University of Essex

Year 3

Part 2 (research task)

Nathan Jordan

1905341

CE310 Evolutionary Computation and Genetic Programming

Riccardo Poli

Date: 08/03/2023

Contents

Note: This report includes a vast collection of information presented through various figures and tables to offer the reader a comprehensive understanding of the research.

List of Figures

List of Tables

1. Methodology

The two problems that this report analyses are:

Problem 1: $x^2 + x - 7 + 3 * \sin(\pi * x)$

Problem 2: $4 * x * \sin(\frac{5*\pi}{4} * x)$

The value of x is from a list of 65 equally spaced values between -π and +π, passed into these functions, which return y values. For example, figure 1 represents the second function's x and y values.

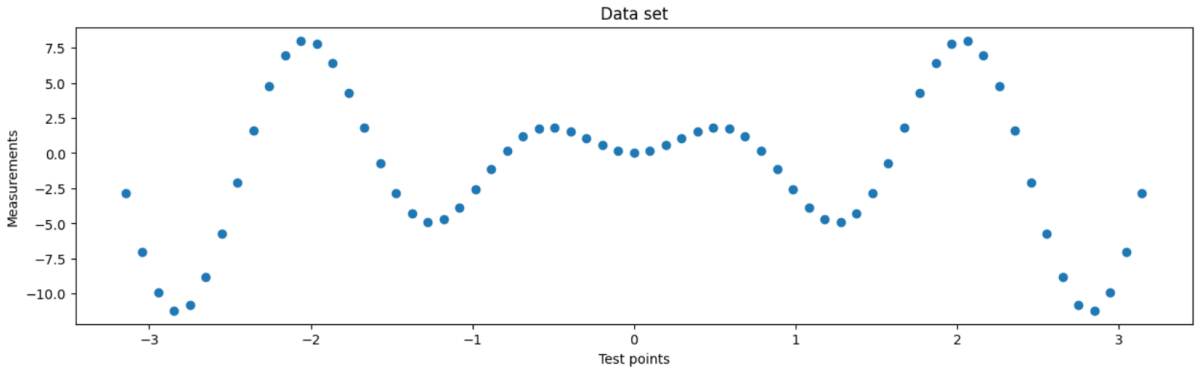

Figure 1: Plot of the x and y values of problem 2.

Overall, eight experiments were carried out, which can be seen in table 1, and each experiment was run ten times for comprehensive results. The other parameters were set to the below:

- Number of generations: 30
- Crossover rate: 0.7
- Mutation rate: 0.3
- Use the least square approach: True

Table 1: Details of experiments carried out.

Experiment Number	Problem	Population Size	Tournament Size
1	1	2000	2
2	1	500	2
3	1	2000	5
4	1	500	5
5	2	2000	2
6	2	500	2
7	2	2000	5
8	2	500	5

2. Results

The results of each of the eight experiments will be compared against each other in this section.

In this section, the results of the experiments mention fitness and size. Fitness measures how close a solution is to the target function, which in this GP is minimised with more minor fitness scores resulting in better solutions. On the other hand, size measures how many instructions are in the resultant function. The maximum fitness statistic was always huge, which invalidates the average and standard deviation as there is no clear correlation; this is why, for the fitness, only the median and minimum are used for analysis; this is likely due to the GP generating functions that are highly exponential and will result in a very high fitness score since it will be very far from the target function.

General observations can be seen after performing all the experiments. For example, as the average median fitness reduces, the average median size increases; this can be seen in figure 2, which is for problem 2. However, this is also the same for problem 1; experiments with the same problem start with roughly the same fitness and size values.

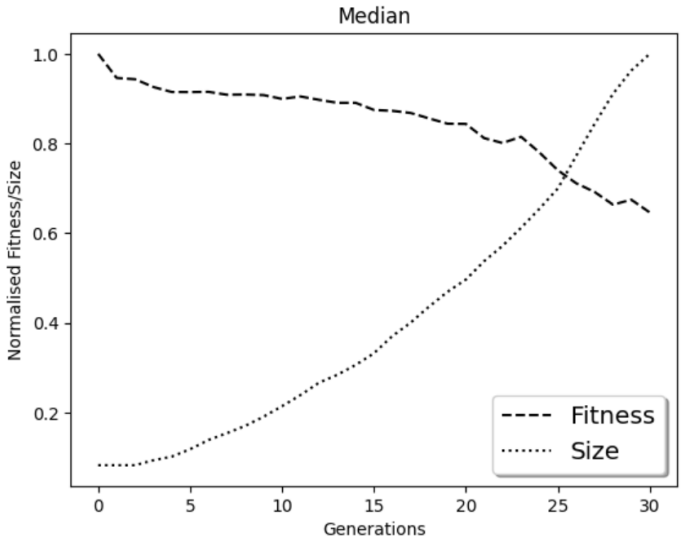

Figure 2: Median fitness plotted against the median size for experiment 8.

The average run time was recorded and displayed in table 2. Based on these results, it can be observed that:

- Problem 1 is more computationally expensive than problem 2.
- As the population and tournament size increase, more time is taken to compute each run.
- There is a trade-off between the population/ tournament size in terms of the time taken, the runs will take more time, but they may result in better solutions.

Table 2: Average runtime of each experiment.

Experiment Number	Problem	Population size	Tournament size	Average time taken per run (seconds)
1	1	2000	2	7.9
2	1	500	2	1.47
3	1	2000	5	6.98
4	1	500	5	2.92
5	2	2000	2	3.57
6	2	500	2	0.87
7	2	2000	5	5.55
8	2	500	5	1.18

Each generation of all runs sometimes took a different amount of time; this is because the earlier generations of a run ran much quicker than the later generations, and this is because the average size of the function increases with each new generation which means more computations are required for when the next generation is processed. However, since the program calls to a third-party library, the individual generations could not be timed. Nevertheless, this can be seen by running the notebook with more generations.

2.1. Experiments 1 and 2:

Both experiments start at approximately the same fitness/size, and as new generations are created, these values start to move apart as the experiment with the higher population size is more likely to produce better solutions since it generates around 1,500 more functions. For example, at generation 30 for experiment 1, the median fitness is 15.9, while in experiment 2, it is 17.27; this shows that experiment 1 can, on average, provide a solution closer to the target than experiment 2.

Table 3: Average results from experiment 1.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	5.6028e+28	6.268e+31	35.066	15.8513	1.57277e+30	3.5856	7	3	2	1.51316
1	1578.40	3.4223e+28	4.2133e+31	31.3713	14.9563	1.05322e+30	3.8376	12.9	3	1	1.88591
2	1579.80	2.8458e+28	2.8312e+31	31.2297	14.6976	7.94865e+29	4.1022	15	3.4	1	2.15182
3	1582.10	3.3436e+28	3.9657e+31	29.0318	14.312	1.02537e+30	4.3638	16.8	4	1	2.42394
4	1580.00	3.1785e+28	3.4148e+31	27.3186	13.9804	8.94324e+29	4.6098	18.3	4	1	2.6676
5	1585.50	3.6884e+28	4.0193e+31	26.5605	13.5631	1.07078e+30	4.8895	18.9	4.2	1	2.91047
6	1572.20	2.7002e+59	5.4005e+62	26.0129	13.4544	1.20728e+61	5.1755	20.9	4.7	1	3.16983
7	1581.40	9.8973e+29	1.9585e+33	25.2833	13.3615	4.38454e+31	5.4646	22.7	4.9	1	3.35602
8	1571.50	2.2198e+28	2.0558e+31	25.1566	13.0663	5.87513e+29	5.8526	23.7	5.1	1	3.60267
9	1578.50	3.1188e+28	3.2785e+31	25.1566	11.9698	8.62736e+29	6.29	27.3	5.5	1	3.85287
10	1589.60	2.4974e+28	2.4756e+31	25.1566	11.8134	6.89104e+29	6.7306	27.2	5.9	1	4.12284
11	1570.80	1.854e+28	2.311e+31	24.4936	10.9563	5.63572e+29	7.3228	30	6.5	1	4.46078
12	1580.50	2.2704e+28	2.5039e+31	23.4658	11.2567	6.56724e+29	7.8882	32	7	1	4.7659
13	1586.90	1.2908e+28	1.6066e+31	22.448	11.0634	3.99426e+29	8.4788	32.2	7.5	1	5.13339
14	1576.40	2.5439e+28	2.5868e+31	21.5938	10.8076	7.01863e+29	9.1341	36.2	8.2	1	5.47989
15	1580.10	2.372e+28	2.9138e+31	21.1906	10.3287	7.31494e+29	9.802	37.6	8.9	1	5.81549
16	1576.80	2.0633e+28	1.8633e+31	20.9059	9.72612	5.21639e+29	10.544	43.4	9.3	1	6.19395
17	1580.70	2.7812e+28	3.7036e+31	20.6586	9.76407	9.03955e+29	11.388	43.7	10.2	1	6.6725
18	1574.00	1.4735e+28	2.0526e+31	20.4229	9.26711	4.85779e+29	12.305	51.4	11.1	1	7.12327
19	1579.70	1.7517e+28	2.1548e+31	20.0837	9.18155	5.7064e+29	13.152	53	12	1	7.57049
20	1582.50	2.6055e+28	3.7084e+31	19.6639	9.15294	9.08772e+29	14.148	54.5	12.9	1	8.10818
21	1582.20	2.5258e+28	2.4702e+31	19.2469	8.12341	7.03457e+29	15.036	59.4	13.8	1	8.6381
22	1583.60	2.3329e+28	3.7624e+31	18.9261	7.98653	8.78825e+29	15.988	61.4	14.4	1	9.19125
23	1577.90	2.4662e+28	2.6907e+31	18.6445	8.10265	6.96706e+29	17.134	65.9	15.6	1	9.80904
24	1578.60	9.6792e+27	1.4603e+31	18.2629	7.54108	3.40493e+29	18.375	67.5	16.8	1	10.4341
25	1576.70	2.4105e+28	2.3215e+31	17.905	7.16029	6.48023e+29	19.707	77.2	17.9	1	11.3097
26	1567.20	6.9302e+56	1.386e+60	17.3929	6.64659	3.09849e+58	21.115	82.7	19.3	1	12.1595
27	1585.30	2.5508e+29	4.6648e+32	17.0067	6.80535	1.05184e+31	22.51	87	20.5	1	13.0961
28	1587.70	1.7598e+28	2.2132e+31	16.6195	6.6504	5.29096e+29	23.953	97.1	21.6	1	14.0475
29	1584.60	4.1607e+28	5.4251e+31	16.2578	6.31187	1.32563e+30	25.663	104.7	23.1	1	15.0159
30	1574.90	3.584e+28	3.6776e+31	15.8971	6.12293	1.00316e+30	27.351	108	24.8	1	16.1015

