COMP SCI 5401 FS2017 Assignment 1c

John Niemeyer
JJNB78@mst.edu

October 9, 2017

Self-Adaptive EA Strategy Parameter

The self-adaptive parameter that was chosen for this EA was the mutation rate. The self adaptation that was created for the mutation rate is very simple, it simply changes based on the number of iterations gone between mutations. So if there are too many mutations happening the code will increase the mutation rate by a factor of 0.001 or 0.1

Investigation

a.)

For problem 1, I would say that the self-adaptivity did not have a very large impact on the average fitness, but it very well may have increased the best fitness by quite a lot due to the ability of getting a population out of a local optimum. For problem 2 it seems like the average fitness is extremely low while the best fitness is doing just fine. I believe this is due to the mutation rate getting too high, though it seems to work fine for Problem 1 and Problem 3. Problem 3, much like Problem 1, seems to be doing quite well. The average gets to a certain point on all three problems and steadies out like it should, however for Problem 2 something is holding it back and I am not entirely sure what.

So there is definitely a differing benefit between the problems, with Problem 1 and Problem 3 looking about the same and Problem 2 being underwhelming in terms of average fitness. The average fitness does reach a certain point on all datasets and hover around that point, which is a good thing when it comes to the averaging of populations.

b.)

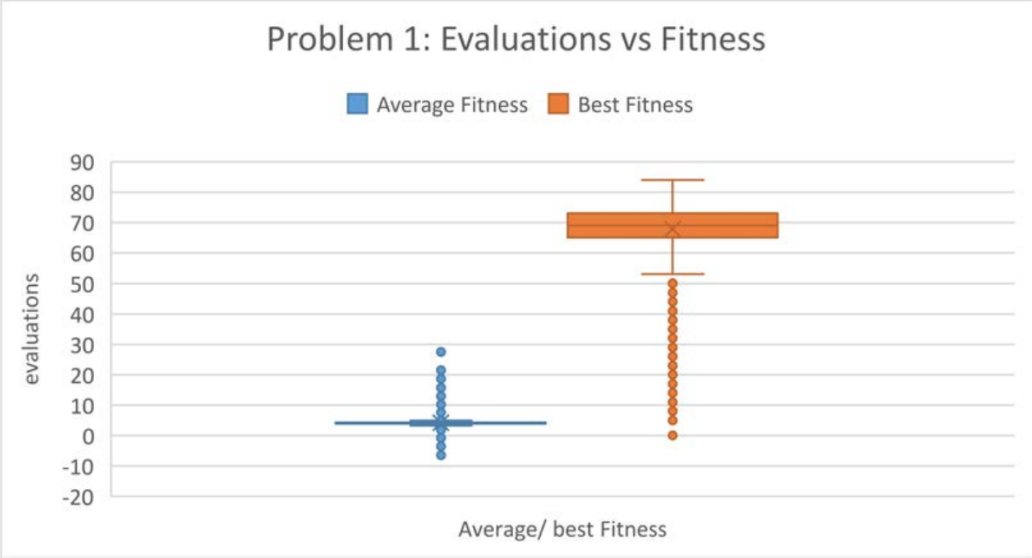
The penalty function coefficient will greatly impact the performance of all three problems. If I was to increase the penalty function coefficient there would be a severe drop in best fitness and average fitness due to how the penalty function currently works, likewise if the penalty function coefficient was decreased the best fitness and average fitness would go up since the penalty function coefficient directly decreases the fitness of a given dataset.

It seems that the penalty coefficient in Problem 1 and Problem 3 could both use a little lower penalty function in order to obtain a better average fitness, but Problem 2 is having a very hard time with a penalty function coefficient of 1. This means that Problem 2 needs to have a smaller penalty function coefficient than Problem 1 and Problem 3, which is mainly due to having double the shapes of Problem 1 and a smaller width of 20 and length while Problem 3 has a larger width of 30 and length to work with.

