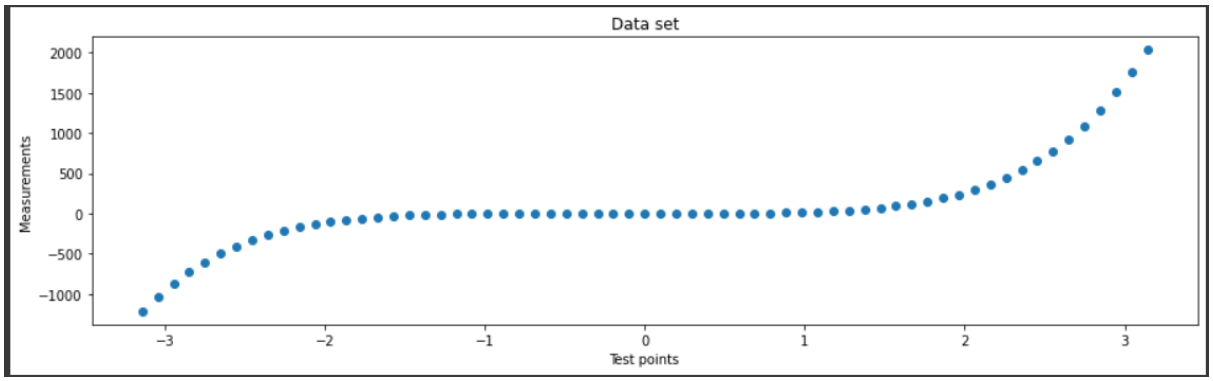
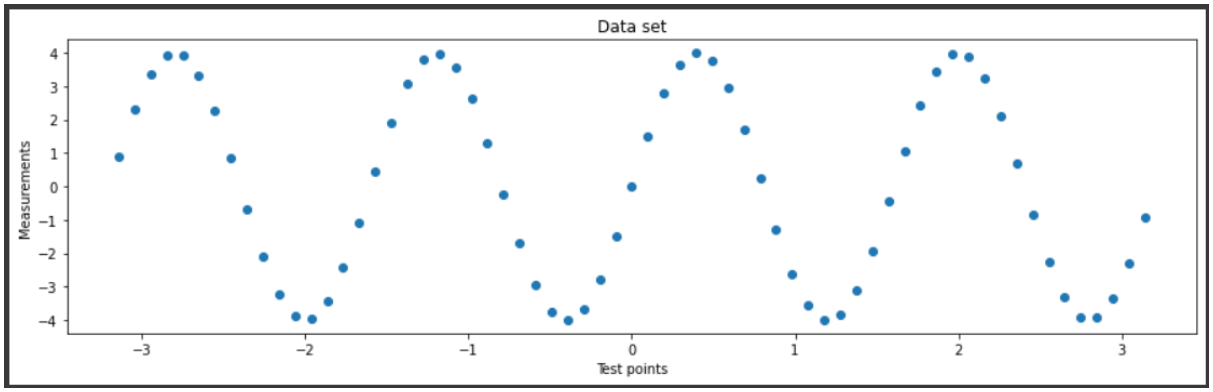
Drew Berry RegNo: 2000878

CE310 Assignment Part 2

Data set for $p_1(x)$



Data set for $p_2(x)$



Drew Berry RegNo: 2000878

Experiments

In this section Max Fitness and Min Size are my control variables.

