Sujay Jakka

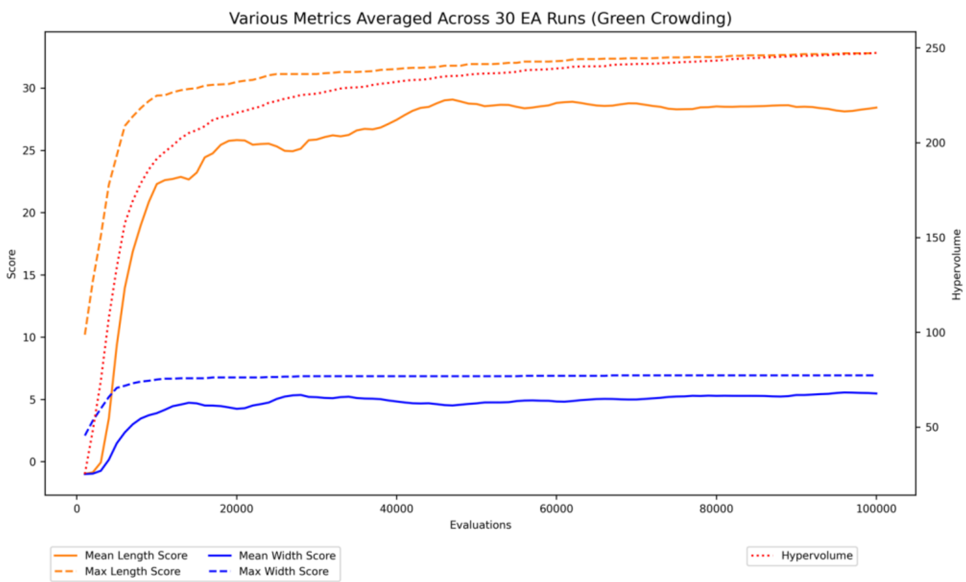
Svj0007@auburn.edu

COMP 5660 Fall 2024 Assignment 1d

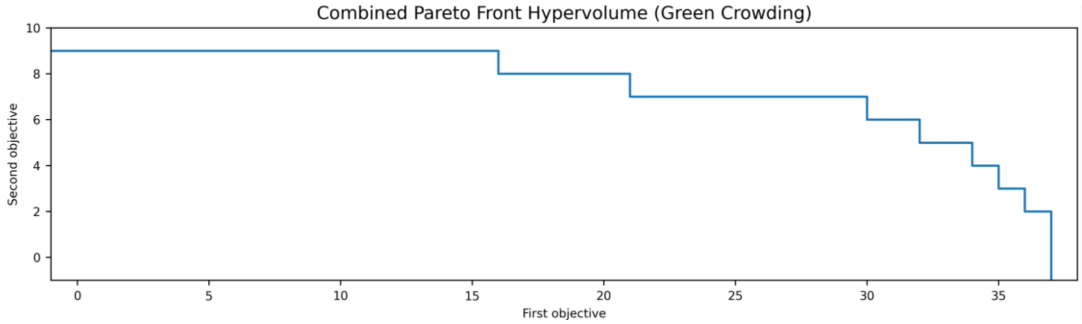
20 October 2024

Assignment 1d Report

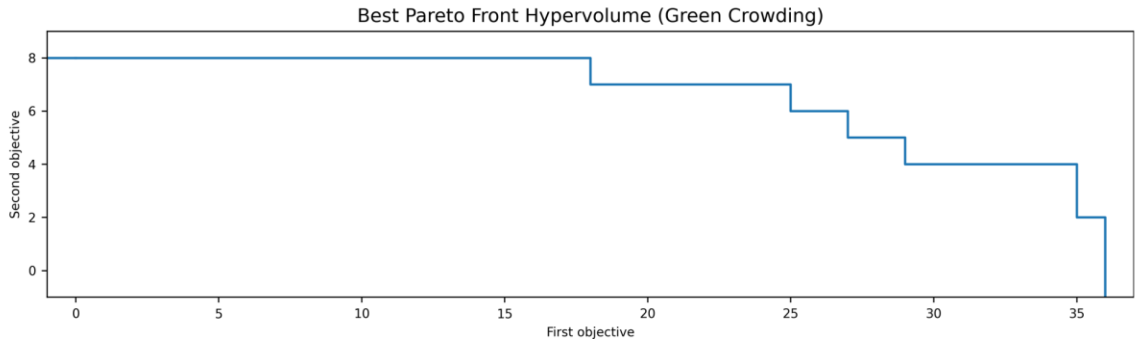
I will first introduce the results of the green deliverable with crowding. The plot below represents my evals-vs-metrics plots displaying the average mean length score, max length score, mean width score, max width score, and hypervolume score across runs per generation.