Table 4: Average results from experiment 2.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	500.00	8.2218e+27	3.7006e+30	34.7164	17.5129	1.69709e+29	3.5358	7	3	2	1.4765
1	394.60	2.4036e+28	1.0787e+31	31.3713	16.0953	4.91084e+29	3.7488	11.5	3	1	1.85411
2	392.30	3.6189e+28	1.5838e+31	30.8462	15.9722	7.34874e+29	4.003	13.3	3.2	1	2.12939
3	397.90	3.151e+28	1.2294e+31	29.0981	15.9283	5.72034e+29	4.2318	15.3	3.65	1	2.38736
4	392.30	2.6446e+28	8.8037e+30	27.242	14.6307	4.31095e+29	4.4948	16.2	4	1	2.62593
5	395.80	2.0305e+28	5.8962e+30	26.3555	14.832	3.13205e+29	4.762	17.9	4.1	1	2.91518
6	392.90	1.5393e+28	4.8508e+30	25.9294	14.1973	2.49626e+29	5.1308	19.1	4.6	1	3.20893
7	401.40	3.2342e+28	1.1819e+31	25.3386	13.4537	5.94178e+29	5.4044	19	4.9	1	3.36451
8	395.20	1.677e+28	6.3332e+30	25.371	13.0993	2.98821e+29	5.9	21.7	5.2	1	3.7008
9	394.70	2.9811e+28	9.1875e+30	24.9259	13.1614	4.78531e+29	6.3094	21.6	5.5	1	3.97365
10	398.60	2.6949e+28	1.0602e+31	24.8487	13.1121	4.9523e+29	6.7482	24.1	5.7	1	4.23919
11	393.00	1.8875e+28	8.4116e+30	24.6116	13.1518	3.80512e+29	7.2	25.4	6.3	1	4.486
12	395.40	1.7956e+28	7.1913e+30	24.2004	11.9765	3.35883e+29	7.823	28.3	7	1	4.7777
13	392.30	1.2954e+28	5.4514e+30	23.5752	12.5047	2.48844e+29	8.2834	30.4	7.4	1	5.02078
14	394.50	1.2342e+28	5.2684e+30	22.7782	12.6211	2.40517e+29	8.8774	32	7.8	1	5.32052
15	400.20	2.2518e+28	1.0189e+31	22.4502	12.4186	4.66462e+29	9.6058	36.3	8.7	1	5.7223
16	394.80	2.0307e+28	8.7173e+30	21.6167	11.6606	4.07176e+29	10.255	37.7	9.4	1	5.93371
17	398.30	5.6741e+27	2.1961e+30	21.2277	11.616	1.066e+29	10.726	40.1	9.55	1	6.31151
18	398.10	2.3646e+28	9.1483e+30	21.0781	11.2134	4.45017e+29	11.386	40.8	10.2	1	6.59591
19	390.40	1.2271e+28	4.2549e+30	20.723	11.1067	2.04809e+29	12.136	43.4	10.95	1	6.97145
20	397.80	1.3143e+28	6.5116e+30	20.612	10.4267	2.91011e+29	12.819	44.7	11.85	1	7.294
21	393.00	2.3162e+28	1.0966e+31	20.13	10.2121	4.91113e+29	13.627	46.6	12.65	1	7.67487
22	398.50	3.656e+28	1.5114e+31	19.6223	10.1155	7.17818e+29	14.481	46.7	13.45	1	8.06186
23	395.30	7.5511e+27	3.3922e+30	19.2729	9.68656	1.54089e+29	15.564	49.7	14.3	1	8.62877
24	389.60	2.2214e+28	1.0542e+31	18.8537	9.54112	4.74706e+29	16.573	54	15.15	1.1	9.24417
25	392.10	1.6539e+28	6.922e+30	18.5723	9.57396	3.18303e+29	17.835	55.8	16.6	1	9.88009
26	396.70	3.2699e+28	1.2552e+31	18.278	9.10617	5.91678e+29	19.019	62.5	17.5	1	10.5606
27	397.10	2.4187e+28	8.2396e+30	17.6571	8.95841	4.14796e+29	19.998	64.4	18.3	1	11.2558
28	398.10	1.8057e+28	5.5725e+30	17.5741	9.09007	2.82008e+29	21.229	68.6	19.8	1	11.951
29	399.30	4.5661e+28	1.5798e+31	17.5629	9.06132	7.4882e+29	22.51	71	20.8	1	12.6389
30	399.80	9.1143e+29	4.5199e+32	17.2699	8.59869	2.02252e+31	23.825	74.3	21.7	1	13.3574

Figures 3 and 4 show that when using higher population size, the solutions generated follow the target function more closely and are even starting to try to match the second hill when x is 0. Interestingly, with experiment 1 there is a spike in some of the solutions when x is 0; this appears to be a singularity as the implementation of the Jupyter notebook provided contains a protected division method which returns one if the term on the right-hand side is 0, this could be why, when x is 0 that the results suddenly spike up by a small amount.

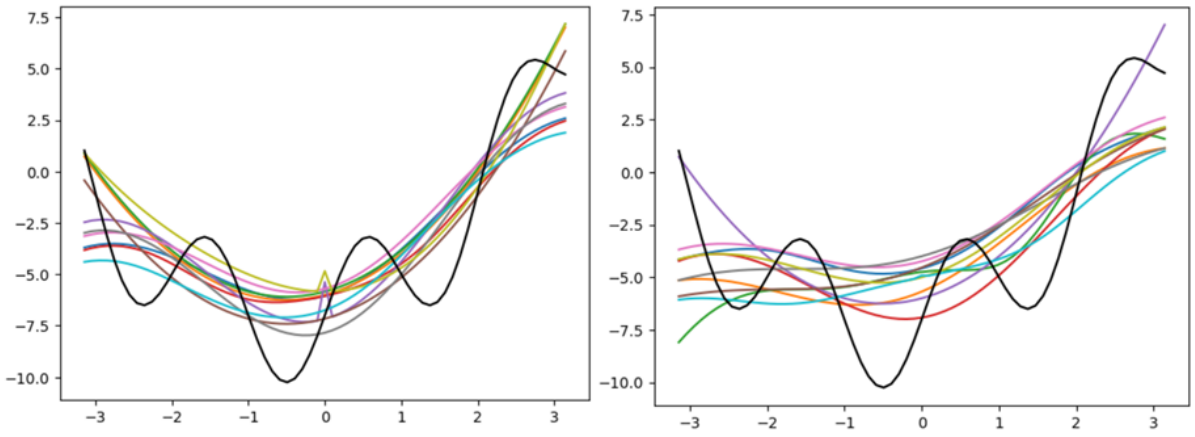

Figure 3: Best solutions, all runs vs the target function: experiment 1 (left), experiment 2 (right).

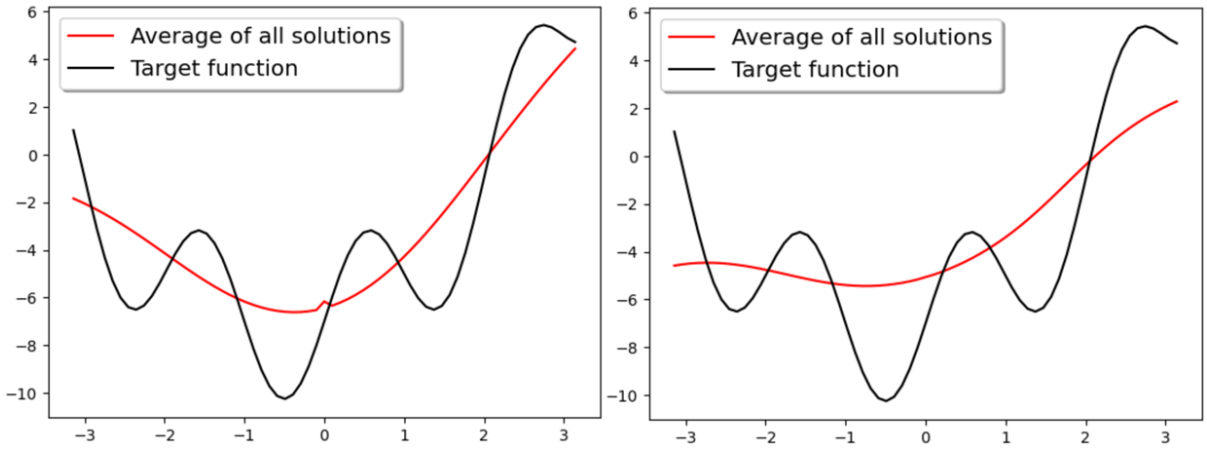

Figure 4: Average solution of all runs vs the target function: experiment 1 (left), experiment 2 (right).

2.2. Experiments 3 and 4:

Experiment 3 produces better solutions than experiment 1, which has a tournament size of 2 compared to the tournament size of 5 for experiment 3. Taking the median fitness of both at generation 30, experiment 3 is 4.12 while experiment 1 is 15.9, showing that with the same population size, a tournament size of 5 is better than 2 by a factor of almost 4.

The closer a solution is to the target function, the more complex the evolved function is because the size of the function increases to around 109 instructions on average for experiment 3, which is very large compared to the target function in problem 1.

Table 5: Average results from experiment 3.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	4.4923e+28	4.1036e+31	34.8817	15.6378	1.11836e+30	3.5872	7	3	2	1.50281
1	1574.90	2.368e+28	2.22e+31	28.3691	13.9513	6.18633e+29	3.9537	13	3	1	1.91492
2	1576.90	2.8668e+28	2.9534e+31	25.3284	13.7577	7.80869e+29	4.4503	16.8	4	1	2.21316
3	1587.10	3.1375e+28	2.2024e+31	22.4089	13.2167	7.04534e+29	5.0873	18.2	5	1	2.47129
4	1590.10	5.3232e+28	5.6976e+31	20.942	12.08	1.49182e+30	6.069	19	5.5	1	2.76676
5	1587.40	4.5653e+28	2.8474e+31	20.2259	10.9725	9.28109e+29	7.1336	21.2	7	1	3.0711
6	1588.20	2.4869e+28	2.784e+31	18.8794	10.2973	7.33438e+29	8.176	24.7	7.3	1	3.526
7	1570.20	2.8732e+28	2.444e+31	17.967	9.09843	7.30659e+29	9.2432	28	8.8	1	3.87923
8	1568.10	2.2641e+28	2.1713e+31	16.5998	8.44439	6.33659e+29	10.254	30.4	9.4	1	4.26045
9	1582.80	1.7488e+28	1.894e+31	16.1127	7.15897	5.15764e+29	11.309	34.3	10.8	1	4.76245
10	1590.40	1.4841e+28	1.4983e+31	15.7868	6.41406	4.26237e+29	12.583	40.1	11.9	1	5.41954
11	1569.60	3.1131e+28	3.0407e+31	15.1205	6.04765	8.23077e+29	14.289	44	13.7	1	6.03655
12	1583.50	4.1726e+28	4.662e+31	14.2571	5.27164	1.20248e+30	16.282	49.6	15.5	1	6.94993
13	1572.70	3.3782e+28	3.8144e+31	13.406	4.99474	9.70542e+29	18.354	53.5	17.7	1	7.77349
14	1595.90	2.5964e+59	5.1928e+62	12.5475	4.83166	1.16086e+61	20.594	62.3	19.7	1	8.65096
15	1579.20	2.5964e+59	5.1928e+62	11.3728	4.5928	1.16086e+61	22.996	64.6	22.3	1	9.37238
16	1588.90	3.431e+28	2.8268e+31	10.5351	4.35988	7.8975e+29	25.648	71.4	24.8	1	10.6039
17	1576.40	9.0485e+58	1.8097e+62	9.51573	4.16591	4.04559e+60	28.559	81.7	27.5	1	11.5644
18	1594.50	1.0944e+59	2.1887e+62	8.95153	4.02379	4.89294e+60	31.199	90.4	30	1	12.5449
19	1578.00	1.3217e+59	2.4002e+62	8.46132	3.85015	5.39291e+60	33.908	98.4	32.3	1	13.8631
20	1577.20	5.0641e+28	3.5324e+31	7.84871	3.59762	1.01973e+30	36.8	115.1	34.9	1	15.549
21	1586.50	8.687e+28	5.3353e+31	7.34223	3.47276	1.60832e+30	39.987	122.8	38.3	1	17.19
22	1576.50	2.8908e+58	4.6253e+61	6.87933	3.21198	1.06568e+60	44.04	138.1	42	1	19.3383
23	1571.40	4.726e+28	4.2121e+31	6.29814	3.13292	1.0329e+30	48.93	157.3	46.4	1	21.9077
24	1580.80	4.1483e+28	2.2232e+31	5.89909	3.13441	6.55106e+29	54.208	187.1	50.4	1	25.0258
25	1592.70	2.135e+58	4.2699e+61	5.44378	2.89727	9.5455e+59	60.751	200.9	55.9	1	28.5137
26	1585.70	3.369e+28	2.0419e+31	5.07555	2.83381	5.9211e+29	68.675	230.1	63.5	1	32.7016
27	1581.50	3.3494e+28	3.6992e+31	4.77058	2.60669	9.45225e+29	77.29	257.3	71.5	1	36.3236
28	1582.20	2.1567e+28	2.32e+31	4.56147	2.58213	5.80503e+29	89.363	269.3	84.4	1	41.0937
29	1575.40	4.5661e+28	5.1523e+31	4.3321	2.28566	1.29034e+30	102.43	321.7	100.65	1.3	43.6439
30	1574.60	3.7946e+59	7.5892e+62	4.11899	2.18729	1.69658e+61	109.64	323.3	105.8	1	45.4362