Experiment 1

Problem = $p_1(x)$

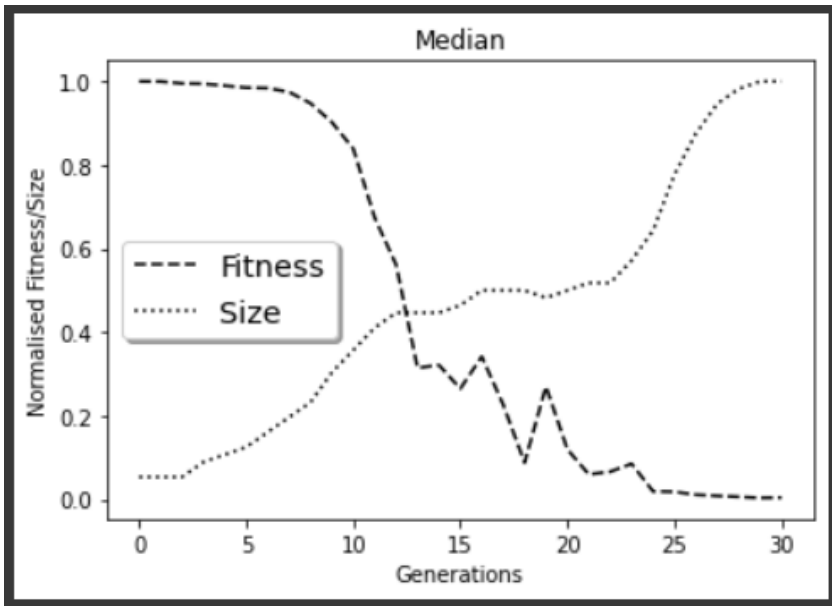
Parameters

- Population = 500
- Tournaments = 5

Test Data:

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	1894.552	307677.047	3.0	65.0	12349
2	3342.965	307580.481	3.0	67.0	12284
3	680.033	307677.047	3.0	74.0	12275
4	17874.904	307580.481	3.0	65.0	12304
5	4001.345	307677.047	3.0	83.0	12411
6	634.147	307677.047	3.0	51.0	12418
7	10133.197	307580.481	3.0	46.0	12362
8	154.565	307580.481	3.0	67.0	12422
9	453.671	307580.481	3.0	90.0	12322
10	10446.573	307585.098	3.0	48.0	12352

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 2

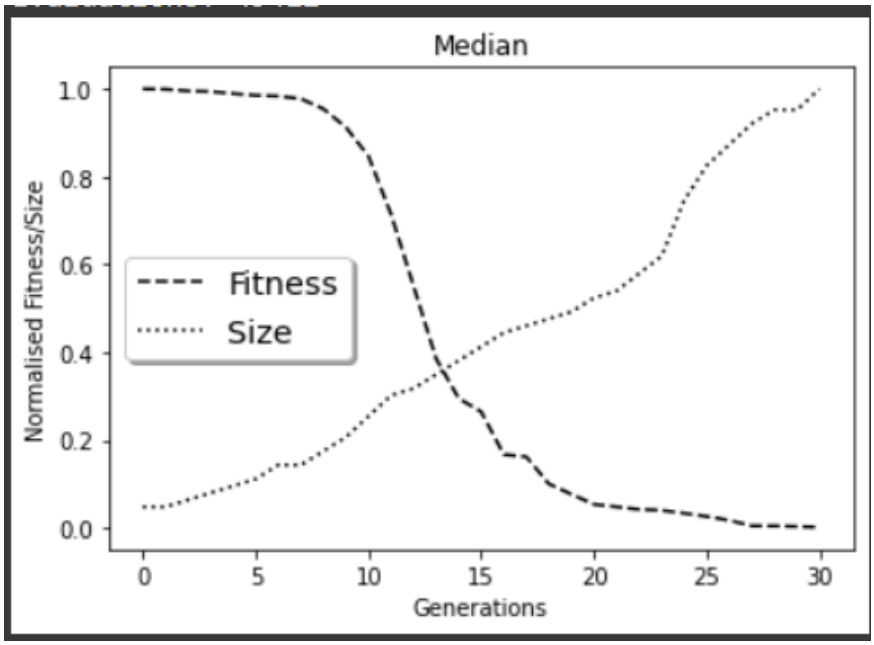
Problem = $p_1(x)$

Parameters

- Population = 2000
- Tournaments = 5

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	428.376	307642.678	3.0	55.0	49498
2	216.009	307596.278	3.0	80.5	49399
3	2220.609	307615.570	3.0	58.0	49665
4	1168.258	307580.481	3.0	73.0	49664
5	2972.447	307580.481	3.0	54.0	49508
6	1324.364	307677.047	3.0	76.0	49503
7	863.569	307585.098	3.0	63.0	49469
8	1945.770	307645.281	3.0	44.0	49386
9	3477.223	307677.047	3.0	44.0	49404
10	1876.974	307580.481	3.0	48.0	49393