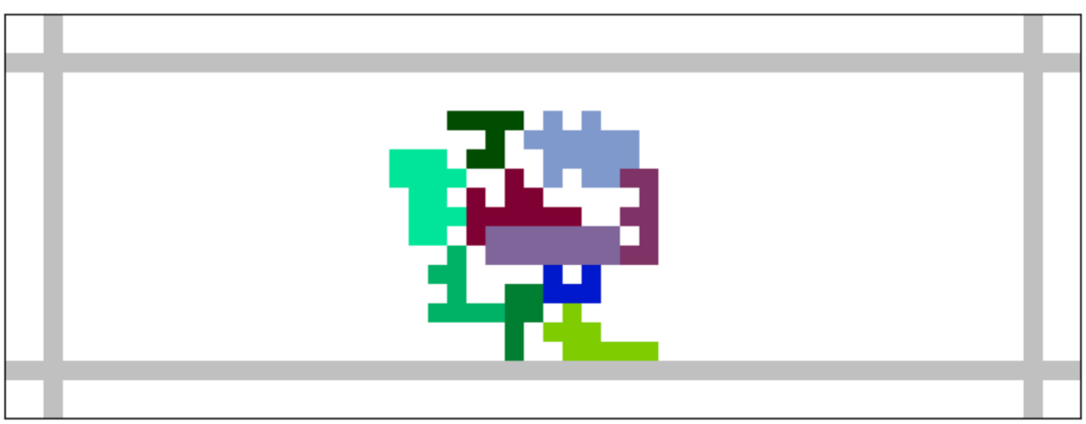
Judging from the plot, it looks like there needs to be more selection pressure as the mean length and mean width did not converge to the max length and max width. This indicates that the MOEA did not do a good job of exploitation. Furthermore, the image below is of my combined Pareto front, which are the solutions that are nondominated across the entire experiment.



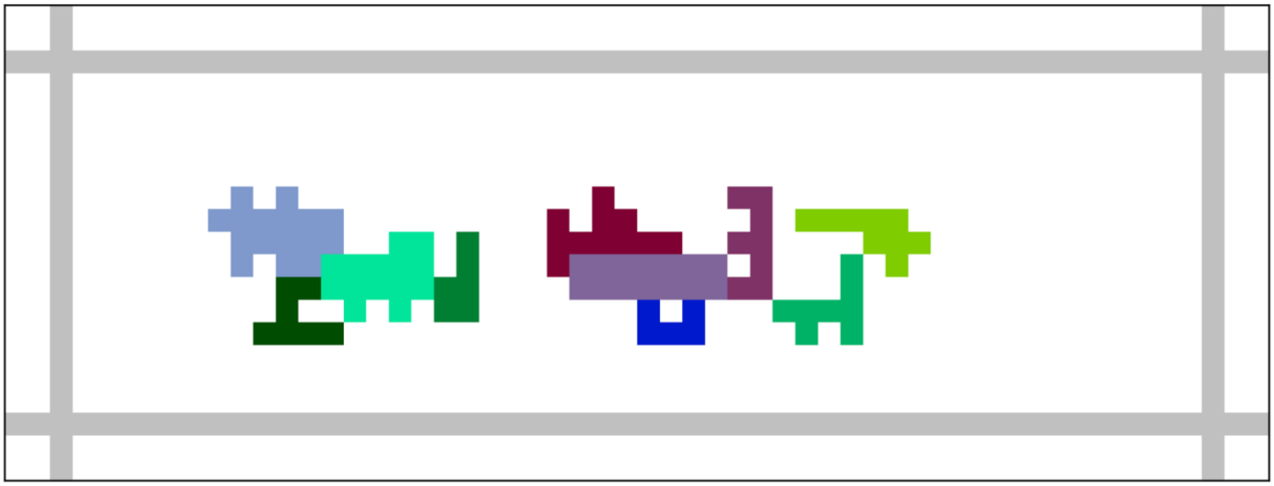
Additionally, plot below is the Pareto Front with the largest hypervolume across 30 runs.