Experiment parameters and graphs

1 Problem 1

1.1 Graphs



1.2 Result Tables

Problem 1: final results

Assignment 1c: Problem 1		
Run	Average Fitness	best Fitness
1	5.4	68
2	4.74	77
3	4.67	79
4	4.4	75
5	4.48	77
6	4.23	69
7	4.36	75
8	4.19	71
9	4.16	65
10	4.13	69
11	4.03	77
12	4.11	73
13	4.32	73
14	4.15	80
15	4.02	76
16	4.21	73
17	4.08	77
18	3.99	73
19	3.89	72
20	3.88	74
21	3.84	77
22	3.82	71
23	4.02	74
24	4.01	75
25	3.85	65
26	3.72	70
27	4.01	66
28	3.65	84
29	3.77	72
30	3.86	69

1.3 Statistical Analysis

Problem 1: Best Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	15	68	-53	2809
2	12	77	-65	4225
3	6	79	-73	5329
4	2	75	-73	5329
5	20	77	-57	3249
6	10	69	-59	3481
7	7	75	-68	4624
8	11	71	-60	3600
9	13	65	-52	2704
10	15	69	-54	2916
11	13	77	-64	4096
12	9	73	-64	4096
13	17	73	-56	3136
14	16	80	-64	4096
15	17	76	-59	3481
16	21	73	-52	2704
17	8	77	-69	4761
18	16	73	-57	3249
19	17	72	-55	3025
20	14	74	-60	3600
21	8	77	-69	4761
22	17	71	-54	2916
23	12	74	-62	3844
24	6	75	-69	4761
25	14	65	-51	2601
26	16	70	-54	2916
27	6	66	-60	3600
28	8	84	-76	5776
29	13	72	-59	3481
30	11	69	-58	3364
SUM:			-1826	112530
	t-value:	-1.23094		df: 29
	P-Value:	0.2282		t-value: 2.045

Problem 1: Average Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	5	5.4	-0.4	0.16
2	5	4.74	0.26	0.0676
3	0	4.67	-4.67	21.8089
4	0	4.4	-4.4	19.36
5	9	4.48	4.52	20.4304
6	3	4.23	-1.23	1.5129
7	4	4.36	-0.36	0.1296
8	1	4.19	-3.19	10.1761
9	4	4.16	-0.16	0.0256
10	1	4.13	-3.13	9.7969
11	7	4.03	2.97	8.8209
12	3	4.11	-1.11	1.2321
13	8	4.32	3.68	13.5424
14	6	4.15	1.85	3.4225
15	3	4.02	-1.02	1.0404
16	9	4.21	4.79	22.9441
17	5	4.08	0.92	0.8464
18	9	3.99	5.01	25.1001
19	3	3.89	-0.89	0.7921
20	6	3.88	2.12	4.4944
21	3	3.84	-0.84	0.7056
22	3	3.82	-0.82	0.6724
23	6	4.02	1.98	3.9204
24	2	4.01	-2.01	4.0401
25	3	3.85	-0.85	0.7225
26	6	3.72	2.28	5.1984
27	2	4.01	-2.01	4.0401
28	1	3.65	-2.65	7.0225
29	6	3.77	2.23	4.9729
30	5	3.86	1.14	1.2996
SUM:			4.01	198.2979
	t-value:	0.0095054		df: 29
	P-Value:	0.9925		t-value: 2.045

So according to the statistical analysis (shown above) the p-value for both best fitness and average fitness is not low enough to say that the results are statistically significant. That means that the t-value of -1.23094 and the t-value of 0.0095054, computed using the tables given, were not far enough apart from the t-value given of 2.045 to make the difference in the fitness values statistically significant.

