861	3.579	3.883
20	0.000	0.687	0.860	1.064	1.325	1.725	2.086	2.528	2.845	3.552	3.850
21	0.000	0.686	0.859	1.063	1.323	1.721	2.080	2.518	2.831	3.527	3.819
22	0.000	0.686	0.858	1.061	1.321	1.717	2.074	2.508	2.819	3.505	3.792
23	0.000	0.685	0.858	1.060	1.319	1.714	2.069	2.500	2.807	3.485	3.768
24	0.000	0.685	0.857	1.059	1.318	1.711	2.064	2.492	2.797	3.467	3.745
25	0.000	0.684	0.856	1.058	1.316	1.708	2.060	2.485	2.787	3.450	3.725
26	0.000	0.684	0.856	1.058	1.315	1.706	2.056	2.479	2.779	3.435	3.707
27	0.000	0.684	0.855	1.057	1.314	1.703	2.052	2.473	2.771	3.421	3.690
28	0.000	0.683	0.855	1.056	1.313	1.701	2.048	2.467	2.763	3.408	3.674
29	0.000	0.683	0.854	1.055	1.311	1.699	2.045	2.462	2.756	3.396	3.659
30	0.000	0.683	0.854	1.055	1.310	1.697	2.042	2.457	2.750	3.385	3.646
40	0.000	0.681	0.851	1.050	1.303	1.684	2.021	2.423	2.704	3.307	3.551
60	0.000	0.679	0.848	1.045	1.296	1.671	2.000	2.390	2.660	3.232	3.460
80	0.000	0.678	0.846	1.043	1.292	1.664	1.990	2.374	2.639	3.195	3.416
100	0.000	0.677	0.845	1.042	1.290	1.660	1.984	2.364	2.626	3.174	3.390
1000	0.000	0.675	0.842	1.037	1.282	1.646	1.962	2.330	2.581	3.098	3.300
z	0.000	0.674	0.842	1.036	1.282	1.645	1.960	2.326	2.576	3.090	3.291
	0%	50%	60%	70%	80%	90%	95%	98%	99%	99.8%	99.9%
	Confidence Level										

Figure 1: t-values table

happens to be lower or equal to the corresponding value on the table, then the null hypothesis is confirmed, and there's no statistical difference between the two groups of observations. Otherwise, if it's greater, the null hypothesis is rejected in favour to the alternative one. In the end, the t-test will determine whether the two observation groups are coming from the same population, or not; in the presented case, it will tell whether the time complexity of the two control algorithms are the same, or one of them outperforms the other.

3 Experimental Results

The following tables depicts the results of the 50 experiments for each of the two different setups.

	Exp #1: Circle		Exp #2: Random	
	Execution Time	Success	Execution Time	Success
Test 1	80	Yes	126	Yes
Test 2	93	Yes	79	Yes
Test 3	195	No	79	Yes
Test 4	172	Yes	108	Yes
Test 5	118	Yes	131	Yes
Test 6	113	Yes	89	Yes
Test 7	78	No	128	Yes
Test 8	95	Yes	89	Yes
Test 9	141	Yes	91	Yes
Test 10	125	Yes	110	Yes
Test 11	136	Yes	73	Yes
Test 12	99	No	70	Yes
Test 13	146	Yes	114	Yes
Test 14	189	Yes	94	Yes
Test 15	105	Yes	93	Yes
Test 16	150	Yes	98	Yes
Test 17	156	No	70	Yes
Test 18	178	Yes	90	Yes
Test 19	154	Yes	70	Yes
Test 20	97	No	90	Yes
Test 21	166	No	71	Yes
Test 22	262	No	73	Yes
Test 23	92	Yes	90	Yes
Test 24	125	Yes	126	Yes
Test 25	191	No	91	Yes
Test 26	190	Yes	126	Yes
Test 27	185	Yes	94	Yes
Test 28	178	Yes	89	Yes
Test 29	129	No	110	Yes
Test 31	148	Yes	110	Yes
Test 32	99	Yes	114	Yes
Test 33	135	No	90	Yes
Test 34	152	Yes	89	Yes
Test 35	130	Yes	110	Yes
Test 36	105	No	114	Yes
Test 37	118	No	90	Yes
Test 38	163	Yes	108	Yes
Test 39	152	Yes	89	Yes
Test 40	191	No	108	Yes
Test 41	178	Yes	126	Yes
Test 42	80	Yes	90	Yes
Test 43	139	Yes	110	Yes
Test 44	190	No	89	Yes
Test 45	105	Yes	108	Yes
Test 46	178	Yes	90	Yes
Test 47	118	Yes	110	Yes
Test 48	215	No	89	Yes
Test 49	170	Yes	79	Yes
Test 50	135	No	90	Yes

Figure 2: Experimental results of the 50 tests, over the two setups

3.1 Statistical Analysis

From these results, the statistical estimators of mean and variance can be computed, leading to the following results:

Experiment 1 $\mu_{E1} = 140.78\,s \quad S_{E1} = 95.37\,s$

Experiment 2 $\mu_{E2} = 95.30\,s \quad S_{E2} = 16.77\,s$

whereas the second control algorithm, devised by the collegue of the author, has the following means and variances:

$$\mu_{E1'} = 161.68\,s \quad S_{E1'} = 78.80\,s$$

$$\mu_{E2'} = 106.94\,s \quad S_{E2'} = 2.94\,s$$

From here, two paired t-test can be devised over the two pairs of experiments. In particular, the setup presented here will be having:

- $\alpha = 0.05$: A level of significance which implies a 5% chance of rejecting the null hypothesis when this is true;
- $df = 50 + 50 - 2 = 98$: Degrees of freedom of the t-distribution.