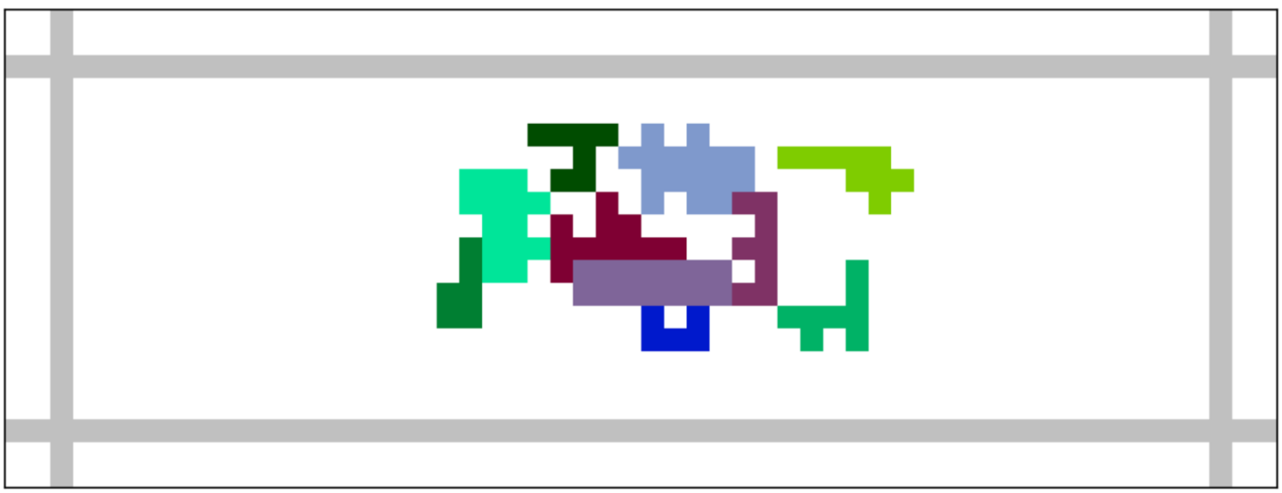
It seems that the Combined Pareto Front has more solutions on its Pareto Front compared to the Best Pareto Front seen across 30 runs. The images below represent the solutions with the highest length objective, highest width objective, and one solution in the middle of the best Pareto Front seen across 30 runs respectively.



(Solution with the highest length objective)



(Solution with the highest width objective)



(Random Solution on the Best Pareto Front)

Honestly there is not much to say comparing these three images besides the fact that is clearly evident that the solution with the highest length objective used a lot less length compared to the other two solutions that are shown from the best Pareto Front. This can also be said with the second solution that has the highest width objective. Furthermore, it is apparent that the random solution from the Best Pareto Front does not use as much length as the solution with the best width objective, but also does not use as much as width as the solution with the best length objective. Lastly the parameters I used for the green deliverable with crowding are shown below.

```
[ea]
mu = 1000
num_children = 1000
mutation_rate = 0.15
parent_selection = k_tournament_with_replacement
survival_selection = k_tournament_without_replacement

[recombination_kwargs]
method = uniform
prob_selecting_parent_1_gene = 0.5

[parent_selection_kwargs]
k = 30

[survival_selection_kwargs]
k = 50

[fitness_kwargs]
crowding = True
```

[mutation_kwargs]

