Jordan Giacone
Jmgv27@mst.edu
COMP SCI 5401 FS2016 Assignment 1d

Multi-objective Evolutionary Algorithms

In this assignment, our task was to create a multi-objective EA that focused on minimizing runtime as well as memory usage. This is accomplished by evaluating each member of the population in terms of the runtime and memory usage, then building what is known as a "Pareto Hierarchy". A Pareto Hierarchy is a list of lists where the individuals in each list do not dominate each other and all individuals in lower lists are dominated by at least one individual in a higher list. An individual is said to dominate another individual if and only if it scores just as highly in all objectives and is better in at least one objective. The highest level – level 1 is called the Pareto Front. The Pareto Front is the list of all non-dominated individuals and is analogous to the best-performing members of the entire population. Using this idea of Pareto domination levels, all of the familiar selection algorithms are easily adapted to the realm of multi-objective EAs by substituting the singular fitness goal with the individual's Pareto level.

Diversity

For most of the pareto fronts, the diversity is fairly high. Since this EA is optimizing three different objectives, there is significant tradeoff between the three, which creates a nice diverse pareto front. A diversity score close to one (1) denotes a highly diverse Pareto front, while a diversity score near zero (0) denotes a non-diverse front. A table of diversities for each experiment can be found in Table 1

Table 1

	Control1	Control2	3Restart1	3Restart2	FitnessProportional1	FitnessProportional2
Run 1	0.999934387	0.998977	0.999021	0.853744	0.963124172	0.999971303
Run 2	0.999901432	0.99998	0.999292	0.999977	0.992754603	0.999779933
Run 3	0.932635294	0.999944	0.895081	0.999962	0.999709011	0.999963454
Run 4	0.974678303	0.999575	0.999692	0.999977	0.998319311	0.997058102
Run 5	0.999457229	0.998796	0.958861	0.999105	0.764375704	0.999976973
Average	0.981321329	0.999454	0.970389	0.970553	0.94365656	0.999349953

Jordan Giacone
Jmgv27@mst.edu
COMP SCI 5401 FS2016 Assignment 1d

Things to note:

- Control Tests 1 and 2 were run with a μ of 50, however due to how long the evaluation of 50 individuals on the test set took, this number was reduced to 25 for subsequent experiments
- All graphs are graphing the "normalized fitness value" vs number of evaluations. This was done because the range of fitness values possible on each of the objectives made for an uninteresting graph. The normalized value is a decimal between zero and 1 where 1 represents the max for that run. Basically, each cell was transformed according to this formula: cell = cell / max(cellStart:cellEnd)

Configurations Used

1. Control Configuration (mu + lambda)
   a. Population Size: 50
   b. Offspring per generation: 25
   c. Parent Selection: tournament selection w/ replacement
   d. Survival Selection: Truncation

   Running on the control configuration was interesting. As seen in figures Figure 1 and Figure 2, the averages for memory usage and number of decisions was all over the place. This could be due to the truncation survival selection, which simply sorts the list by pareto level and then takes the top mu individuals