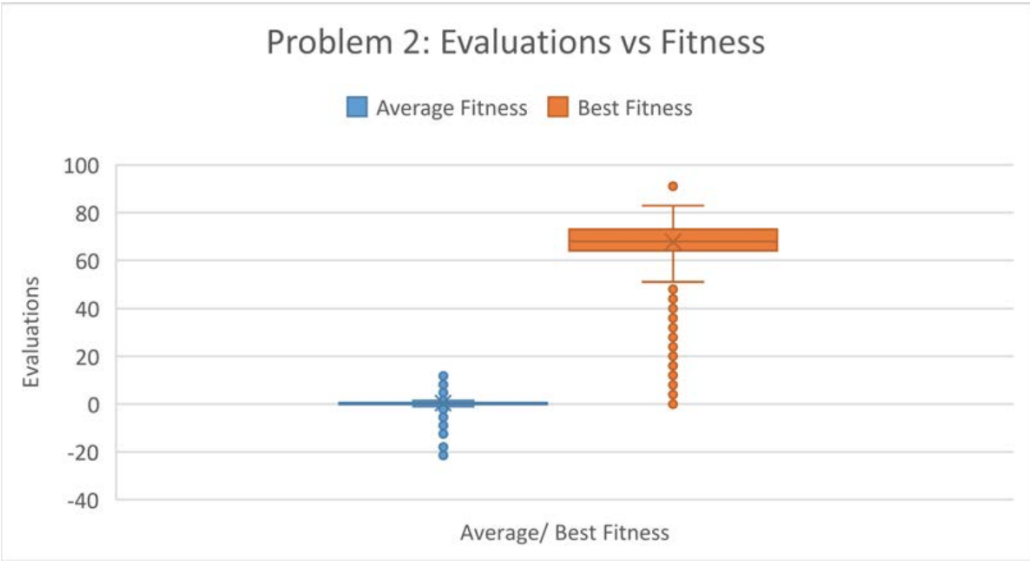
1.4 EA Configurations

If you want to get the same results you have to change the newSeed variable to 0 (Zero) in the configuration file in order to use the previous seed.

```
Random = 0
EA = 1
newSeed = 1

mu: 20
lambda: 10
penalty: 1
runs: 30
mutation_rate: 0.01
fitness_evaluations: 10000
prob_log_random: logs/prob1_random_log.txt
prob_log_EA: logs/prob1_EA_log.txt
number_of_evals_till_termination: 5
tournament_size_for_parent_selection: 10
tournament_size_for_survival_selection: 10
n_for_termination_convergence_criterion: 5
prob_solution_random: solutions/prob1_random_solution.txt
prob_solution_EA: solutions/prob1_EA_solution.txt
seed: time.time()

selfAdaptive: adaptMutation: 1

Initialization: Uniform_Random: 1

Parent_Selection: Uniform_random_parent: 0,
    Fitness_Proportional_Selection: 0, k-
    Tournament_Selection_with_replacement: 1

Survival_Strategy: plus: 0, comma: 1

Survival_Selection: Uniform_random_survival: 0, Truncation: 0, k-
    Tournament_Selection_without_replacement: 1

Termination: Number_of_evals: 0,
    no_change_in_average_population_fitness_for_n_generations: 0,
    no_change_in_best_fitness_in_population_for_n_generations: 1
```

2 Problem 2

2.1 Graphs



2.2 Result Tables

Problem 2: final results

Assignment 1c: Problem 2		
Run	Average Fitness	best Fitness
1	2.96	78
2	1.57	77
3	1.3	91
4	0.83	69
5	0.47	71
6	0.63	72
7	1.02	72
8	0.12	70
9	0.57	74
10	0.34	67
11	0.36	75
12	0.06	92
13	0.36	73
14	0.28	74
15	0.08	66
16	0.18	70
17	-0.15	65
18	-0.04	69
19	0	69
20	0.1	68
21	0.04	75
22	0.02	78
23	-0.25	73
24	-0.25	75
25	-0.07	83
26	-0.03	66
27	0.03	68
28	-0.17	68
29	-0.12	70
30	-0.32	81

2.3 Statistical Analysis

Problem 2: Best Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	17	68	-51	2601
2	16	77	-61	3721
3	11	79	-68	4624
4	15	75	-60	3600
5	11	77	-66	4356
6	17	69	-52	2704
7	8	75	-67	4489
8	17	71	-54	2916
9	12	65	-53	2809
10	9	69	-60	3600
11	12	77	-65	4225
12	16	73	-57	3249
13	6	73	-67	4489
14	9	80	-71	5041
15	15	76	-61	3721
16	30	73	-43	1849
17	10	77	-67	4489
18	11	73	-62	3844
19	7	72	-65	4225
20	22	74	-52	2704
21	15	77	-62	3844
22	23	71	-48	2304
23	17	74	-57	3249
24	12	75	-63	3969
25	14	65	-51	2601
26	17	70	-53	2809
27	11	66	-55	3025
28	16	84	-68	4624
29	11	72	-61	3721
30	16	69	-53	2809
SUM:			-1773	106211
	t-value:	-1.26389		df: 29
	P-Value:	0.2163		t-value: 2.045

Problem 2: Average Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	6	2.96	3.04	9.2416
2	8	1.57	6.43	41.3449
3	5	1.3	3.7	13.69
4	5	0.83	4.17	17.3889
5	2	0.47	1.53	2.3409
6	1	0.63	0.37	0.1369
7	3	1.02	1.98	3.9204
8	4	0.12	3.88	15.0544
9	5	0.57	4.43	19.6249
10	4	0.34	3.66	13.3956
11	3	0.36	2.64	6.9696
12	6	0.06	5.94	35.2836
13	2	0.36	1.64	2.6896
14	2	0.28	1.72	2.9584
15	4	0.08	3.92	15.3664
16	5	0.18	4.82	23.2324
17	4	-0.15	4.15	17.2225
18	4	-0.04	4.04	16.3216
19	2	0	2	4
20	7	0.1	6.9	47.61
21	3	0.04	2.96	8.7616
22	12	0.02	11.98	143.5204
23	6	-0.25	6.25	39.0625
24	6	-0.25	6.25	39.0625
25	5	-0.07	5.07	25.7049
26	5	-0.03	5.03	25.3009
27	5	0.03	4.97	24.7009
28	5	-0.17	5.17	26.7289
29	7	-0.12	7.12	50.6944
30	7	-0.32	7.32	53.5824
SUM:			133.08	744.912
	t-value:	0.382909		df: 29
	P-Value:	0.7046		t-value: 2.045