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 3

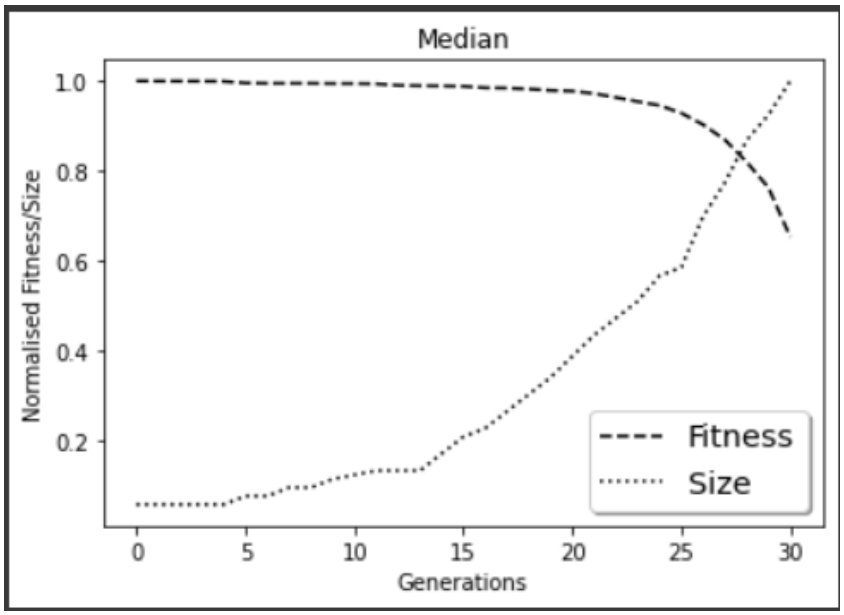
Problem = $p_1(x)$

Parameters

- Population = 500
- Tournaments = 2

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	145908.454	307580.481	3.0	53.0	12369
2	208489.836	307636.726	3.0	49.0	12283
3	171633.027	307580.481	3.0	60.0	12301
4	153582.920	307580.481	3.0	56.0	12329
5	178128.551	307677.047	3.0	71.0	12330
6	235666.839	307677.047	3.0	42.0	12422
7	276823.119	307580.481	3.0	40.0	12389
8	183133.754	307580.481	3.0	47.0	12247
9	228647.392	307677.047	3.0	57.0	12398
10	189693.967	307630.751	3.0	43.0	12295