Table 6: Average results from experiment 4.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	500.00	2.8505e+28	9.0471e+30	34.6701	16.7737	4.552e+29	3.5714	7	3	2	1.48536
1	392.80	1.8872e+28	5.9752e+30	28.9618	15.4381	3.17454e+29	3.9048	12.1	3	1	1.87088
2	393.10	1.7896e+28	7.777e+30	25.5629	15.357	3.56156e+29	4.2252	13.3	3.2	1	2.11652
3	397.00	8.4282e+27	3.1883e+30	23.4211	13.6754	1.58355e+29	4.899	15.6	4.6	1	2.4832
4	390.70	6.3452e+28	3.0162e+31	21.5653	12.979	1.35437e+30	5.7986	18.2	5.15	1	2.79596
5	395.60	2.9643e+28	9.8974e+30	20.4707	12.7419	5.1028e+29	6.8768	20.9	6.4	1	3.2003
6	396.00	4.9078e+28	1.7854e+31	19.312	11.9349	8.88061e+29	7.9594	22.8	7.3	1	3.63347
7	396.90	5.6883e+28	2.1876e+31	18.6634	10.3835	1.0842e+30	9.0374	25.1	8.5	1	4.03341
8	386.90	4.8256e+28	1.1348e+31	17.2871	10.2443	6.88786e+29	10.288	26.9	10	1	4.30633
9	389.70	2.0949e+28	8.0472e+30	16.5419	9.56025	3.83201e+29	11.322	30.1	11	1	4.75274
10	394.60	2.4321e+28	7.3323e+30	15.9283	8.90791	3.79612e+29	12.561	32.8	11.9	1	5.3208
11	395.40	8.5633e+27	4.0765e+30	15.3176	7.95919	1.83676e+29	14.235	39.9	13.65	1	6.2097
12	393.50	1.751e+29	8.5396e+31	14.2704	6.96444	3.82288e+30	15.943	44.7	15.25	1	6.99779
13	391.00	1.058e+28	4.2641e+30	13.4334	6.81643	1.99138e+29	18.054	53.1	17.3	1	7.96992
14	403.20	1.2962e+28	5.9917e+30	12.6461	6.13014	2.6908e+29	20.057	56.7	19.05	1	8.8702
15	396.90	1.9897e+28	7.9554e+30	11.6496	5.70702	3.80357e+29	22.428	59.6	21.55	1.2	9.846
16	394.00	2.2815e+28	6.5792e+30	10.8577	5.43048	3.48632e+29	25.274	69.7	24.3	1	10.7287
17	395.00	2.4919e+28	7.975e+30	10.1246	5.15805	4.06947e+29	27.665	72.3	27.05	1.1	11.3768
18	395.90	2.3045e+28	9.5547e+30	9.38731	4.88471	4.47375e+29	30.447	75.8	29.95	1.4	11.9796
19	395.90	2.6136e+28	1.0435e+31	8.77962	4.59175	4.97449e+29	32.941	86.3	32.2	1.2	13.1812
20	394.60	1.4877e+28	5.015e+30	8.09787	4.36804	2.51025e+29	36.169	89.7	35.25	1.1	14.4708
21	401.90	3.4635e+28	6.5384e+30	7.63413	4.2134	4.41457e+29	39.123	95.9	37.85	1.2	15.5563
22	400.90	1.5602e+28	5.6566e+30	7.17945	4.05183	2.77318e+29	43.059	110.7	41.6	1.6	17.4037
23	399.90	5.0073e+28	1.0006e+31	6.71925	3.9104	5.79442e+29	47.774	118.7	46.8	1.4	18.9966
24	395.10	2.7486e+28	1.0256e+31	6.35971	3.65109	5.1903e+29	52.7	140.9	50.9	1.4	21.2364
25	399.80	3.2101e+28	8.0435e+30	6.0048	3.42574	4.71144e+29	58.47	143	56.8	1.2	23.0157
26	395.80	2.4566e+28	4.9838e+30	5.63965	3.19909	3.21605e+29	64.534	159.7	62.9	1.1	25.0318
27	394.90	3.4024e+27	1.6413e+30	5.44085	3.10652	7.34194e+28	69.832	178.3	69.6	1.8	26.1072
28	397.90	6.5584e+28	1.9689e+31	5.17775	2.84698	1.02951e+30	75.51	191.1	73.95	2.1	28.6293
29	401.30	2.118e+28	7.0842e+30	4.89622	2.74167	3.69149e+29	81.193	207.1	78.9	3.2	30.8721
30	394.90	1.8757e+28	6.8402e+30	4.72542	2.67769	3.37782e+29	88.734	217.2	86.7	2.3	32.4286

From observing figure 5, experiment 3 is more uniform than experiment 4, where the solutions are typically further away from the target function. However, both experiments follow the function better than experiments 1 and 2, respectively, which indicates that changing the tournament size has a better impact than the population size. Similarly, compared to experiment 1, experiment 3 contains spikes where the hills/ valleys are, which results in a better fitness score; this can also be seen in figure 6, where the lines are not smooth.

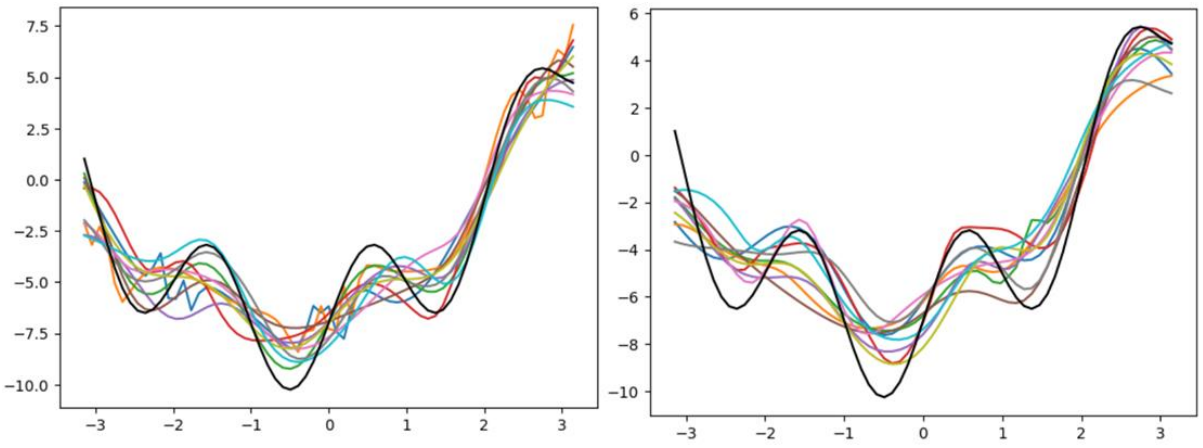

Figure 5: Best solutions, all runs vs the target function: experiment 3 (left), experiment 4 (right).

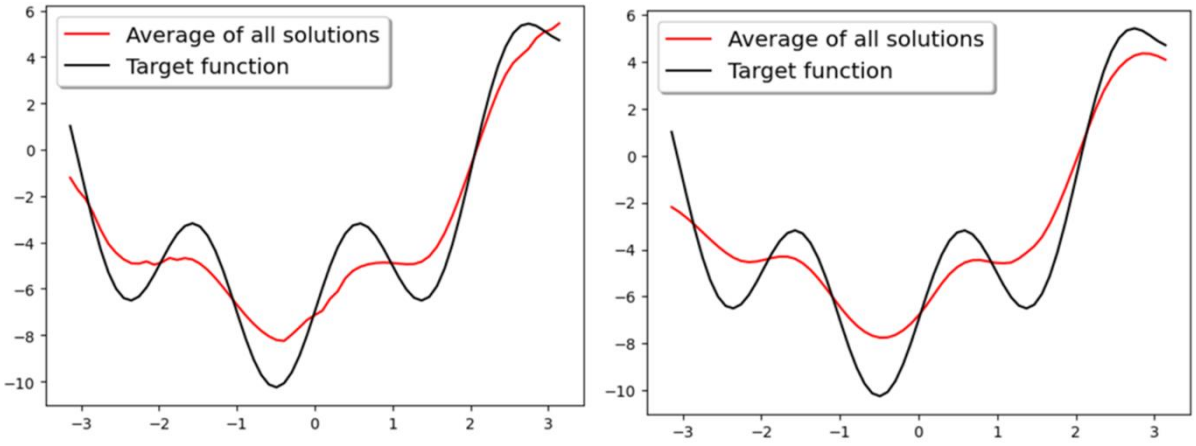

Figure 6: Average solution of all runs vs the target function: experiment 3 (left), experiment 4 (right).

2.3. Experiments 5 and 6:

Compared to problem 1, problem 2 is a more complex function to emulate, and this can be seen by looking at the average minimum and median fitness of experiments 1-4 compared to experiments 5-8. For example, Table 7 contains the median and minimum of the fitness scores at generation 30 in experiments 1 and 5, and since experiments 1 and 5 both have the same population and tournament size, this shows that problem 2 is more complicated as the fitness scores are lower for experiment 1.

Table 7: Average fitness scores of experiments 1 and 5.

Statistic	Experiment 1	Experiment 5
Median fitness	15.9	25.65
Minimum fitness	6.12	19.12

Population size does not significantly impact the fitness of problem 2, and this can be seen in the final generation of both experiments, where the median and minimum fitness is relatively similar. However, slightly better solutions are generated by the experiment with the higher population size, with the best solution found at generation 28 of experiment 5 with a minimum fitness of 19.072 compared to the best in experiment 6, 21.076.

