

COMP 6660 Fall 2022 Assignment 2b

Joshua Boyd
jcb0216@auburn.edu

November 13, 2022

1 GPac Player Controller Evolution (Green)

The experiment was run 30 times for 2000 evaluations per run. Number of nodes in solution trees was used as the parsimony pressure metric. The config file used for the experiment can be found at `configs/green2b_config.txt`. Best Solution Raw Fitness Score per Run are given in table 1 on page 3. The best solution found was given a fitness score of **166**. This fitness was hit in the **12th run**. A plot showing the average of the generational average fitness and average generational best fitness per run is shown in figure 1 on page 4. The tree for the best solution can be found at `/solutions/bestGreenRunTree.txt`. The worldfile for the best run can be found at `/worldFiles/overallGreenBest.txt`.

Run Number	Best Raw Solution Fitness Score
1	70
2	110
3	76
4	63
5	76
6	73
7	83
8	83
9	90
10	93
11	50
12	166
13	133
14	56
15	60
16	160
17	73
18	73
19	100
20	100
21	46
22	43
23	46
24	66
25	126
26	96
27	93
28	53
29	50
30	83

Table 1: GPac Green Best Solution Per Run

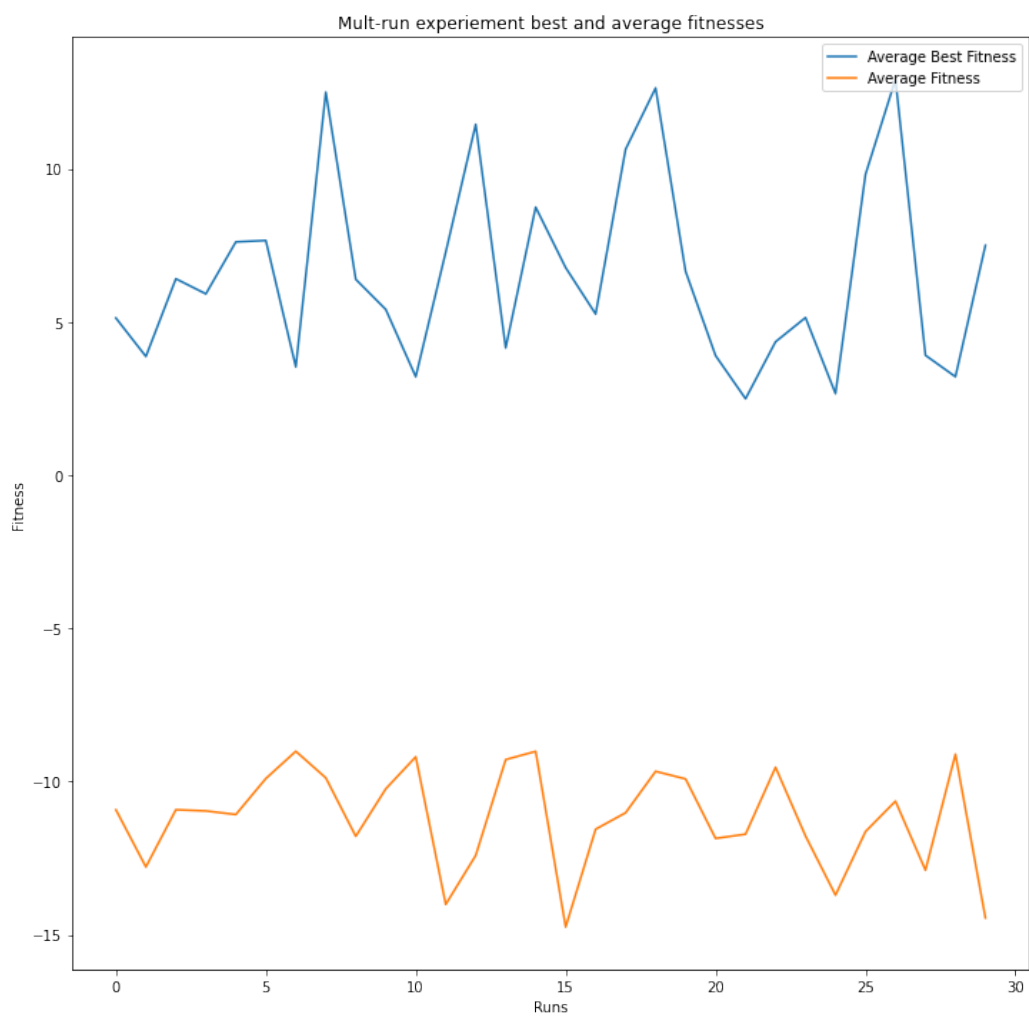


Figure 1: GPac Green Generational Average Fitness & Best Fitness per Run

1.1 GPac Player Controller Evolution Statistical Analysis (Green)

In this analysis, the best fitness per run data from the GPac 2a assignment was compared against the best fitness per run data from the GPac 2b assignment green experiment. The distribution of the data is not known to be normal and the sample size is greater than 29. Therefore, a two sample F-Test was performed for variances. The table showing the results of this test can be found in figure 2 on page 6. The chosen value for α was 0.025. The sample size for each sample was 30. The variance for sample 1 was 63.8620689655172 and the variance for sample 2 was 1063.97816091954. The calculated test statistic F was valued at 0.0600219734870541. The nearest critical value to F(F Critical one-tail) was 0.475964774310031. According to the test results, the value of F is smaller than one and is less than F Critical one-tail. This means that the test has rejected the null hypothesis of equal variances, and determined that the two populations have significantly unequal variances.

With the variances of the data sets being significantly unequal, a Two-tailed two-sample t-test assuming unequal variances was used next for comparison. The table showing the results of this test can be found in figure 3 on page 6. The chosen value for α was 0.05. The sample size for each sample was 30. The sample mean for sample 1 was 82 and the sample mean for sample 2 was 79.4333333333333. The sample variance for sample 1 was 63.8620689655172 and the sample variance for sample 2 was 1063.97816091954. The calculated test statistic t was found to be 0.418606827030786. The upper critical value (t Critical two-tail) was 2.0369333434601.