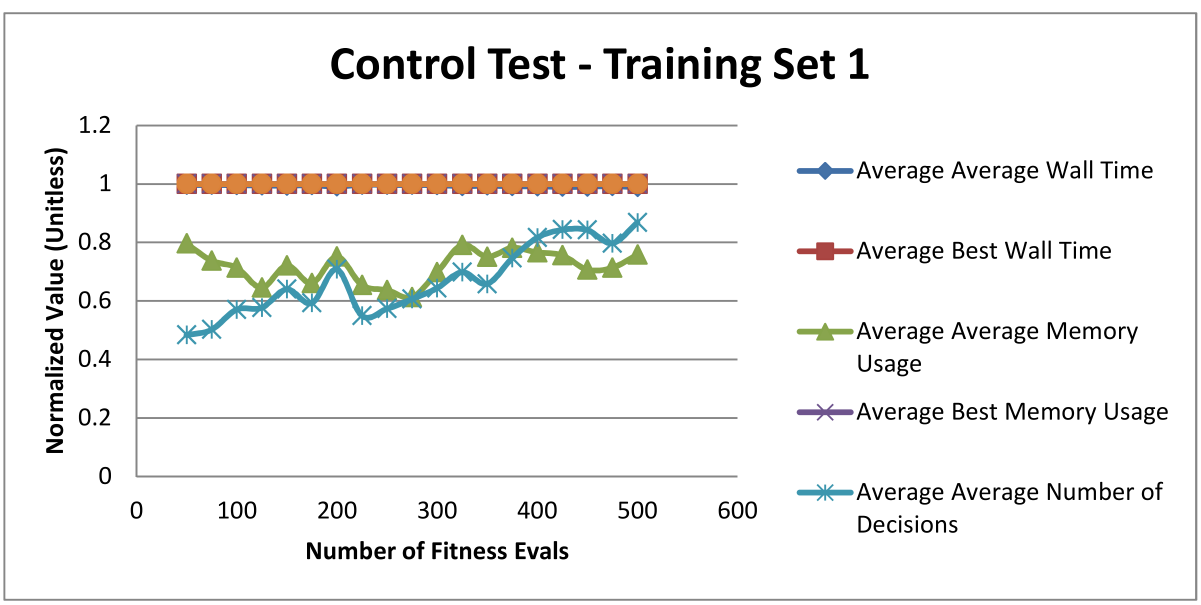

Figure 1

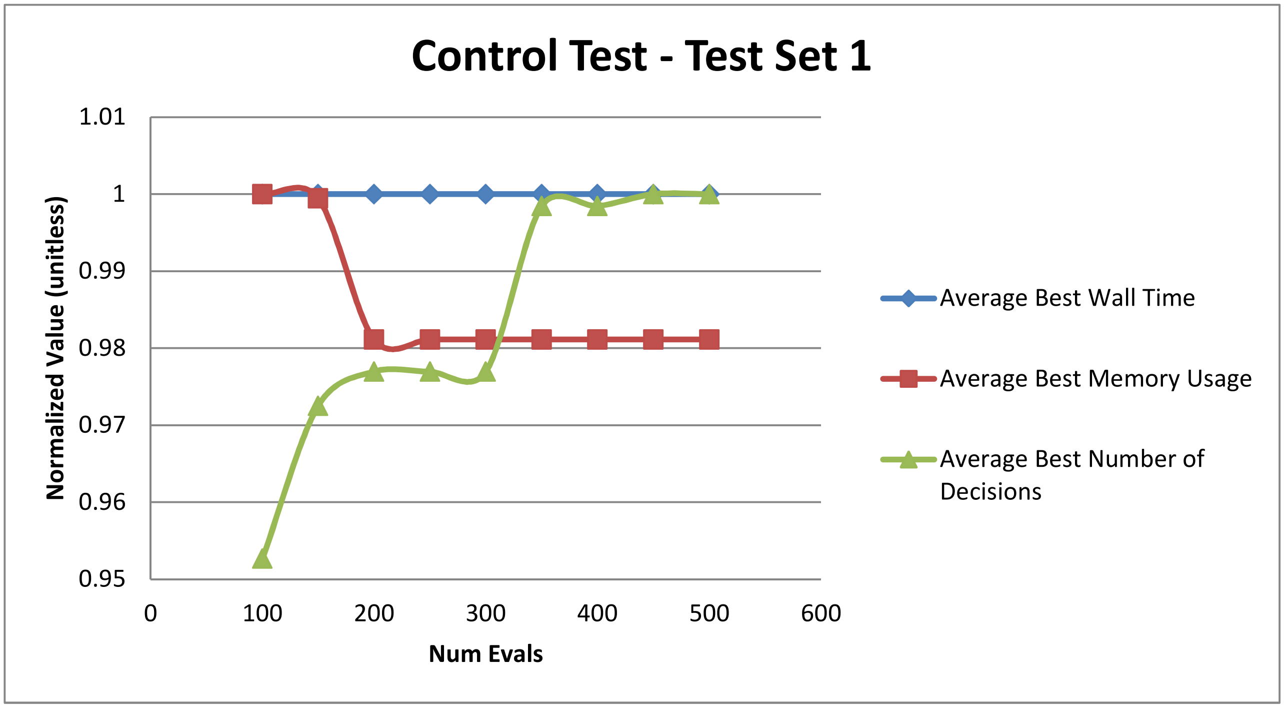

Figure 2

2. 3-Restart Configuration (mu + lambda)
    a. Population Size: 25
    b. Offspring per generation: 10
    c. Parent Selection: tournament selection w/ replacement
    d. Survival Selection: Truncation
    e. R-Restart Value: 3