Table 8: Average results from experiment 5.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	3.9437e+28	3.0301e+31	28.6835	24.1088	8.97498e+29	3.5671	7	3	2	1.50338
1	1576.20	1.8115e+28	1.7199e+31	27.9331	23.9743	4.92743e+29	3.703	12.8	3	1	1.8366
2	1571.50	2.5516e+28	2.148e+31	27.3399	23.212	6.08024e+29	3.8797	15.8	3	1	2.12131
3	1588.00	2.5058e+28	2.6475e+31	27.1629	23.1728	7.17633e+29	4.0705	15.8	3.5	1	2.35147
4	1580.80	2.3907e+28	3.075e+31	27.1324	23.1972	7.66626e+29	4.2245	17.4	3.9	1	2.55462
5	1583.80	1.9803e+28	1.7009e+31	27.1324	22.7704	5.01393e+29	4.3203	18.3	3.9	1	2.73949
6	1571.40	2.9729e+28	3.3774e+31	27.1215	22.8028	8.65466e+29	4.3911	19.2	3.9	1	2.92338
7	1577.80	1.7069e+28	1.936e+31	27.0509	22.7597	5.15511e+29	4.4484	21.7	3.8	1	3.16325
8	1594.30	1.3298e+28	1.4215e+31	26.7342	22.5094	3.94523e+29	4.4887	22.4	3.8	1	3.34107
9	1572.80	1.4304e+28	1.4486e+31	26.4001	22.5225	3.88647e+29	4.4772	23.7	3.7	1	3.50584
10	1582.50	1.5649e+28	1.792e+31	26.2993	22.3621	4.66497e+29	4.5194	25.2	3.4	1	3.68254
11	1584.30	2.4676e+28	3.7681e+31	26.1733	22.2357	8.79744e+29	4.5845	25	3.4	1	3.84166
12	1582.20	1.3876e+28	1.3612e+31	26.1199	22.3239	3.77835e+29	4.6039	24.7	3.4	1	3.97925
13	1574.30	1.8023e+28	2.3453e+31	26.0906	22.1143	5.80687e+29	4.7892	26	3.45	1	4.24868
14	1571.20	2.2068e+28	2.7992e+31	26.0814	21.8167	6.82491e+29	5.0035	27.1	3.7	1	4.53079
15	1574.20	1.7762e+28	1.7332e+31	26.078	21.9796	4.95667e+29	5.3048	31.3	3.7	1	4.82948
16	1589.30	2.0746e+28	3.3585e+31	26.081	21.397	7.71001e+29	5.7196	35.5	4.3	1	5.21997
17	1578.50	2.3395e+29	4.5684e+32	26.0782	20.8583	1.02615e+31	6.237	36.3	4.5	1	5.62823
18	1574.10	1.4745e+28	1.5647e+31	26.0856	20.7722	4.319e+29	6.7826	36.7	5.2	1	5.97114
19	1583.20	2.1401e+28	1.792e+31	26.0846	20.3777	5.44533e+29	7.4452	40.4	5.6	1	6.45055
20	1583.10	2.1975e+28	2.2149e+31	26.0852	20.8636	6.01858e+29	8.1104	41.8	6.4	1	6.86886
21	1573.10	5.6599e+28	9.3365e+31	26.1228	20.6196	2.1795e+30	8.8211	45	7	1	7.32214
22	1596.60	2.8581e+28	2.8867e+31	26.1057	20.8283	7.56378e+29	9.5841	48.6	7.55	1	7.83949
23	1578.90	1.6612e+28	1.4724e+31	26.1139	20.2782	4.20317e+29	10.329	51.8	8.3	1	8.32833
24	1567.70	4.738e+57	9.476e+60	26.0984	20.0844	2.11837e+59	11.256	55.5	9.3	1	8.96275
25	1579.80	2.2245e+28	1.9906e+31	26.0731	20.1815	5.97457e+29	12.326	62.2	10.1	1	9.6283
26	1592.00	2.4489e+28	2.7612e+31	26.013	19.9981	7.43911e+29	13.469	69.7	11.2	1	10.4103
27	1565.20	3.1129e+28	4.7533e+31	25.9788	19.4791	1.13508e+30	14.735	76	12.3	1	11.1605
28	1580.70	4.5657e+28	6.3208e+31	25.8446	19.072	1.53487e+30	16.095	84.2	13.4	1	12.0582
29	1579.00	2.7002e+59	5.4005e+62	25.7641	19.2379	1.20728e+61	17.679	84	14.8	1	13.0474
30	1581.40	1.352e+29	2.2159e+32	25.6491	19.1225	5.14063e+30	19.221	91.1	16.2	1	14.0104

Table 9: Average results from experiment 6.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	500.00	5.2651e+28	1.8593e+31	28.9786	24.516	8.7794e+29	3.5722	7	3	2	1.527
1	392.80	4.1403e+28	1.2521e+31	28.2385	24.1266	6.93664e+29	3.6844	12	3	1	1.8489
2	395.20	6.2192e+28	2.7297e+31	27.5149	23.3239	1.24495e+30	3.8858	13.2	3.15	1	2.10654
3	395.80	3.8007e+27	1.5499e+30	27.2918	23.3602	7.24094e+28	4.1066	13.9	3.8	1	2.3413
4	396.30	1.6441e+28	8.1606e+30	27.1324	23.4518	3.6468e+29	4.255	14.7	3.7	1	2.5674
5	398.20	1.1935e+28	5.3329e+30	27.1121	23.5971	2.4376e+29	4.3034	15.3	3.8	1	2.72275
6	393.20	1.8748e+28	6.2721e+30	27.0161	23.8329	3.23149e+29	4.3208	17.2	3.65	1	2.88653
7	395.30	2.482e+28	1.0508e+31	26.7962	23.5359	4.84676e+29	4.3758	17.2	3.8	1	2.99783
8	398.30	2.2136e+28	9.2699e+30	26.6215	23.3651	4.32578e+29	4.4046	18.4	3.7	1	3.13668
9	397.20	1.8003e+28	5.9841e+30	26.5504	22.9798	3.18313e+29	4.4528	19.7	3.7	1	3.32791
10	393.30	1.0256e+28	4.5127e+30	26.3979	23.1742	2.07262e+29	4.5804	19.3	3.6	1	3.49757
11	396.70	3.1375e+28	1.3816e+31	26.3225	23.4019	6.31178e+29	4.7248	20.9	3.8	1	3.75721
12	397.60	5.454e+27	1.0857e+30	26.1819	23.4456	7.52175e+28	4.7994	23.1	3.8	1	3.98241
13	395.30	9.6331e+27	4.6053e+30	26.1747	23.5445	2.09537e+29	4.9176	22.7	3.7	1	4.21551
14	399.80	1.6138e+28	6.5728e+30	26.1039	23.2304	3.01898e+29	5.0622	24.6	4	1	4.33261
15	391.50	2.2387e+28	1.1194e+31	26.15	23.0528	5.00091e+29	5.329	24.7	4	1	4.53315
16	400.30	3.1389e+28	1.2995e+31	26.1083	22.8458	6.07595e+29	5.751	28	4.3	1	4.98647
17	396.40	1.0272e+28	3.4877e+30	26.1537	22.9102	1.82532e+29	6.156	28.5	4.7	1	5.29716
18	391.00	1.4357e+28	6.1357e+30	26.1284	23.0211	2.86801e+29	6.5572	31.2	4.9	1	5.66255
19	393.70	1.3608e+28	5.9835e+30	26.1553	22.8702	2.744e+29	7.0518	31.9	5.5	1	5.97079
20	390.30	1.6398e+28	6.5116e+30	26.0918	22.683	3.01608e+29	7.525	36	5.9	1	6.24435
21	389.40	1.1436e+28	4.2472e+30	26.085	23.039	2.01116e+29	8.239	39.7	6.45	1	6.87141
22	403.70	2.3503e+28	6.9564e+30	26.1061	22.2881	3.77134e+29	9.0006	40	7.3	1	7.24433
23	399.30	1.648e+28	6.5116e+30	26.122	22.3217	3.11476e+29	9.8056	40.6	8.1	1	7.74438
24	391.90	7.295e+28	3.4712e+31	26.1236	21.6421	1.56125e+30	10.607	45.8	8.6	1	8.3405
25	402.30	1.8157e+28	6.4491e+30	26.0761	21.6595	3.11559e+29	11.443	50.6	9.5	1	9.11356
26	395.50	4.1845e+28	1.1819e+31	26.0621	21.9143	6.55074e+29	12.517	52.6	10.1	1	9.67484
27	396.10	3.7451e+60	1.8726e+63	25.9555	21.076	8.36599e+61	13.564	57.9	11.4	1	10.2606
28	395.40	5.5228e+27	1.7012e+30	25.9742	21.3921	9.22021e+28	14.312	58.4	11.8	1	10.7836
29	397.40	2.2973e+28	9.8487e+30	25.9008	21.2231	4.48863e+29	15.435	61.5	12.95	1	11.3094
30	398.20	3.7225e+28	1.8217e+31	25.8327	21.167	8.14281e+29	16.685	64.9	14.2	1	11.8728

Figure 7 shows that most resultant functions are having trouble with the two big hills at x=-2 and x=2. However, some of the solutions show that it is possible, for example, the purple line in the left plot and the orange in the right plot; more spikes are being explored in experiment 5, which is why, typically, the results from this experiment are better as they result in being closer to the target function. Figure 8 shows that the centre part of the function is similar in both experiments; however, the start and the end differ. For example, in the plot for experiment 5, the line starts and ends at around y=-7.5, whereas, with experiment 6, this is between -2.5 and -5.

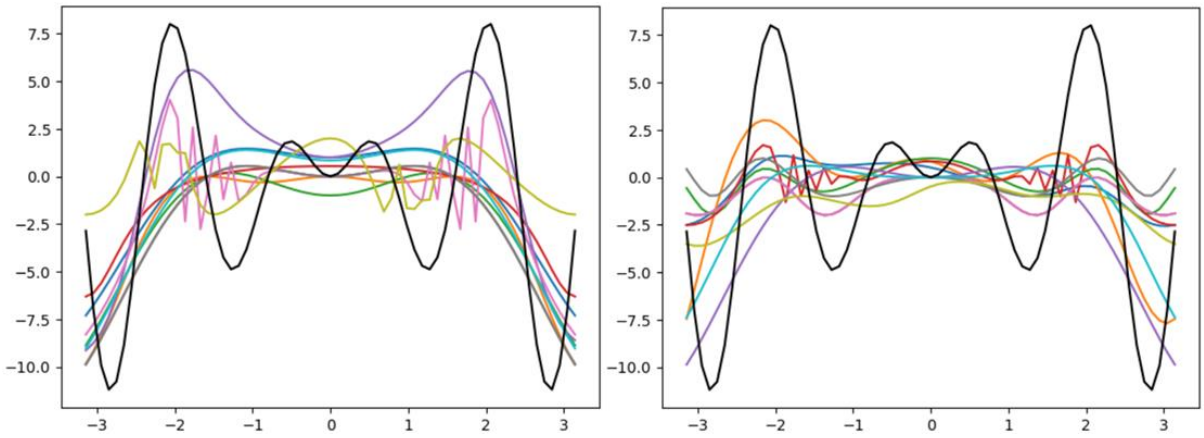

Figure 7: Best solutions, all runs vs the target function: experiment 5 (left), experiment 6 (right).