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 4

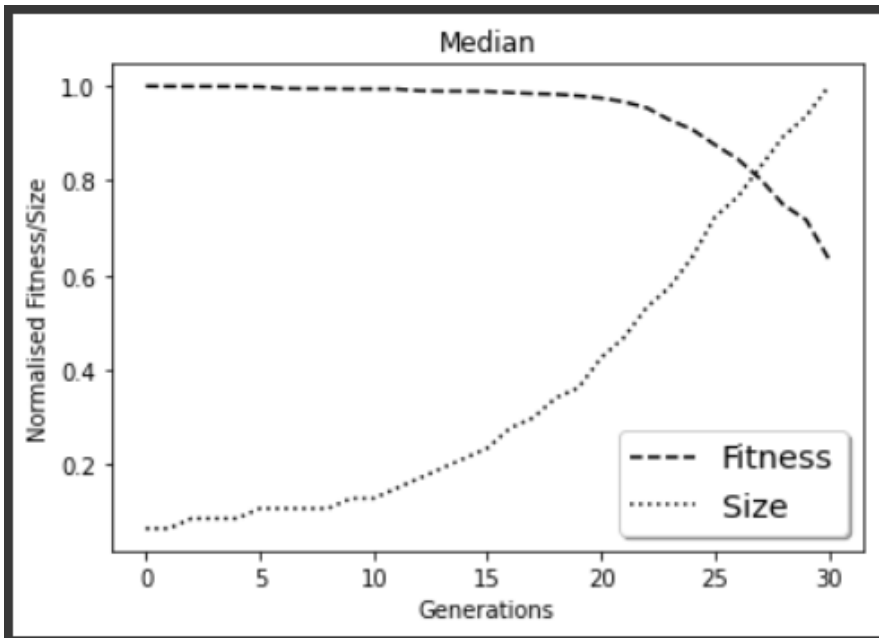
Problem = $p_1(x)$

Parameters

- Population = 2000
- Tournaments = 2

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	200495.268	307580.481	3.0	46.0	49562
2	202074.816	307677.047	3.0	41.0	49331
3	214300.976	307587.669	3.0	46.0	49516
4	235792.731	307600.553	3.0	52.0	49387
5	164247.688	307580.481	3.0	52.0	49300
6	233530.930	307594.850	3.0	36.0	49569
7	208675.840	307677.047	3.0	44.0	49376
8	215007.990	307634.885	3.0	69.0	49522
9	210096.639	307585.098	3.0	57.0	49530
10	218506.648	307649.800	3.0	40.0	49299

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 5

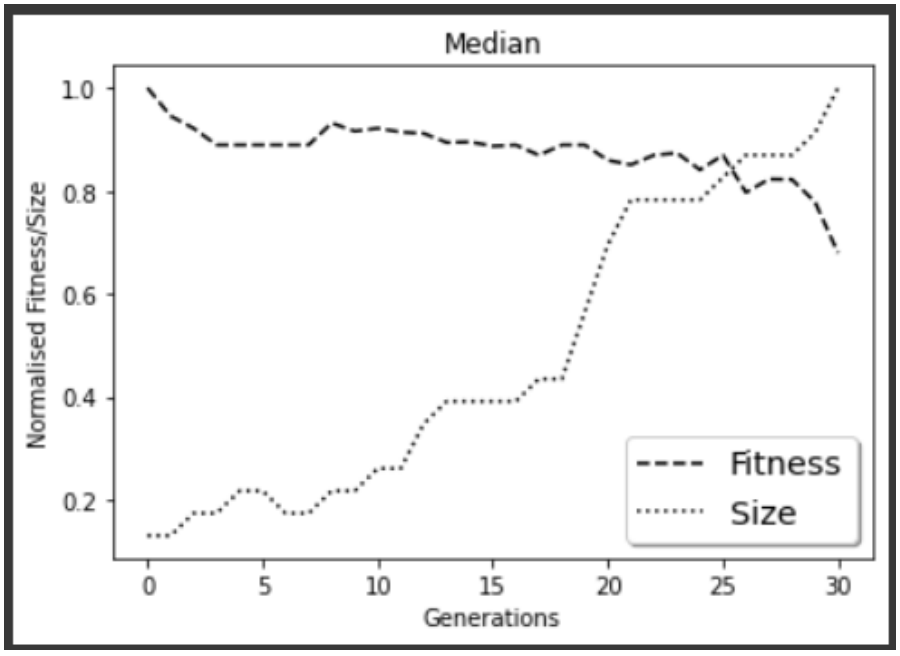
Problem = $p_2(x)$

Parameters

- Population = 500
- Tournaments = 5

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	6.731	9.023	3.0	93.0	12373
2	6.424	9.023	3.0	22.0	12323
3	4.941	9.023	3.0	34.0	12356
4	7.292	9.023	3.0	18.0	12217
5	6.573	9.023	3.0	39.0	12425
6	5.876	9.023	3.0	34.0	12251
7	7.397	9.023	3.0	24.0	12376
8	4.490	9.023	3.0	87.0	12421
9	7.576	9.023	3.0	24.0	12276
10	7.610	9.023	3.0	18.0	12405