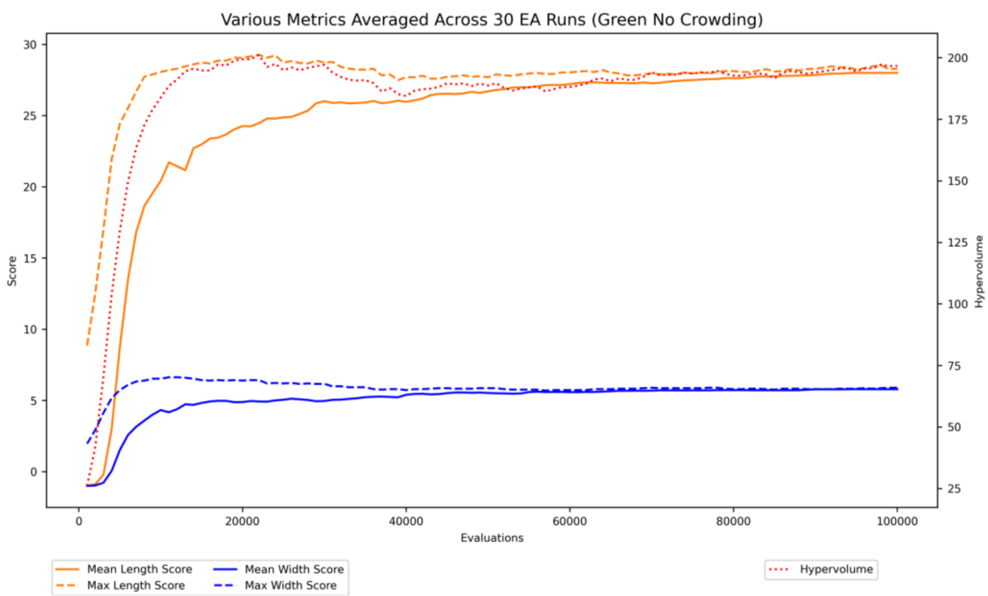
prob_creep_mutation = .15

mean_of_change_dist = 0

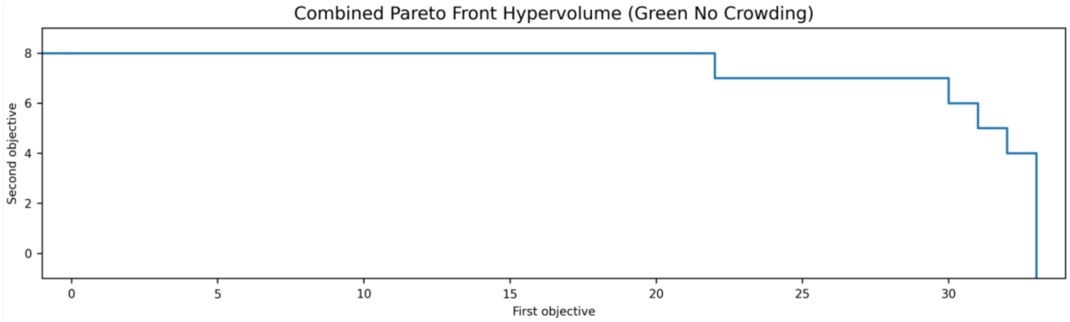
std_of_change_dist = 3

prob_random_reset = .15

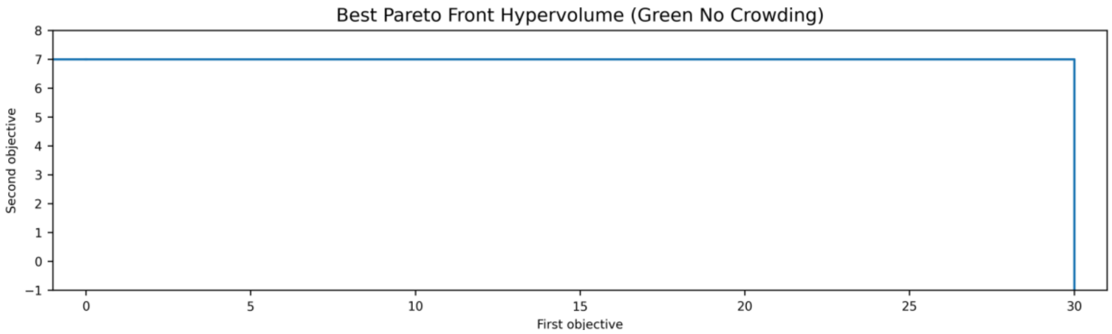
I will now introduce the results of the green deliverable without crowding. Similar to the green deliverable with crowding, the plot below represents my evals-vs-metrics plots displaying the average mean length score, max length score, mean width score, max width score, and hypervolume score across runs per generation.