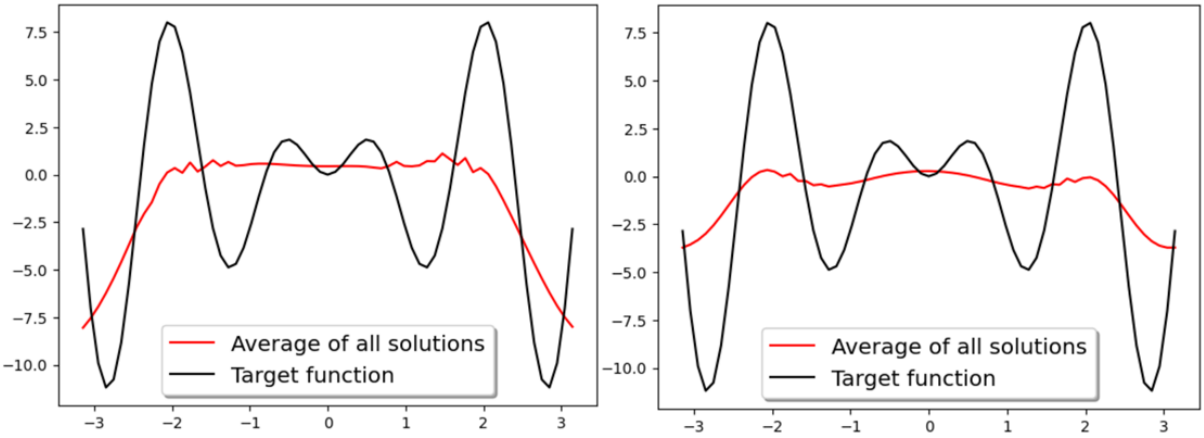

Figure 8: Average solution of all runs vs the target function: experiment 5 (left), experiment 6 (right).

2.4. Experiments 7 and 8:

Similar observations from comparing experiments 1 and 3 for problem 1 can also be seen with experiments 5 and 7, whereby the experiment with the larger tournament size generates better solutions. For instance, the median fitness at generation 30 for experiments 5 and 7 is 25.65 and 10.21, respectively; this means that, on average, for problem 2, a tournament size of 5 is a factor of 2.5 times better than a GP with a tournament size of 2, however, this appears to have a more significant impact when applied to problem 1.

Table 10: Average results from experiment 7.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	4.6713e+28	2.9078e+31	28.842	24.0911	9.43011e+29	3.5866	7	3	2	1.50891
1	1586.60	3.5605e+28	4.1259e+31	27.1324	23.6007	1.03875e+30	3.6585	13.2	3	1	1.84397
2	1578.30	2.2568e+28	1.7436e+31	27.1215	23.2807	5.33991e+29	3.6791	15.3	3	1	2.02786
3	1579.60	2.6851e+28	2.3923e+31	26.8509	22.548	6.33354e+29	3.7782	15.8	3.1	1	2.20905
4	1585.40	6.8398e+57	1.368e+61	26.1815	22.2827	3.05809e+59	3.9433	17.4	3.4	1	2.4044
5	1591.30	1.552e+28	1.7207e+31	26.0942	21.6784	4.45786e+29	4.3244	19.4	3.8	1	2.65731
6	1575.20	3.8567e+28	4.8736e+31	25.9756	19.8121	1.2217e+30	4.964	20.1	4.35	1	2.93166
7	1584.50	5.2308e+28	3.7209e+31	26.2633	19.5432	1.18189e+30	5.8225	23.2	5.2	1	3.2488
8	1569.70	8.7584e+28	2.89e+31	26.2096	18.4479	1.29055e+30	6.6508	24.1	6.1	1	3.54664
9	1582.00	6.6626e+59	1.3325e+63	26.1059	18.1723	2.97885e+61	7.5116	26.9	6.9	1	3.92655
10	1572.10	1.0801e+60	1.0801e+63	25.8542	16.8538	3.29731e+61	8.5339	31.3	7.9	1	4.38306
11	1581.90	4.368e+60	7.6558e+63	25.7473	16.3702	1.77122e+62	9.6037	32.6	8.8	1	4.80694
12	1593.80	7.2018e+91	1.4404e+95	25.6933	14.5687	3.21994e+93	10.797	35.5	10	1	5.27375
13	1576.00	6.3721e+59	7.3436e+62	25.3888	13.2724	2.03731e+61	12.045	41.5	11.1	1	5.78536
14	1581.00	6.6626e+59	1.3325e+63	25.0112	13.085	2.97885e+61	13.543	42.1	12.4	1	6.52378
15	1576.60	6.6626e+59	1.3325e+63	24.6844	12.3546	2.97885e+61	15.451	50.3	14.5	1	7.33859
16	1575.20	2.9767e+59	3.89e+62	24.2544	12.8668	9.84176e+60	17.567	54.3	17	1	8.08336
17	1581.10	2.9188e+91	5.8376e+94	23.5729	11.6145	1.30499e+93	19.744	61.4	19.25	1	8.67652
18	1581.30	1.6667e+29	2.6565e+32	22.37	10.4635	6.15064e+30	21.921	69.2	21.4	1	9.5002
19	1586.70	2.8807e+92	5.7614e+95	21.6184	9.86328	1.28798e+94	23.887	74.2	22.9	1	10.2233
20	1569.60	1.7285e+58	3.4569e+61	20.985	8.84004	7.72804e+59	25.697	75.9	24.6	1	10.9723
21	1576.70	5.4463e+28	4.859e+31	20.0309	8.50022	1.36517e+30	27.481	81	26.3	1	11.7843
22	1572.90	3.1438e+59	6.2875e+62	18.9542	7.92147	1.40559e+61	29.582	92.6	27.95	1	12.8598
23	1578.40	1.476e+29	1.6957e+32	17.5792	7.67081	4.35129e+30	32.013	101.6	30.5	1	14.0176
24	1568.50	6.6626e+59	1.3325e+63	16.0422	6.98668	2.97885e+61	35.492	104.9	34.4	1	15.398
25	1583.30	2.3217e+60	4.6434e+63	14.6935	6.60885	1.03803e+62	39.515	113.1	38.5	1	16.7757
26	1568.40	1.2984e+29	1.7352e+32	13.7234	6.47284	4.12166e+30	43.224	133.3	41.8	1	18.0696
27	1576.60	3.0027e+29	5.6839e+32	12.7444	6.22868	1.28257e+31	46.483	154	44.15	1	20.0336
28	1583.20	1.7774e+29	3.1886e+32	11.4547	6.13707	7.19218e+30	49.328	161.3	46.2	1	21.878
29	1579.70	9.406e+58	1.8812e+62	10.8475	5.76656	4.20543e+60	51.568	171.3	47.8	1	23.7167
30	1572.50	2.7002e+59	5.4005e+62	10.2137	5.48179	1.20728e+61	53.914	192.6	49.6	1	25.2946

Table 11: Average results from experiment 8.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	500.00	9.5425e+28	4.3239e+31	28.7237	24.6611	1.97433e+30	3.5476	7	3	2	1.48801
1	398.80	2.3493e+28	8.1737e+30	27.1816	23.5452	4.08862e+29	3.6254	12	3	1	1.83064
2	391.30	2.2157e+28	7.3858e+30	27.0998	23.0658	3.88361e+29	3.7214	13.1	3	1	2.06643
3	390.10	1.9995e+28	5.5126e+30	26.6007	22.8111	3.03588e+29	3.9078	15	3.4	1	2.3754
4	395.30	1.7785e+28	5.0454e+30	26.2812	22.788	2.71433e+29	4.2272	15.6	3.7	1	2.50526
5	402.00	5.918e+28	1.7167e+31	26.2755	22.5152	8.61776e+29	4.792	16.3	4.3	1	2.69445
6	396.10	1.099e+29	2.6383e+31	26.2916	22.3535	1.42035e+30	5.5012	17.3	5.05	1	3.03645
7	392.40	8.0513e+28	1.6925e+31	26.1046	22.1172	9.42589e+29	6.2022	19	5.6	1	3.19753
8	391.30	6.8525e+28	1.3175e+31	26.1174	21.0423	8.49426e+29	6.7616	20.7	6.2	1	3.36103
9	397.70	6.178e+28	1.5234e+31	26.0823	19.7347	8.53416e+29	7.5184	22.7	6.95	1	3.68966
10	394.40	9.7657e+28	1.1048e+31	25.839	19.3551	9.03139e+29	8.318	24.6	7.8	1	3.93907
11	399.20	5.8458e+60	1.3028e+63	25.9949	18.833	8.3419e+61	9.2164	25.7	8.7	1	4.2681
12	396.00	3.8156e+60	1.1451e+63	25.7804	17.7833	6.5225e+61	10.187	28.3	9.7	1	4.62271
13	398.00	1.4687e+60	7.3436e+62	25.5878	17.1756	3.28089e+61	10.875	30.9	10.35	1	4.83706
14	393.70	1.0801e+60	5.4005e+62	25.5876	17.3024	2.41276e+61	11.77	32	11.15	1.1	5.17157
15	387.20	7.4522e+60	2.2963e+63	25.1315	16.6869	1.22949e+62	12.897	35.5	12.1	1	5.64541
16	396.30	2.1602e+60	5.4005e+62	25.0737	16.4975	3.40873e+61	14.214	37.6	13.5	1	6.19442
17	400.70	1.0801e+60	5.4005e+62	24.9363	15.8547	2.41276e+61	15.393	41.6	14.6	1	6.73618
18	393.10	1.0801e+60	5.4005e+62	24.5873	15.3968	2.41276e+61	16.696	46.1	15.9	1.1	7.30599
19	393.90	1.0801e+60	5.4005e+62	24.2487	14.9889	2.41276e+61	17.909	50.5	17.1	1.2	7.76931
20	391.30	7.8835e+28	1.7077e+31	24.2329	15.0962	1.01318e+30	19.039	50.1	18.1	1.3	7.86316
21	391.80	3.2403e+60	1.0801e+63	23.3207	14.7754	5.82149e+61	20.271	51.8	19.6	1	8.13645
22	388.70	6.5964e+28	1.801e+31	23.0129	14.408	9.75176e+29	21.537	53	20.85	1.2	8.48728
23	401.10	1.5932e+29	4.8479e+31	23.4148	14.0618	2.51392e+30	22.769	57	22.3	1.1	8.9696
24	392.40	8.7076e+28	1.8675e+31	22.4072	13.8281	1.09366e+30	24.376	65	23.85	1	9.79232
25	400.60	1.0801e+60	5.4005e+62	21.2671	13.3504	2.41276e+61	26.333	70.2	25.5	1.7	10.7247
26	399.90	2.8807e+92	1.4404e+95	20.4324	13.1004	6.43504e+93	28.795	71.5	28.15	1	11.4881
27	397.40	2.1602e+60	5.4005e+62	19.8759	12.9911	3.40873e+61	31.187	80.6	30.75	1.2	12.2264
28	400.10	1.0801e+60	5.4005e+62	19.0636	12.8276	2.41276e+61	33.7	92.1	33.2	1.9	13.2089
29	400.20	1.0801e+60	5.4005e+62	19.377	12.7764	2.41276e+61	36.461	96.8	35.1	1.2	14.6428
30	395.50	3.5004e+28	9.69e+30	18.562	12.3375	5.14955e+29	38.805	104.6	36.45	1.3	16.1442

The results from experiment 8 are generally very erratic and jump up and down rapidly, whereas the runs from experiment 7 contain smoother curves; this can be seen in figures 9 and 10. This could be due to the resulting functions from experiment 8 applying arithmetic operations to simulate a curve rather than working out what curve function fits best, which makes sense as the GP is given no initial details of the target function and is simply trying to get as close to it as it can.