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 6

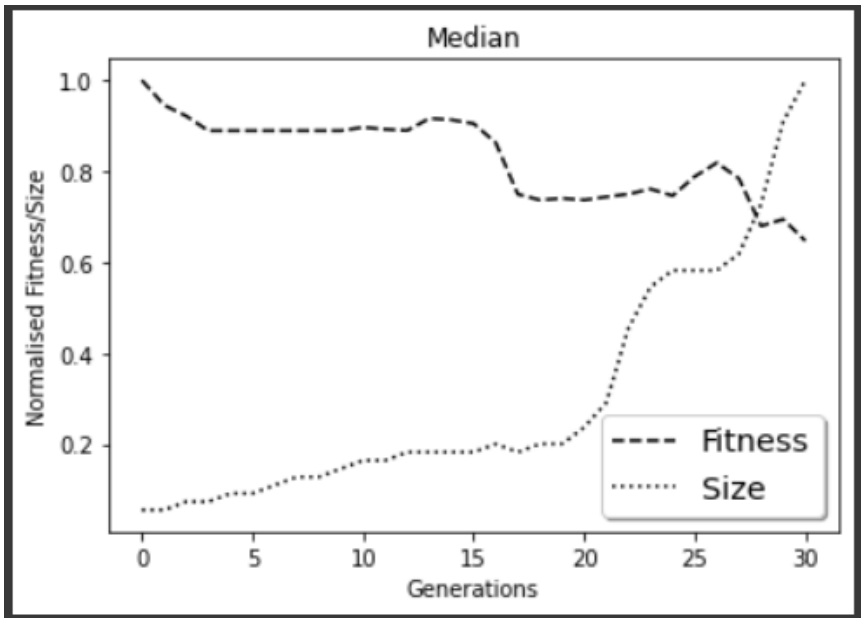
Problem = $p_2(x)$

Parameters

- Population = 2000
- Tournaments = 5

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	4.356	9.023	3.0	29.0	49517
2	4.505	9.023	3.0	70.0	49290
3	4.169	9.023	3.0	47.0	49381
4	6.695	9.023	3.0	25.0	49307
5	4.464	9.023	3.0	24.0	49359
6	4.092	9.023	3.0	27.0	49457
7	6.465	9.023	3.0	28.0	49276
8	4.241	9.023	3.0	47.0	49538
9	2.807	9.023	3.0	47.0	49617
10	4.476	9.023	3.0	53.0	49283

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 7

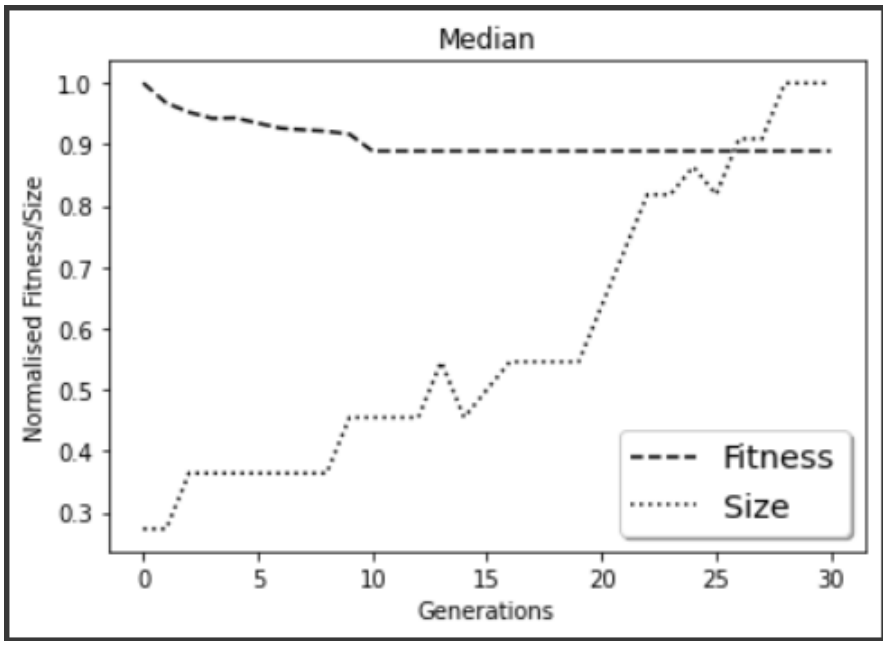
Problem = $p_2(x)$

Parameters

- Population = 500
- Tournaments = 2

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	8.023	9.023	3.0	21.0	12367
2	8.023	9.023	3.0	15.0	12269
3	8.023	9.023	3.0	9.0	12353
4	8.023	9.023	3.0	14.0	12306
5	8.004	9.023	3.0	20.0	12391
6	8.023	9.023	3.0	15.0	12405
7	8.023	9.023	3.0	15.0	12264
8	8.023	9.023	3.0	14.0	12330
9	8.023	9.023	3.0	12.0	12416
10	8.023	9.023	3.0	20.0	12271