So according to the statistical analysis (shown above) the p-value for both best fitness and average fitness is not low enough to say that the results are statistically significant. That means that the t-value of -1.26389 and the t-value of 0.382909, computed using the tables given, were not far enough apart from the t-value given of 2.045 to make the difference in the fitness values statistically significant.

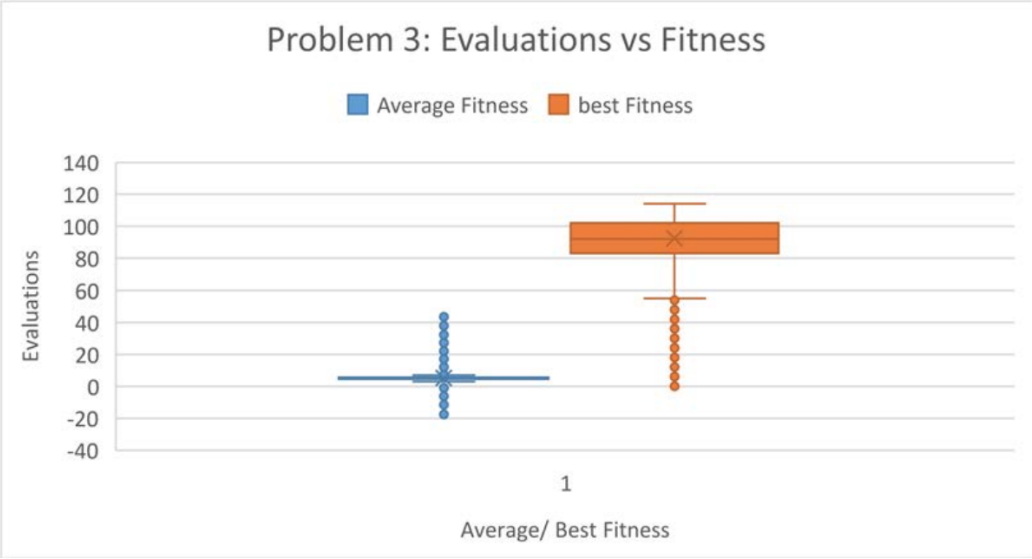
2.4 EA Configurations

If you want to get the same results you have to change the newSeed variable to 0 (Zero) in the configuration file in order to use the previous seed.

```
Random = 0
EA = 1
newSeed = 1

mu: 20
lambda: 10
penalty: 1
runs: 30
mutation_rate: 0.01
fitness_evaluations: 10000
prob_log_random: logs/prob2_random_log.txt
prob_log_EA: logs/prob2_EA_log.txt
number_of_evals_till_termination: 5
tournament_size_for_parent_selection: 10
tournament_size_for_survival_selection: 10
n_for_termination_convergence_criterion: 5
prob_solution_random: solutions/prob2_random_solution.txt
prob_solution_EA: solutions/prob2_EA_solution.txt
seed: time.time()

selfAdaptive: adaptMutation: 1

Initialization: Uniform_Random: 1

Parent_Selection: Uniform_random_parent: 0,
    Fitness_Proportional_Selection: 0, k-
    Tournament_Selection_with_replacement: 1

Survival_Strategy: plus: 0, comma: 1

Survival_Selection: Uniform_random_survival: 0, Truncation: 1, k-
    Tournament_Selection_without_replacement: 0

Termination: Number_of_evals: 0,
    no_change_in_average_population_fitness_for_n_generations: 0,
    no_change_in_best_fitness_in_population_for_n_generations: 1
```

3 Problem 3

3.1 Graphs



3.2 Result Tables

Problem 3: final results

Assignment 1c: Problem 3		
Run	Average Fitness	best Fitness
1	7.42	89
2	6.66	91
3	5.69	90
4	6.1	92
5	5.59	101
6	5.64	90
7	5.3	102
8	5.42	99
9	5.23	106
10	5.48	108
11	4.51	94
12	4.54	98
13	4.75	102
14	5.04	95
15	5.19	113
16	5.04	114
17	4.64	92
18	5.41	93
19	4.18	113
20	4.41	106
21	4.71	108
22	5.01	92
23	4.76	82
24	5.06	106
25	4.58	113
26	4.45	111
27	4.54	112
28	4.67	112
29	4.49	106
30	4.39	95