Another reflection upon looking at figure 9, experiment 7, is that about half of the solutions' minimum y values are lower than the target function, which makes the generated graph use a different y-axis since the solutions from experiment 7 have a bigger amplitude compared to experiment 8.

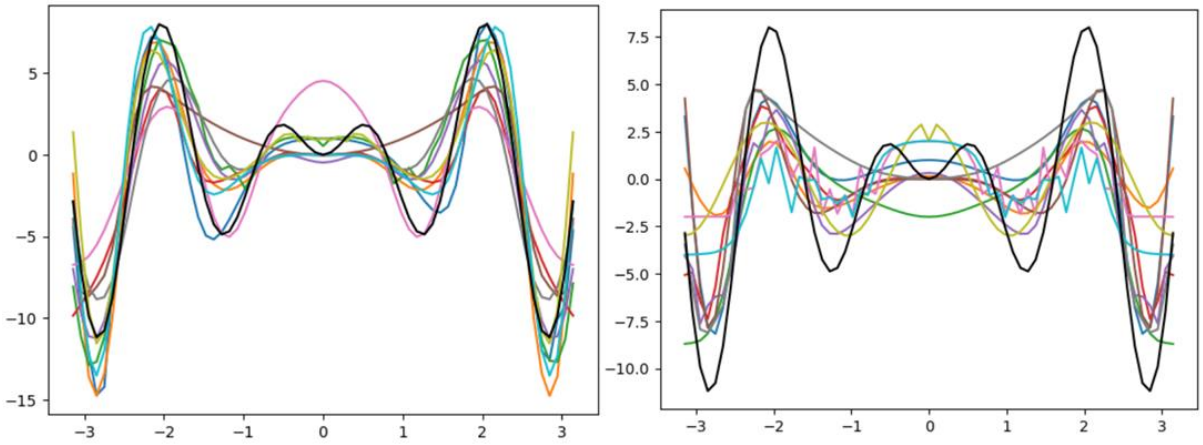

Figure 9: Best solutions, all runs vs the target function: experiment 7 (left), experiment 8 (right).

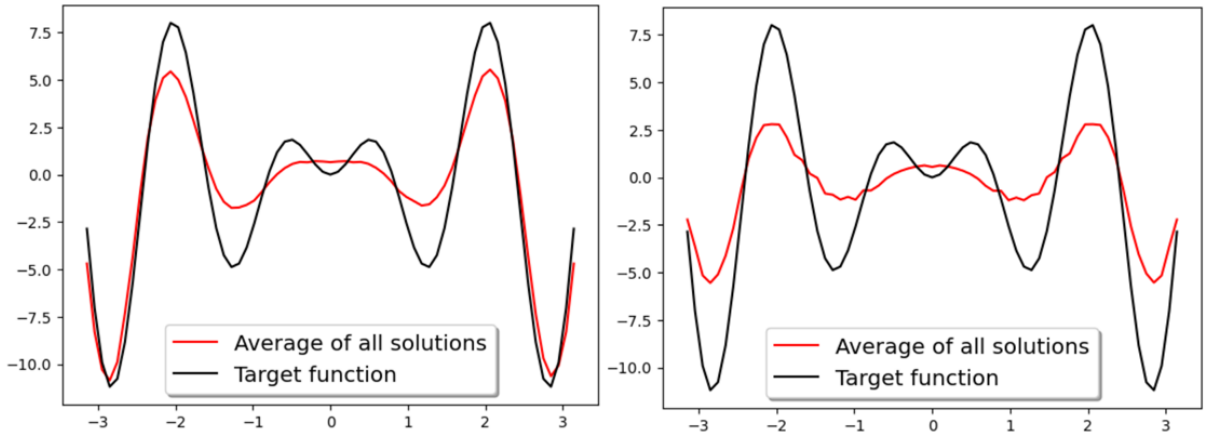

Figure 10: Average solution of all runs vs the target function: experiment 7 (left), experiment 8 (right).

Run six from experiment 8 showed an interesting similarity to the target function. The evolved function is relatively small compared to the other runs of this experiment, and although it is small, the result is close to the target function, which can be seen in Figure 11. The solution follows the pattern of a sine wave, so it has correctly identified that the target function contains a sin function and can follow the function pattern because, as x approaches 0, y is closer to 0, however as x moves away from 0, a typical sine wave pattern is produced, the further away from 0 the amplitude and frequency is increased exponentially due to x being squared. The evolved function from run six can be seen below, and a derived version of this equation produces an interesting-looking wave, see figure 12.

$$protectedDiv(sin(mul(x,x)), protectedDiv(neg(1), mul(x,x)))$$

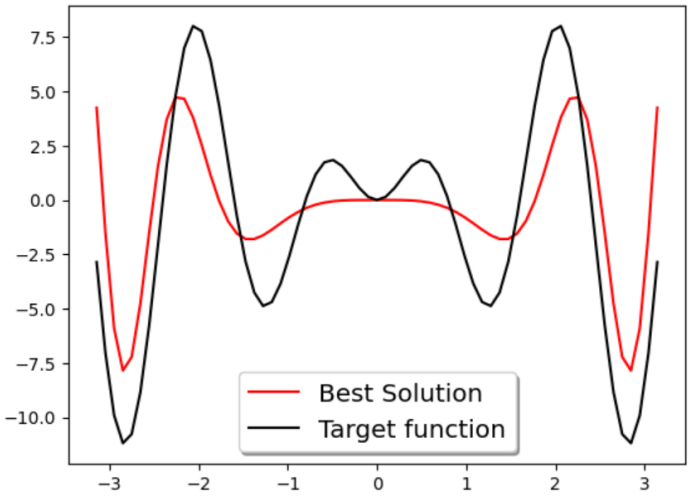

Figure 11: Best solution vs the target function from run 6 of experiment.

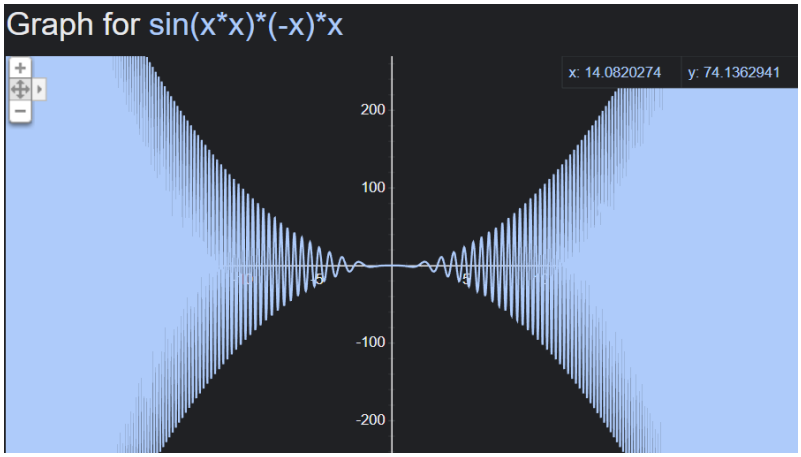

*Figure 12: Graph of sin(x*x)*-x*x*

3. Selection of Best Parameter Configuration

For the best parameter configuration, time taken is not considered, firstly because this section will only analyse the best solutions, not the speed, and secondly because the GP already runs quickly. The size is only usually considered if the time taken or the storage in computer memory is essential, which in this case, neither is.

The “best solutions” are defined as the functions closest to the target function; this is evaluated using the experiment’s lowest median fitness (in any generation). This statistic is chosen instead of the minimum fitness because the minimum could be an anomaly that can skew the results, whereas the median provides a better overall picture of all ten runs. In addition, the results from the experiments contain massive outliers, which makes using the average fitness not viable since it is sensitive to outlier values.

Table 12 shows that the best parameter configurations are experiments 3 and 7 for problems 1 and 2, respectively.

Table 12: Lowest median fitness of each experiment (any generation).

Experiment	Problem	Lowest median fitness
1	1	15.8971
2	1	17.2699
3	1	4.11899
4	1	4.72542
5	2	25.6491
6	2	25.8327
7	2	10.2137
8	2	18.562

3.1. Problem 1 further experiment

Experiment 3 has the population size set to 2,000, tournament size set to 5, and the target function is problem 1; this configuration has been rerun ten more times, and this section will discuss the difference between them and experiment 3. Table 13 displays the full results from the extra runs.