3.1.1 Experiment 1

The first experiment's t-value is the following:

$$t_1 = \frac{\mu_{E1} - \mu_{E1'}}{\sqrt{\frac{S^2_{E1}}{n_1} + \frac{S^2_{E1'}}{n_{1'}}}} = 1.19$$

By comparing this value with the corresponding $(\alpha, df) = (0.05, 98)$, which is roughly equal to $(\alpha, df) = (0.05, 100) = 1.984$, it's found that $t_{computed} < t_{table}$, therefore the null hypothesis is confirmed.

3.1.2 Experiment 2

The second experiment's t-value is:

$$t_1 = \frac{\mu_{E2} - \mu_{E2'}}{\sqrt{\frac{s^2_{E2}}{n_2} + \frac{s^2_{E2'}}{n_{2'}}}} = 2.01$$

This means that the computed t-value is greater that $t_{100} = 1.984$, concluding that there is statistical difference between the two control algorithms. In particular, the null hypothesis is rejected, while the alternative is confirmed: this concludes that the control algorithm proposed by the author outperforms the one proposed by their collegue, in the random-scatter experiment. This fact tells that the proposed algorithm is more robust than the other one, handling the presented task in a faster way.

3.2 Accuracy and failure rate

The following tables present the accuracy and fail rate of the two algorithms, over the two different experimental setups.

random-token	Task Completed	Failures	Error%
my-script	34	16	32%
colleague-script	22	28	56%

circular-token	Task Completed	Failures	Error%
my-script	50	0	0
colleague-script	42	8	16%

Figure 3: Accuracy and failure rate, as well as percentage error

4 Sources of errors

After carefully observing the robot's behavour during the trials, the author has found two main sources of error that may affect significantly the results of the analysis:

- **Dragging tokens**: during the robot's motion, it might happen that it collides with other silver tokens, or more seriously, it drags around tokens that have not been chosen as the current target. This leads to more random token positions, as the time goes on, which in the end might cause the robot to not conclude the task within the 300 seconds deadline, and therefore a loss in accuracy.

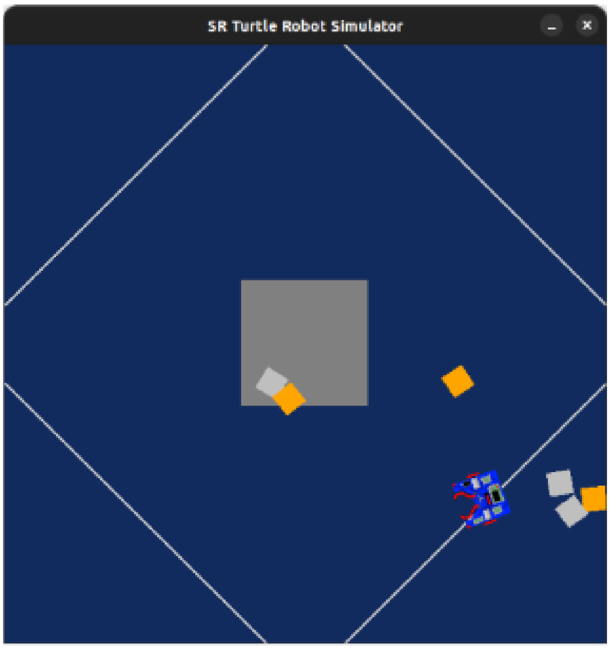

Figure 4: Dragging around tokens might cram them up so much, that the robot might not be able to grab them, or might take too much time

- **Undesired final state**: it may happen that, while driving, the robot drags around silver tokens already placed near the corresponding golden token; this means that, in the end, the trial will be counted as a success, but a silver token might actually be too far away from its golden partner, to be actually considered correct. This would lead to false positives, which increase the accuracy metric without having any significant improvement in the algorithm performances.

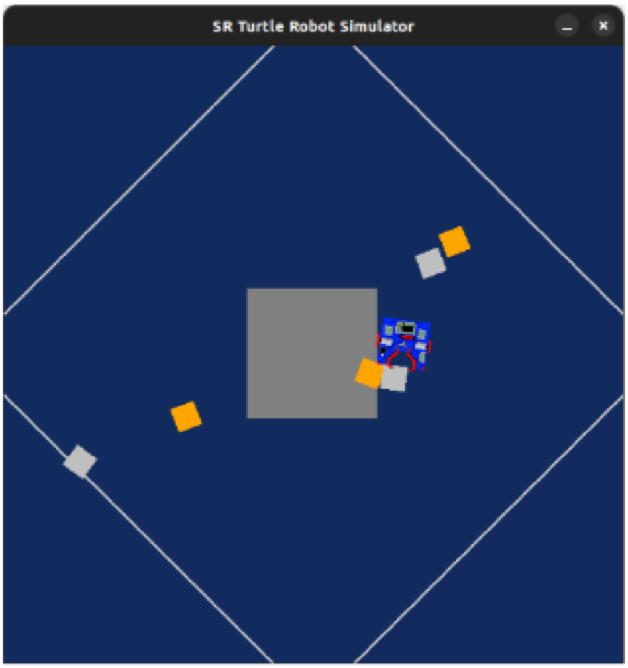

Figure 5: False positives, the silver token is too far away from the golden one, to consider the trial as correct, see Dragging tokens

5 Conclusions

The control algorithm proposed by the author has overall better performances than the one from the author's collegue, being more robust to more complex situations, and thus having lower time complexity than the other one. Nonetheless, the algorithm can be significally improved, by adding a module for collision avoidance, which will drive the robot around the tokens that are not selected as the current target; in this way, both sources of errors that are presented in the section above can be reduced in a single action.