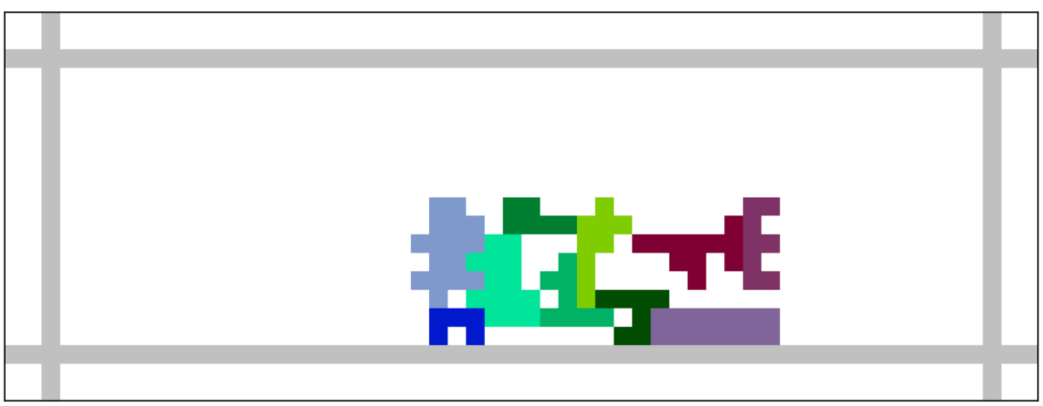
Judging from the plot, it looks like the mean width score converged with the max width score pretty early on. This could be because there was too much selection pressure. However, the mean length reached close to convergence toward the end of the life of the MOEA which displays a good balance between exploration and exploitation. What is concerning is that the MOEA seems to lose the optimal Pareto Front near 20000 evaluations.



Additionally, plot below is the Pareto Front with the largest hypervolume across 30 runs.



The plot above shows that there was only one solution in the Pareto Front. This is undesirable as we would like a diverse set of solutions across the Pareto Front displaying the tradeoff of the two conflicting objectives. The image below represents the solution with the highest length objective, highest width objective, and the solution in the middle of the best Pareto Front seen across 30 runs respectively.



There is only one image because my best Pareto Front only had one solution. This displays the importance of crowding because crowding promotes diversity across the Pareto Front. With crowding, we could have seen the MOEA exploit not just the crowded regions in the current Pareto Front but also the sparse regions. This would also increase the chances of survival of these solutions' genetic material that reside in the sparse areas of the Pareto Front as they will have a higher fitness than its crowded peers on the Pareto Front, allowing them to beat out these solutions if they happen to be in the same K-tournament.

Lastly the parameters I used for the green deliverable without crowding are shown below. They are identical to the parameters I used with crowding.

```
--------------------------------------------------------------------------------
[ea]
mu = 1000
num_children = 1000
mutation_rate = 0.15
parent_selection = k_tournament_with_replacement
survival_selection = k_tournament_without_replacement

[recombination_kwargs]
method = uniform
prob_selecting_parent_1_gene = 0.5

[parent_selection_kwargs]
k = 30
```

[survival_selection_kwargs]

k = 50

[fitness_kwargs]

crowding = False

[mutation_kwargs]

prob_creep_mutation = .15

mean_of_change_dist = 0

std_of_change_dist = 3

prob_random_reset = .15

Lastly, the mean of hypervolume of the sample of 30 runs(sample size 30) of the 1d MOEA with crowding was 247.43, while the mean of the hypervolume of the sample of 30 runs(sample size 30) of my 1d MOEA without crowding was 196.73. The standard deviation of the crowding MOEA and the without crowding MOEA was 16.25 and 27.62 respectively. I performed a two sample independent t-test where the null hypothesis was that the two algorithms are equally effective at solving the problem, and the alternative hypothesis is that the two algorithms are not equally effective at solving the problem instance. I performed the t-test with a significance level(alpha) of 0.05 and a degrees of freedom of 30. I got a p-value of 2.66582e-11, which is significantly smaller than the significance level of 0.05. We can then reject the null hypothesis, and claim that the two algorithms have statistically significant differences in performance. Because it is highly unlikely that both sample distributions have the same population mean and that the MOEA with crowding sample has a higher mean than the MOEA sample without crowding, we can conclude that the MOEA with crowding performed better.