Table 13: Average results from the extra runs for problem 1.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	3.6434e+28	2.379e+31	34.8585	14.9654	7.64568e+29	3.5755	7	3	2	1.50964
1	1565.90	3.1812e+28	3.7144e+31	28.2232	14.342	9.59173e+29	3.9256	13.2	3	1	1.89897
2	1575.50	4.0169e+28	3.1958e+31	25.1666	13.2713	9.44082e+29	4.4474	14.2	4	1	2.18325
3	1580.20	3.4715e+28	2.7224e+31	22.4904	13.1359	8.35769e+29	5.1748	16.9	5	1	2.46432
4	1573.20	3.4631e+28	2.1095e+31	20.9143	12.0062	7.62826e+29	6.1414	20	5.5	1	2.83466
5	1571.10	3.5964e+28	2.9051e+31	20.1597	10.7681	8.72596e+29	7.2626	22.9	6.9	1	3.16439
6	1573.70	3.9651e+28	3.6491e+31	18.9003	10.1118	1.03028e+30	8.4272	26.3	8	1	3.57404
7	1581.90	2.0488e+28	2.0576e+31	18.1005	9.41876	5.77126e+29	9.5884	28	9.2	1	4.01529
8	1583.30	5.9896e+28	3.2167e+31	16.8172	8.08537	1.10325e+30	10.767	33.6	10.2	1	4.47608
9	1576.50	3.2284e+28	3.4649e+31	16.0235	7.44294	8.86329e+29	11.942	36.2	11.4	1	5.04415
10	1587.60	5.2628e+28	3.8398e+31	15.5509	6.6713	1.13914e+30	13.46	42.9	12.8	1	5.80037
11	1586.00	3.4941e+28	3.7049e+31	14.5779	6.09192	9.70588e+29	15.319	47.9	14.7	1	6.63796
12	1579.20	7.2492e+28	7.4827e+31	13.7989	5.4856	1.9145e+30	17.381	49.8	16.6	1	7.43509
13	1582.10	7.7786e+28	9.726e+31	12.7438	5.26023	2.3793e+30	19.819	58.9	19	1	8.53331
14	1582.30	4.8684e+28	4.0551e+31	11.704	5.02029	1.12194e+30	22.323	70.5	21.2	1	9.50125
15	1582.40	5.302e+28	4.2341e+31	10.5713	4.85228	1.32108e+30	24.909	74.3	23.9	1	10.4411
16	1569.00	7.6988e+28	1.0455e+32	9.82618	4.707	2.46845e+30	27.542	81.3	26.5	1	11.4976
17	1577.20	5.5602e+28	5.1434e+31	9.07869	4.36767	1.3982e+30	30.367	96.7	28.8	1	12.7186
18	1586.80	7.0173e+28	6.885e+31	8.16917	3.88321	1.83915e+30	33.554	105.9	31.9	1	14.208
19	1589.10	4.3252e+28	4.2018e+31	7.58912	3.75247	1.11579e+30	36.997	113	35.5	1	15.6128
20	1574.00	5.0563e+28	6.6956e+31	7.12517	3.69968	1.60771e+30	40.375	127.4	38.4	1	17.2502
21	1586.40	4.3346e+28	3.4817e+31	6.78867	3.52007	9.71326e+29	44.185	129.9	41.65	1.1	19.3902
22	1572.20	4.4013e+28	5.3056e+31	6.37936	3.20387	1.30756e+30	48.855	153.4	46.1	1	21.7519
23	1589.20	4.548e+28	4.0933e+31	6.19702	3.02597	1.14978e+30	54.648	165	51.4	1	24.2091
24	1577.00	2.8636e+28	3.6508e+31	5.85996	3.01195	9.03874e+29	61.544	192.2	58.5	1.2	26.7226
25	1593.90	1.2503e+28	1.1311e+31	5.61566	2.78632	3.27528e+29	67.993	197.9	65.6	1	28.5889
26	1594.10	1.5443e+28	1.9581e+31	5.39205	2.64233	4.77131e+29	74.013	236.2	71.1	1	31.2196
27	1578.10	3.2769e+28	2.5653e+31	5.11317	2.50583	7.91447e+29	80.348	259.3	76.05	1.1	34.2008
28	1576.20	2.9404e+28	2.1409e+31	4.83251	2.33316	6.71334e+29	87.697	295.6	81.9	1	37.747
29	1571.80	1.9052e+28	3.0736e+31	4.5245	2.20025	7.08398e+29	95.076	295	87.7	1.1	41.3897
30	1587.50	4.6653e+28	7.619e+31	4.33858	1.95028	1.75969e+30	102.06	335.2	93.1	1.1	45.189

As can be seen in Table 14, the results are consistent, with the maximum difference being around 14%. If both collections of runs were run many times, then the difference would likely reduce as the results' randomness

would be minimised. However, the difference in time could result from not running both sets of runs in a controlled environment; they have both been run on a computer which, almost certainly, would have had different CPU loads at runtime, which could have impacted the performance.

Table 14: Comparison of statistics between experiment 3 and the extra runs.

Statistic	Experiment 3	Extra runs	Difference
Time taken (seconds)	6.98	7.79	10.40%
Lowest median fitness	4.11899	4.33858	5.06%
Lowest minimum fitness	2.18729	1.95028	-12.15%
Average size at gen 30	109.64	102.06	-7.43%
Max size at gen 30	323.3	335.2	3.55%
Median size at gen 30	105.8	93.1	-13.64%
Standard deviation at gen 30	45.4562	45.189	-0.59%

Figures 13 and 14 contain the plots of the original experiment and the plots of the extra runs, showing that the resulting solutions from each collection of runs are consistent.

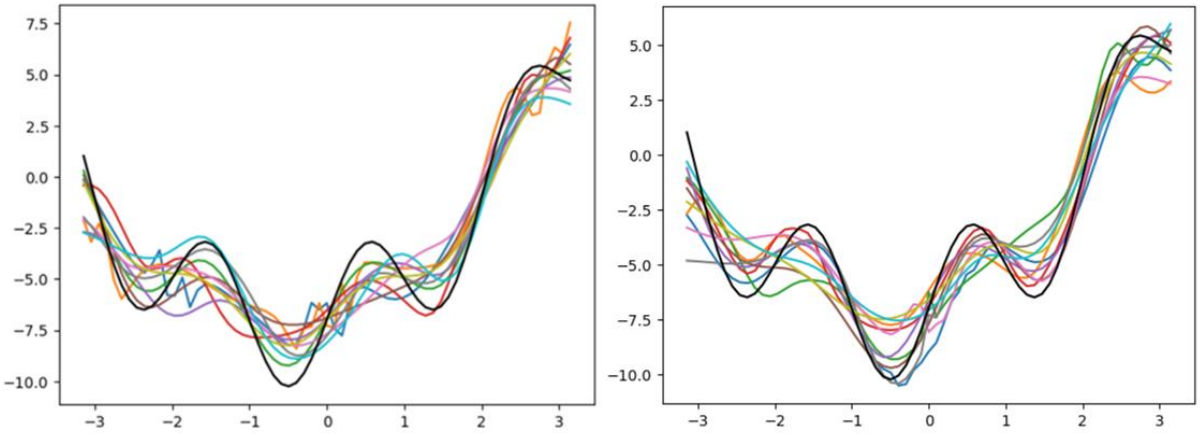

Figure 13: Best solutions, all runs vs the target function: experiment 3 (left), extra runs for problem 1 (right).

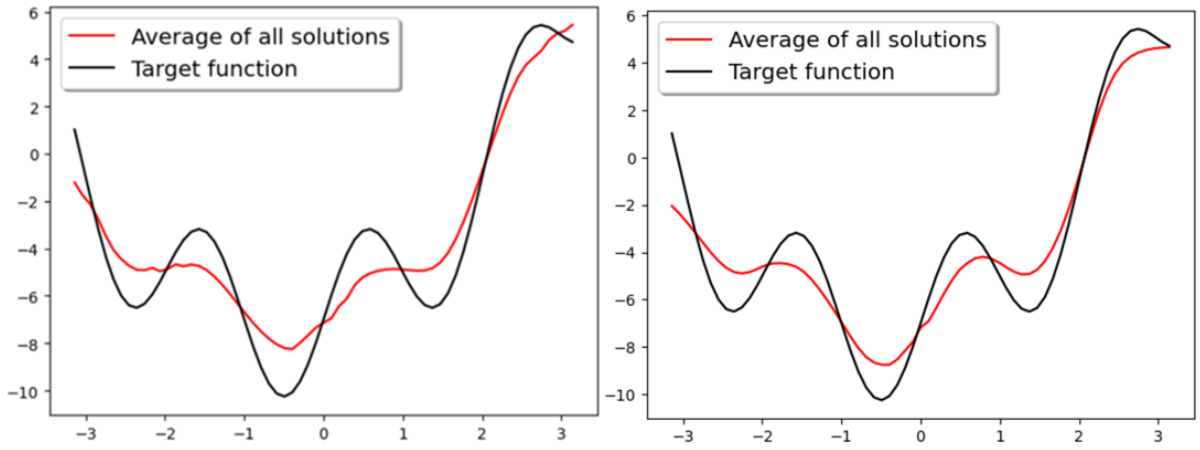

Figure 14: Average solution of all runs vs the target function: experiment 3 (left), extra runs for problem 1 (right).

3.2. Problem 2 further experiment

Experiment 7 has the population size set to 2,000, the tournament size set to 5, and the target function is problem 2. Again, this configuration was run ten more times, and the difference between the two will be discussed. Table 15 displays the full results from the additional runs.

Table 15: Average results from the extra runs for problem 2.

		Fitness					Size				
gen	nevals	avg	max	mdn	min	std	avg	max	mdn	min	std
0	2000.00	2.7978e+28	2.2639e+31	28.8367	24.1065	6.77376e+29	3.5847	7	3	2	1.50799
1	1597.60	1.8978e+28	2.7467e+31	27.1324	23.3479	6.58472e+29	3.6455	12.6	3	1	1.85767
2	1582.70	2.0844e+28	1.6958e+31	27.1215	22.7899	5.00443e+29	3.6803	14.5	3	1	2.01173
3	1587.80	3.2193e+28	2.9555e+31	26.9407	22.1087	8.1599e+29	3.8121	15.1	3.1	1	2.23358
4	1585.90	1.8449e+28	1.9151e+31	26.2847	21.9867	5.10619e+29	4.0422	16.9	3.8	1	2.43888
5	1584.30	3.7022e+28	3.9077e+31	26.1502	21.2503	1.03634e+30	4.4916	20.5	3.9	1	2.72107
6	1581.60	3.2215e+28	2.8529e+31	26.1261	20.3494	8.08474e+29	5.2152	23.1	4.8	1	3.0446
7	1580.80	9.5092e+28	6.005e+31	26.1987	20.0865	1.82978e+30	6.0428	23.8	5.4	1	3.31667
8	1571.40	4.4353e+58	8.8706e+61	26.1812	19.906	1.98303e+60	6.9373	25.8	6.3	1	3.62419
9	1577.40	6.8398e+57	1.368e+61	26.1869	18.9191	3.05809e+59	7.867	27.2	7.2	1	4.0358
10	1590.70	2.6323e+59	4.3775e+62	26.0907	17.8438	1.01715e+61	8.9244	30.1	8.5	1	4.49398
11	1574.80	2.7002e+59	5.4005e+62	25.8391	17.4524	1.20728e+61	10.128	33	9.6	1	4.99145
12	1585.00	9.6386e+59	1.6201e+63	25.6122	16.232	3.73189e+61	11.355	37.4	10.9	1	5.43236
13	1577.20	7.2204e+59	6.8515e+62	25.155	15.3682	2.09972e+61	12.676	40.2	12	1	5.94326
14	1590.20	2.9188e+91	5.8376e+94	25.0391	14.3071	1.30499e+93	13.943	43.9	13.2	1	6.4605
15	1564.40	1.4404e+92	2.8807e+95	24.4754	13.3462	6.43988e+93	15.43	48.7	14.65	1	7.1582
16	1575.30	5.9213e+59	9.7229e+62	24.1802	12.7534	2.32707e+61	16.897	53.7	16.1	1	7.80646
17	1572.80	3.9265e+59	4.0008e+62	23.5369	11.6735	1.14683e+61	18.589	57.1	18	1	8.24405
18	1582.90	1.2609e+60	8.7895e+62	22.8166	11.2609	2.92347e+61	20.351	61.8	19.5	1	8.75509
19	1578.80	1.3317e+60	2.4086e+63	22.3115	10.6517	5.41389e+61	22.317	70	21.3	1	9.58506
20	1576.40	9.9031e+59	1.916e+63	21.4142	10.4415	4.29681e+61	24.396	74.7	23.1	1	10.2775
21	1582.00	4.204e+59	6.8314e+62	21.1633	9.97836	1.5774e+61	26.225	75.9	24.9	1	10.962
22	1576.50	9.6139e+91	1.9228e+95	19.9833	9.51275	4.2984e+93	28.39	83.3	26.9	1	11.7197
23	1584.20	1.5789e+59	3.1579e+62	19.0407	8.93744	7.05948e+60	30.308	97.6	29	1	12.6461
24	1576.60	4.5519e+59	9.1038e+62	17.9773	8.63259	2.03516e+61	32.637	98.3	31.2	1	13.6417
25	1571.30	6.419e+58	1.2838e+62	16.3697	7.67702	2.86997e+60	35.007	100.4	33.2	1	14.4866
26	1584.90	1.7629e+59	3.5259e+62	15.3495	7.43553	7.88212e+60	37.264	113	35	1	15.871
27	1588.70	1.8021e+59	2.3204e+62	14.1015	6.95544	5.99228e+60	39.6	120.2	37.6	1	16.7822
28	1574.10	4.0928e+59	5.7462e+62	13.2338	6.6826	1.38603e+61	42.428	139.2	40.7	1	18.3773
29	1574.70	1.1644e+59	2.3288e+62	12.3988	6.39009	5.20614e+60	45.487	139.6	43.9	1	20.0372
30	1583.60	1.3006e+59	2.2819e+62	11.7974	6.1057	5.17969e+60	48.263	162.2	46.65	1	21.421

Table 16 shows that the results from the extra runs are consistent with experiment 7, with a maximum difference of 18%, and as mentioned previously, increasing the number of runs would likely reduce the difference. However, the findings from the extra runs show that the runs were worse by around 10-13% when comparing the fitness statistics; this means that the target function for problem 2 is harder to simulate as with 30 generations, there is much randomness in the results and is likely to have not converged fully by the 30th generation. Therefore, a higher generation count should also reduce the difference between sets of runs.