As shown in figures Figure 3 and Figure 4, the 3-restart configuration seems to have done the best for all three objectives, with each one steadily increasing over the run. This may be due to the restarting nature of the run, where stagnation is rewarded with a nuke and then having the three best individuals survive the fallout to build the population up again

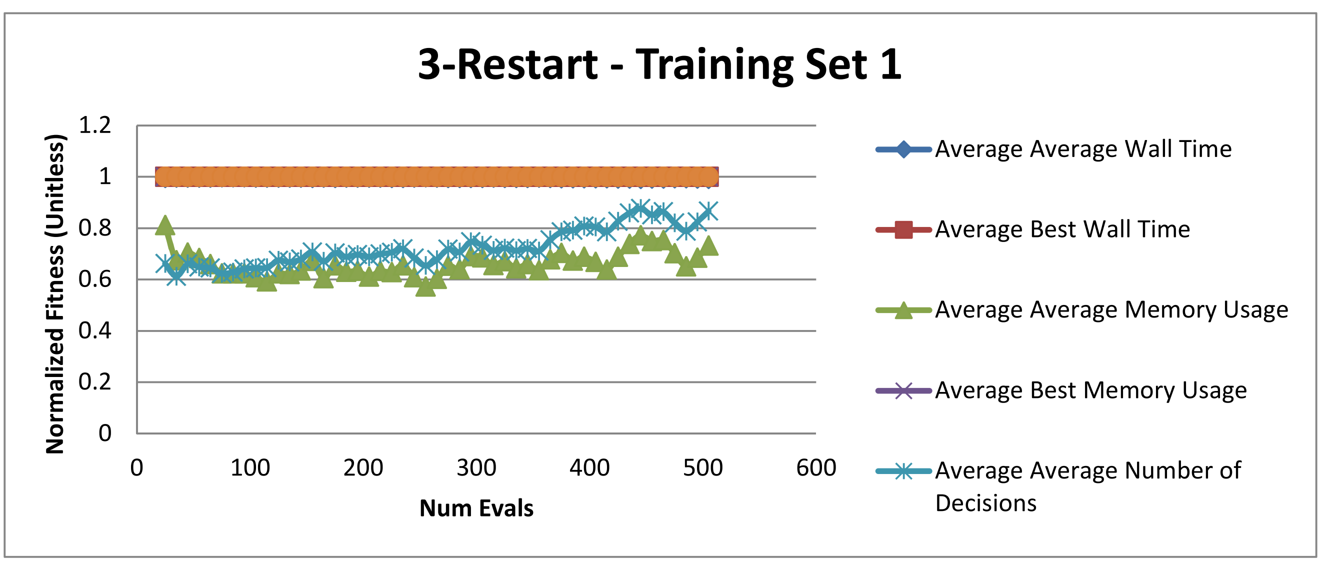

Figure 3

Figure 4

3. Fitness-Proportional Configuration (mu + lambda)
    a. Population Size: 25
    b. Offspring per generation: 10
    c. Parent Selection: Fitness-proportional
    d. Survival Selection: Fitness-proportional

Strangely enough on the fitness-proportional configuration, memory usage seemed to suffer. Not quite sure why, but the suffering was quick and painful as shown in figures Figure 5 and Figure 6