3.3 Statistical Analysis

Problem 3: Best Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	50	68	-18	324
2	28	77	-49	2401
3	31	79	-48	2304
4	43	75	-32	1024
5	41	77	-36	1296
6	34	69	-35	1225
7	34	75	-41	1681
8	25	71	-46	2116
9	32	65	-33	1089
10	35	69	-34	1156
11	19	77	-58	3364
12	23	73	-50	2500
13	50	73	-23	529
14	19	80	-61	3721
15	30	76	-46	2116
16	36	73	-37	1369
17	30	77	-47	2209
18	30	73	-43	1849
19	43	72	-29	841
20	36	74	-38	1444
21	27	77	-50	2500
22	22	71	-49	2401
23	29	74	-45	2025
24	55	75	-20	400
25	41	65	-24	576
26	51	70	-19	361
27	28	66	-38	1444
28	40	84	-44	1936
29	55	72	-17	289
30	29	69	-40	1600
SUM:			-1150	48090
	t-value:	-0.76874		df: 29
	P-Value:	0.4483		t-value: 2.045

Problem 3: Average Fitness				
subject #	Fitness 1b	fitness 1c	x-y	(x-y)^2
1	16	7.42	8.58	73.6164
2	5	6.66	-1.66	2.7556
3	4	5.69	-1.69	2.8561
4	15	6.1	8.9	79.21
5	13	5.59	7.41	54.9081
6	14	5.64	8.36	69.8896
7	21	5.3	15.7	246.49
8	6	5.42	0.58	0.3364
9	14	5.23	8.77	76.9129
10	14	5.48	8.52	72.5904
11	5	4.51	0.49	0.2401
12	7	4.54	2.46	6.0516
13	16	4.75	11.25	126.5625
14	6	5.04	0.96	0.9216
15	4	5.19	-1.19	1.4161
16	10	5.04	4.96	24.6016
17	15	4.64	10.36	107.3296
18	4	5.41	-1.41	1.9881
19	5	4.18	0.82	0.6724
20	16	4.41	11.59	134.3281
21	6	4.71	1.29	1.6641
22	12	5.01	6.99	48.8601
23	7	4.76	2.24	5.0176
24	20	5.06	14.94	223.2036
25	7	4.58	2.42	5.8564
26	15	4.45	10.55	111.3025
27	5	4.54	0.46	0.2116
28	15	4.67	10.33	106.7089
29	19	4.49	14.51	210.5401
30	12	4.39	7.61	57.9121
SUM:			175.1	1854.9542
	t-value:	0.206653		df: 29
	P-Value:	0.8377		t-value: 2.045

So according to the statistical analysis (shown above) the p-value for both best fitness and average fitness is not low enough to say that the results are statistically significant. That means that the t-value of -0.76874 and the t-value of 0.206653, computed using the tables given, were not far enough apart from the t-value given of 2.045 to make the difference in the fitness values statistically significant.

3.4 EA Configurations

If you want to get the same results you have to change the newSeed variable to 0 (Zero) in the configuration file in order to use the previous seed.

```
Random = 0
EA = 1
newSeed = 1

mu: 20
lambda: 10
penalty: 1
runs: 30
mutation_rate: 0.01
fitness_evaluations: 10000
prob_log_random: logs/prob3_random_log.txt
prob_log_EA: logs/prob3_EA_log.txt
number_of_evals_till_termination: 5
tournament_size_for_parent_selection: 10
tournament_size_for_survival_selection: 10
n_for_termination_convergence_criterion: 5
prob_solution_random: solutions/prob3_random_solution.txt
prob_solution_EA: solutions/prob3_EA_solution.txt
seed: time.time()

selfAdaptive: adaptMutation: 1

Initialization: Uniform_Random: 1

Parent_Selection: Uniform_random_parent: 1,
    Fitness_Proportional_Selection: 0, k-
    Tournament_Selection_with_replacement: 0

Survival_Strategy: plus: 0, comma: 1

Survival_Selection: Uniform_random_survival: 1, Truncation: 0, k-
    Tournament_Selection_without_replacement: 0

Termination: Number_of_evals: 0,
    no_change_in_average_population_fitness_for_n_generations: 0,
    no_change_in_best_fitness_in_population_for_n_generations: 1
```