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Experiment 8

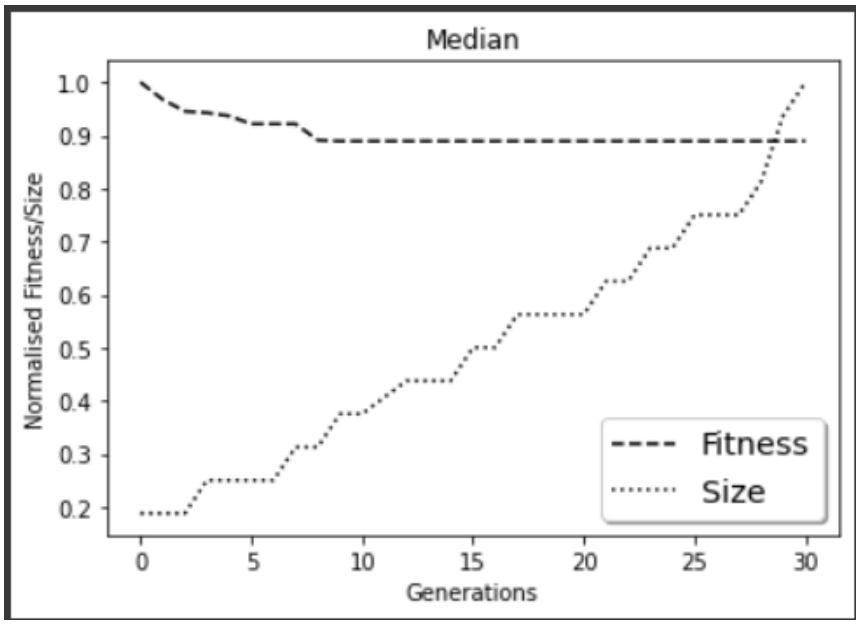
Problem = $p_2(x)$

Parameters

- Population = 2000
- Tournaments = 2

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	8.023	9.023	3.0	15.0	49328
2	8.023	9.023	3.0	14.0	49664
3	8.023	9.023	3.0	11.0	49413
4	8.023	9.023	3.0	11.0	49330
5	8.023	9.023	3.0	22.0	49177
6	8.023	9.023	3.0	16.0	49461
7	7.783	9.023	3.0	24.0	49273
8	8.023	9.023	3.0	12.0	49561
9	8.023	9.023	3.0	18.0	49173
10	8.023	9.023	3.0	15.0	49492

This graph is an example of the graphs generated by the program. It shows that the GP is minimizing the fitness.



Drew Berry RegNo: 2000878

Best Parameter Configuration

For both problems, the parameters below enabled the best outcomes.

- Population = 2000
- Tournaments = 5

So, these were the results after running another 10 tests for each problem:

Problem = p1(x)

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	2220.609	307615.570	3.0	58.0	49665
2	216.009	307596.278	3.0	67.5	49399
3	1935.770	307645.281	3.0	45.0	49386
4	1158.258	307580.481	3.0	73.0	49664
5	2971.447	307580.481	3.0	54.0	49508
6	1223.364	307677.047	3.0	77.0	49503
7	428.376	307642.678	3.0	65.0	49498
8	1915.770	307645.281	3.0	42.0	49386
9	3474.223	307677.047	3.0	44.0	49404
10	1068.258	307580.481	3.0	72.0	49664

Problem = p2(x)

Tests	Min Fitness	Max Fitness	Min Size	Max Size	Evaluations
1	4.229	9.023	3.0	35.0	49371
2	4.595	9.023	3.0	40.0	49160
3	4.508	9.023	3.0	33.0	49356
4	4.608	9.023	3.0	50.0	49208
5	4.499	9.023	3.0	31.0	49323
6	7.554	9.023	3.0	25.0	49298
7	3.865	9.023	3.0	103.0	49329
8	1.545	9.023	3.0	85.0	49440
9	4.207	9.023	3.0	45.0	49255
10	5.338	9.023	3.0	52.0	49441