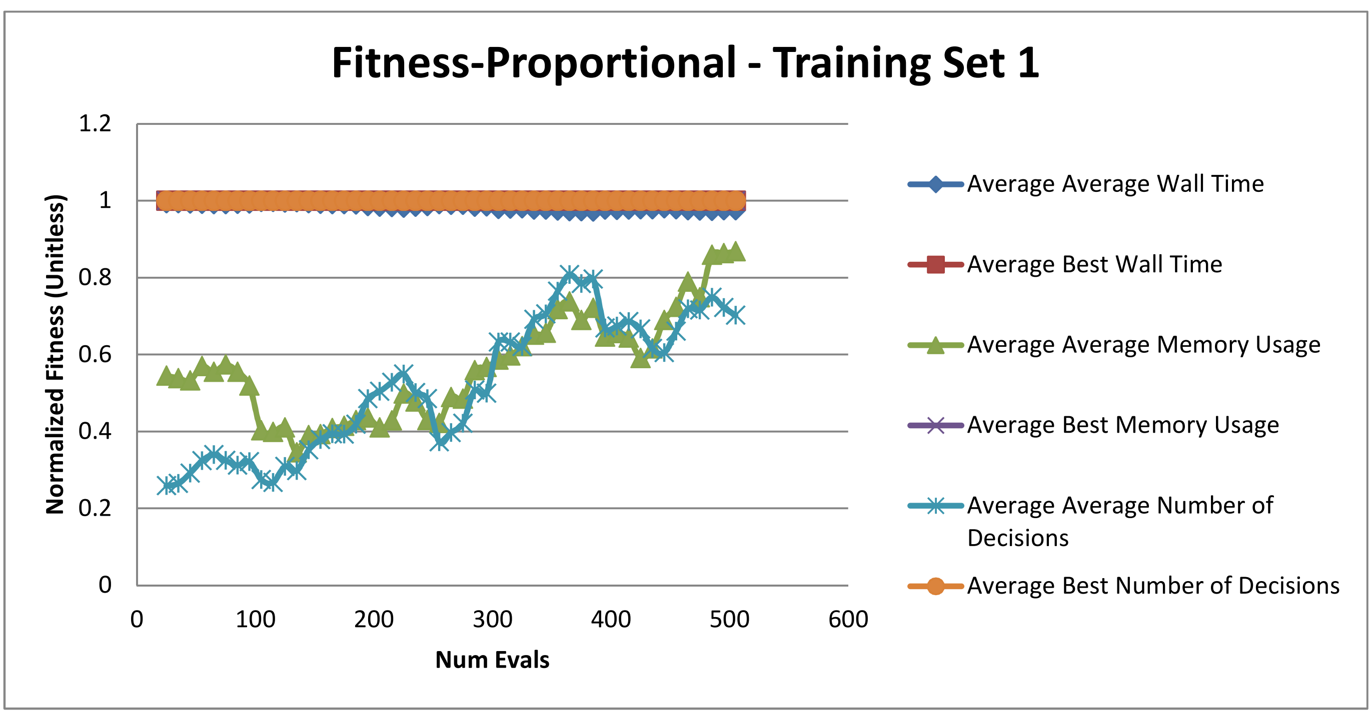

Figure 5

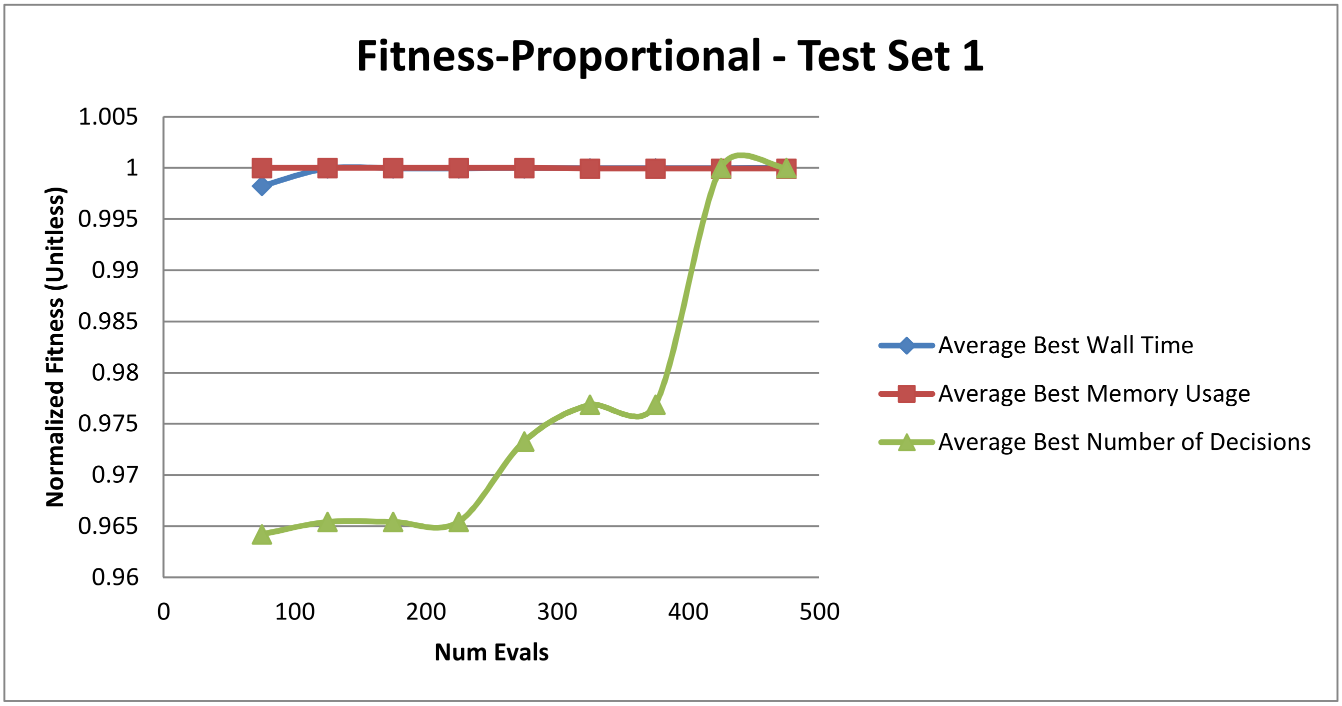

Figure 6

Statistical Analysis:

1. Control Test Vs. 3-Restart Test on Dataset 1

    a. When comparing the control test versus the 3-restart test on dataset 1, the results were surprising. After an initially good outlook from the F-Test, a T-Test assuming equal variances was performed. However, the T-Test showed that there is a significant statistical difference between the two. The control test is statistically better that the 3-restart test on dataset 1. The statistics can be seen in Table 2.

F-Test Two-Sample for Variances

	Control1	*3Restart1*
Mean	0.76	0.24
Variance	0.068	0.028
Observations	5	5
df	4	4
F	2.428571429	
P(F<=f) one-tail	0.205584491	
F Critical one-tail	6.388232909	
Assume Equal Variances by step 1		

t-Test: Two-Sample Assuming Equal Variances

	Control1	*3Restart1*
Mean	0.76	0.24
Variance	0.068	0.028
Observations	5	5
Pooled Variance	0.048	
Hypothesized Mean Difference	0	
df	8	
t Stat	3.75277675	
P(T<=t) one-tail	0.002800961	
t Critical one-tail	1.859548038	
P(T<=t) two-tail	0.005601922	
t Critical two-tail	2.306004135	
Reject null hypothesis		

Table 2

2. Control Test Vs. Fitness-Proportional Test on Dataset 1

   a. After performing an F-Test and not being able to assume equal variances, the T-Test then delivered more bad news. The T-Test failed to reject the null hypothesis so there is no statistical difference between the control and fitness-proportional tests on dataset 1. The statistics can be found in Table 3.

F-Test Two-Sample for Variances		
	Control1	*FitnessProportional1*
Mean	0.4	0.6
Variance	0.16	0.04
Observations	5	5
df	4	4
F	4	
P(F<=f) one-tail	0.104	
F Critical one-tail	6.388232909	
Assume Unequal Variances by step 3		