Table 16: Comparison of statistics between experiment 7 and the extra runs.

Statistic	Experiment 7	Extra runs	Difference
Time taken (seconds)	5.55	5.49	-1.09%
Lowest median fitness	10.2137	11.7974	13.42%
Lowest minimum fitness	5.48179	6.1057	10.22%
Average size at gen 30	53.914	48.263	-11.71%
Max size at gen 30	192.6	162.2	-18.74%
Median size at gen 30	49.6	46.65	-6.32%
Standard deviation at gen 30	25.2946	21.421	-18.08%

Compared to figures 13 and 14, a visible difference can be seen in figures 15 and 16, mainly figure 15, where many of the resultant functions from experiment 7 are noticeably different from the functions in the extra runs.

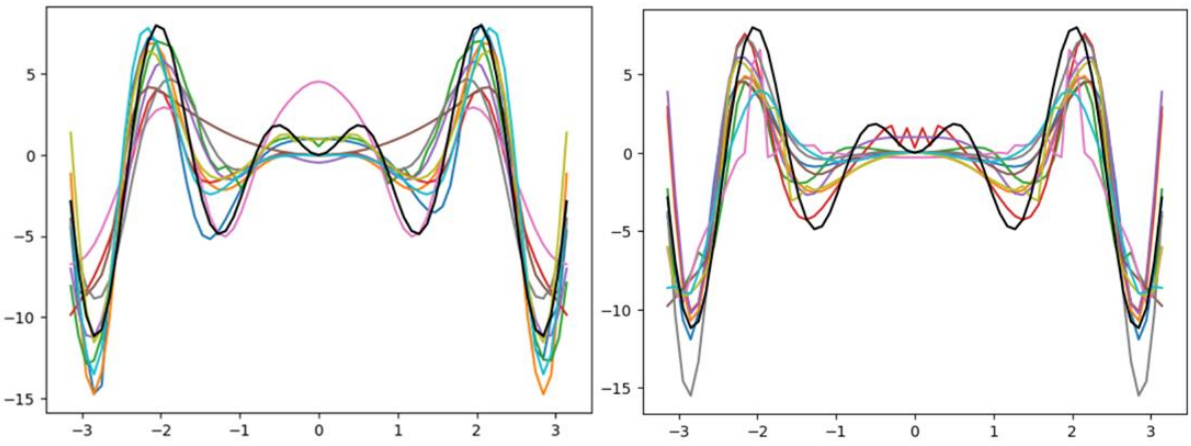

Figure 15: Best solutions, all runs vs the target function: experiment 7 (left), extra runs for problem 2 (right).

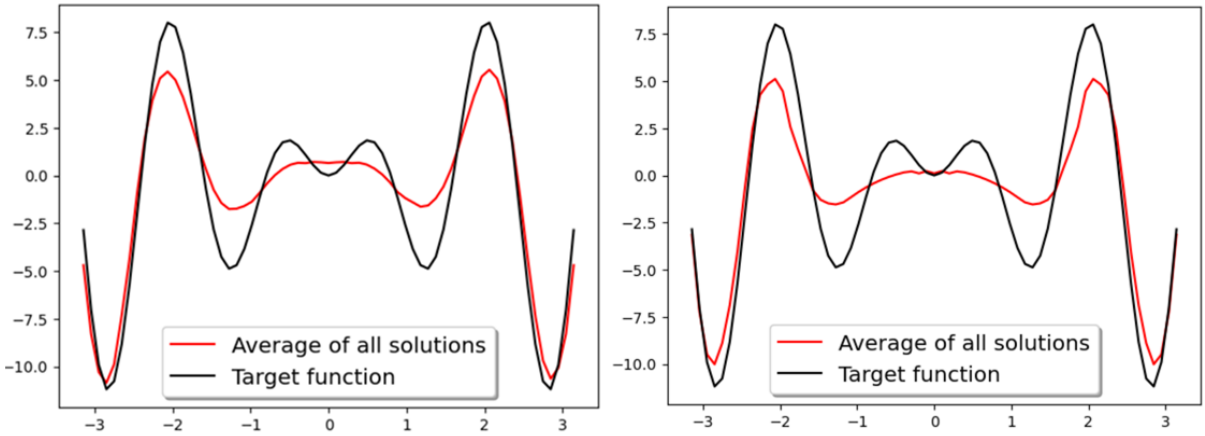

Figure 16: Average solution of all runs vs the target function: experiment 3 (left), extra runs for problem 2 (right).

4. Conclusion

This report concludes that changing the tournament size is more effective than changing the population size for evolving functions to reach a target function. As the generations progress, the median fitness decreases, indicating that the solution is getting closer to the target function since the fitness is being minimised; also, during the generations, the size increases, meaning that the resulting functions are more complex and contain more instructions, this indicates that the time taken to compute is increased. It is pointed out that not every run of the GP produces consistent functions and may differ in appearance from other runs. It is theorised that increasing the number of runs will make the results more consistent. The report also suggests that generating solutions for problem 2 is more complex than problem 1 and that the best statistic to identify the best parameter configuration is the median fitness, as large numbers impact other statistic measures, such as average, because it is sensitive to outlier values.

Generally, the GP solves the problems with reasonable accuracy with the discussed parameters. However, these results should be interpreted carefully as GP is a stochastic process, and different runs may produce different results. Moreover, further experiments can be conducted to explore the impact of the other parameters, such as the number of generations, mutation rate and crossover rate and to investigate the effect of increasing the number of runs in terms of consistency.

5. Supplementary Analysis and Further Experiments

Experiment 3 was repeated twice with 50 runs to analyse the theory that an increased number of runs will reduce the difference between the statistics of 2 different sets of runs, improving overall consistency. Table 17 proves that the theory is verified as the difference is much less than the difference from the previous comparison; this also shows that ten runs do not result in a comprehensive range of results.

Table 17: Comparison of statistics between 2 sets of runs (increased run count) and the difference and the previous difference.

Statistic	Set 1	Set 2	Difference	Previous difference
Time taken (seconds)	7.13	7.24	1.52%	10.40%
Lowest median fitness	4.28511	4.15585	-3.11%	5.06%
Lowest minimum fitness	2.02347	2.10862	4.04%	-12.15%
Average size at gen 30	89.2	92.944	4.03%	-7.43%
Max size at gen 30	270.72	283.54	4.52%	3.55%
Median size at gen 30	84.73	87.28	2.92%	-13.64%
Standard deviation at gen 30	36.7552	39.8273	7.71%	-0.59%

As mentioned in the report, a further experiment was conducted to analyse if problem 2 can become more consistent with a higher number of generations. Two sets of runs consisting of 10 runs each were performed using the parameters below:

- Problem 2
- 50 generations
- 2,000 population size
- 5 tournament size

The results of this experiment can be seen compared in Table 18. With the increased number of generations, there has been an improvement in the consistency of the two sets of runs. However, the results still show a slight inconsistency and are only about half as good at improving consistency as the previous experiment.

Table 18: Comparison of statistics between 2 sets of runs (generations) and the difference and the previous difference.

Statistic	Set 1	Set 2	Difference	Previous difference
Time taken (seconds)	15.60	15.76	1.02%	-1.09%
Lowest median fitness	2.81695	2.69053	-4.70%	13.42%
Lowest minimum fitness	1.7715	1.84306	3.88%	10.22%
Average size at gen 50	128.21	143.39	10.59%	-11.71% (gen 30)
Max size at gen 50	390.7	431.2	9.39%	-18.74% (gen 30)
Median size at gen 50	121.35	134.35	9.68%	-6.32% (gen 30)
Standard deviation at gen 50	54.7838	59.1831	7.43%	-18.08% (gen 30)

Other observations made during this experiment include the average time taken per run has increased compared to experiment 7, which has an average time per run at 5.55 seconds, whereas this experiment has an average time per run of 15.68; this is an increase by factor of 3 primarily due to the increase in the size of the evolved function towards the later generations; the average size at the final generation for this experiment is 135.8, compared to that of 53.194, which is also nearly increased by a factor of 3 (2.55). Furthermore, as a result of running more generations, the resultant functions also more closely match the target function, as shown in figure 17.

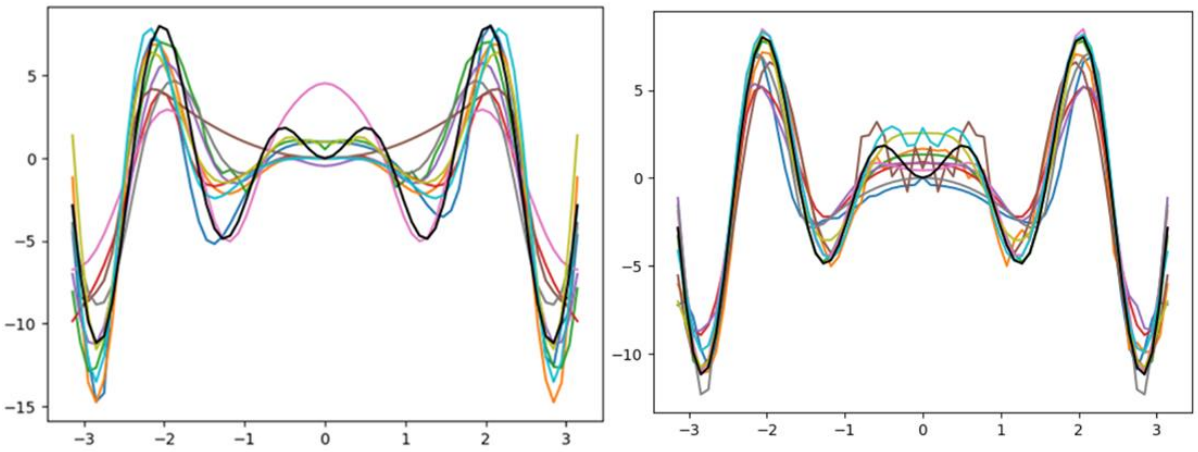

Figure 17: Best solutions, all runs vs the target function: experiment 7 (left), set 1 (right).