Drew Berry RegNo: 2000878

Questions

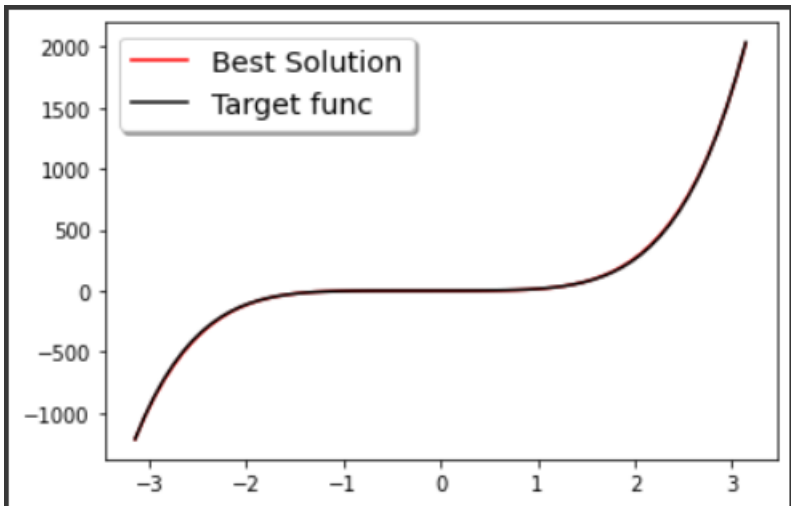
- GP Behaviour
  - When changing the problem, the results change completely as the have different data sets.
  - When using a population size of 2000 (compared to 500), the program takes longer to complete the 30 generations but finds a better solution. This is because more evaluations have to be performed.
  - When using a tournament size of 2 (compared to 5), the program finds a weaker solution. This is because tournaments are less likely to find good individuals to perform crossover and mutation on.

- Variations
  - Through all experiments, the max size obtained seemed to variate often.
  - However, for both problems, the average max fitness was the same.
  - Also, for $p_2(x)$ with a population of 2000 (experiments 7-8), the average min fitness was the same.

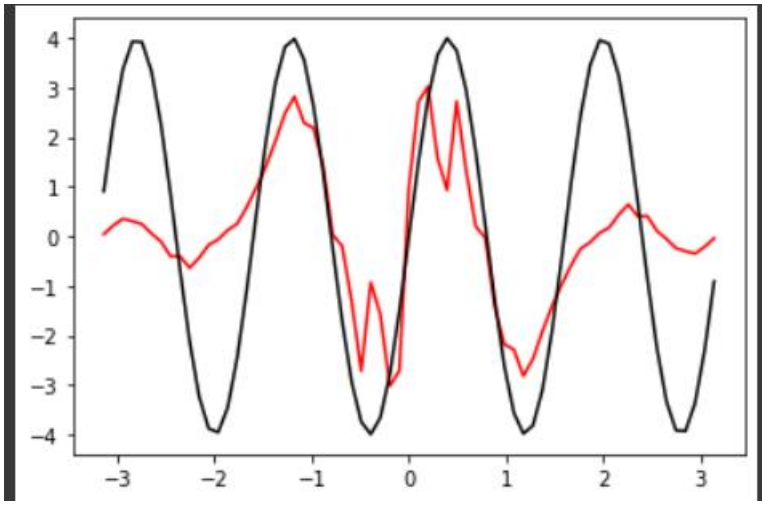
To choose the best parameter configuration, I looked at which configuration generated the lowest min fitness, which happened to be a population size of 2000 and a tournament size of 5 for both problems.

My results from the additional 10 runs seem to be consistent with the first 10 runs. However, two solutions were found with surprisingly low fitness values (see test 2 for $p_1(x)$ and test 8 for $p_2(x)$)

The GP was quite suitable for solving $p_1(x)$, as shown below:



However, the GP was not very suitable for solving $p_2(x)$, as shown below:



The best solution line in red does seem to follow the sign curve to some extent but not nearly enough to say that the GP is suitable.

Conclusion

In this report, I have concluded that using a larger population size and a reasonable tournament size can lead to the best solutions to problems, having compiled a suitable amount of evidence to back up this claim. Also, I have discovered that this GP can be suitable for basic problems such as $p_1(x)$, but it may not be suitable for more complex problems such as $p_2(x)$.