t-Test: Two-Sample Assuming Unequal Variances		
	Control1	*FitnessProportional1*
Mean	0.4	0.6
Variance	0.16	0.04
Observations	5	5
Hypothesized Mean Difference	0	
df	6	
t Stat	-1	
P(T<=t) one-tail	0.177958842	
t Critical one-tail	1.943180281	
P(T<=t) two-tail	0.355917684	
t Critical two-tail	2.446911851	
Cannot reject null hypothesis		

Table 3

3. 3-Restart Test Vs. Fitness-Proportional Test on Dataset 1

   a. After performing an F-Test and assuming equal variances, the T-Test revealed that there is a statistically significant difference between the Fitness-Proportional and 3-Restart tests. Since the fitness-proportional test has the higher mean, it can be concluded that it performs better. The statistics can be found in Table 4

F-Test Two-Sample for Variances		
	FitnessProportional1	*3Restart1*
Mean	0.72	0.2
Variance	0.112	0.04
Observations	5	5
df	4	4
F	2.8	
P(F<=f) one-tail	0.171307771	
F Critical one-tail	6.388232909	
Assume Equal Variances by step 1		

t-Test: Two-Sample Assuming Equal Variances		
	FitnessProportional1	*3Restart1*
Mean	0.72	0.2
Variance	0.112	0.04
Observations	5	5
Pooled Variance	0.076	
Hypothesized Mean Difference	0	
df	8	
t Stat	2.98240454	
P(T<=t) one-tail	0.008767976	
t Critical one-tail	1.859548038	
P(T<=t) two-tail	0.017535952	
t Critical two-tail	2.306004135	
Reject null hypothesis		

Table 4

Jordan Giacone
Jmgv27@mst.edu

4. Control Test Vs. 3-Restart Test on Dataset 2

   a. When comparing the control test versus the 3-restart test on dataset 2, the results were surprising. After an initially good outlook from the F-Test, a T-Test assuming equal variances was performed. However, the T-Test showed that there is not a significant statistical difference between the two. The statistics can be seen in Table 5.

F-Test Two-Sample for Variances		
	Control2	*3Restart2*
Mean	0.68	0.32
Variance	0.012	0.172
Observations	5	5
df	4	4
F	0.069767442	
P(F<=f) one-tail	0.012[illegible]05145	
F Critical one-tail	0.156537812	
Assume Equal Variances by step 1		

t-Test: Two-Sample Assuming Equal Variances		
	Control2	*3Restart2*
Mean	0.68	0.32
Variance	0.012	0.172
Observations	5	5
Pooled Variance	0.092	
Hypothesized Mean Difference	0	
df	8	
t Stat	1.876629727	
P(T<=t) one-tail	0.048704552	
t Critical one-tail	1.859548038	
P(T<=t) two-tail	0.097409105	
t Critical two-tail	2.306004135	
Cannot reject null hypothesis		

Table 5

5. Control Test Vs. Fitness-Proportional Test on Dataset 2

   a. After performing an F-Test and being able to assume equal variances, the T-Test then delivered bad news. The T-Test failed to reject the null hypothesis so there is no statistical difference between the control and fitness-proportional tests on dataset 2. The statistics can be found in Table 6.

F-Test Two-Sample for Variances		
	Control2	*FitnessProportional2*
Mean	0.56	0.44
Variance	0.108	0.108
Observations	5	5
df	4	4
F	1	
P(F<=f) one-tail	0.5	
F Critical one-tail	6.388232909	
Assume Equal Variances by step 1		

t-Test: Two-Sample Assuming Equal Variances		
	Control2	*FitnessProportional2*
Mean	0.56	0.44
Variance	0.108	0.108
Observations	5	5
Pooled Variance	0.108	
Hypothesized Mean Difference	0	
df	8	
t Stat	0.577350269	
P(T<=t) one-tail	0.289792	
t Critical one-tail	1.859548038	
P(T<=t) two-tail	0.579584	
t Critical two-tail	2.306004135	
Cannot reject null hypothesis		

Table 6

6. 3-Restart Test Vs. Fitness-Proportional Test on Dataset 2

    a. After performing an F-Test and assuming unequal variances, the T-Test revealed that there is not a statistically significant difference between the Fitness-Proportional and 3-Restart tests. The statistics can be found in Table 7.

F-Test Two-Sample for Variances		
	FitnessProportional2	*3Restart2*
Mean	0.68	0.32
Variance	0.032	0.152
Observations	5	5
df	4	4
F	0.210526316	
P(F<=f) one-tail	0.08021698	
F Critical one-tail	0.156537812	
Assume Unequal Variances by step 2		

t-Test: Two-Sample Assuming Unequal Variances		
	FitnessProportional2	*3Restart2*
Mean	0.68	0.32
Variance	0.032	0.152
Observations	5	5
Hypothesized Mean Difference	0	
df	6	
t Stat	1.876629727	
P(T<=t) one-tail	0.054833117	
t Critical one-tail	1.943180281	
P(T<=t) two-tail	0.109666233	
t Critical two-tail	2.446911851	
Cannot reject null hypothesis		

Table 7

Jordan Giacone
Jmgv27@mst.edu
COMP SCI 5401 FS2016 Assignment 1d

Bonus 1

Bonus #1 requires the addition of a third objective to the multi-objective EA. This third objective is maximizing the number of decisions made by minisat. By increasing the number of objectives for the EA to optimize, the pareto hierarchy grows in physical dimension. Imagine a three-way optimization problem as cubes in space. Cubes that engulf others completely are said to dominate them, and this third dimension significantly decreases the chances that an individual dominates another. The largest pareto level seen during this whole assignment was level 4 because it is just statistically more difficult for three variables to dominate three others than it is for two variables to dominate each other. These flatter pareto hierarchies lead to more solutions being considered equivalent because they share a pareto level, leading to more genetic diversity than very tall pareto fronts.