According to the t Test results, t Stat was greater than 0 and less than t Critical two-tail. Therefore, the t-test concludes the null hypothesis, and finds that there was no significant difference detected between the two experiments.

F-Test Two-Sample for Variances		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	82	79.43333333
Variance	63.86206897	1063.978161
Observations	30	30
df	29	29
F	0.060021973	
P(F<=f) one-tail	1.56917E-11	
F Critical one-tail	0.475964774	

Figure 2: GPac Green F Test

t-Test: Two-Sample Assuming Unequal Variances		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	82	79.43333333
Variance	63.86206897	1063.978161
Observations	30	30
Hypothesized Mean Difference	0	
df	32	
t Stat	0.418606827	
P(T<=t) one-tail	0.339149909	
t Critical one-tail	1.693888748	
P(T<=t) two-tail	0.678299817	
t Critical two-tail	2.036933343	

Figure 3: GPac Green T Test

1.2 GPac Player Controller Evolution Informal Analysis (Green)

Compared to the random search in assignment 2a, the run from 2b green seems to not make as many back and forth non-purposeful movements. However, the non-purposeful movements are not gone completely. Still, there are many times where it seems the pac player has targeted a particular pill and is decidedly headed towards it. Additionally, the 2b pac player was actually able to complete the game by collecting all of the pills. In the random search results from 2a, the pac player did not complete the game and was killed by a ghost.

2 Parsimony Investigation (Yellow)

The experiment was run 30 times for 2000 evaluations per run. Since the number of nodes in a solution tree was used as the green experiment parsimony pressure metric, tree depth was used in this experiment. The config file used for the experiment can be found at `configs/yellow2b_config.txt`. Best Solution Raw Fitness Score per Run are given in table 2 on page 9. The best solution found was given a fitness score of **194**. This fitness was hit in the **18th run**. A plot showing the average of the generational average fitness and average generational best fitness per run is shown in figure 4 on page 10. The tree for the best solution can be found at `/solutions/bestYellowDepthParsimonyRunTree.txt`. The worldfile for the best run can be found at `/worldFiles/overallBestYellowDepthParsimony.txt`.

Run Number	Best Raw Solution Fitness Score
1	133
2	165
3	120
4	93
5	113
6	180
7	103
8	143
9	123
10	100
11	164
12	177
13	119
14	93
15	106
16	93
17	106
18	194
19	96
20	130
21	146
22	167
23	110
24	113
25	160
26	123
27	153
28	120
29	126
30	145

Table 2: Parsimony Investigation Yellow Best Solution Per Run

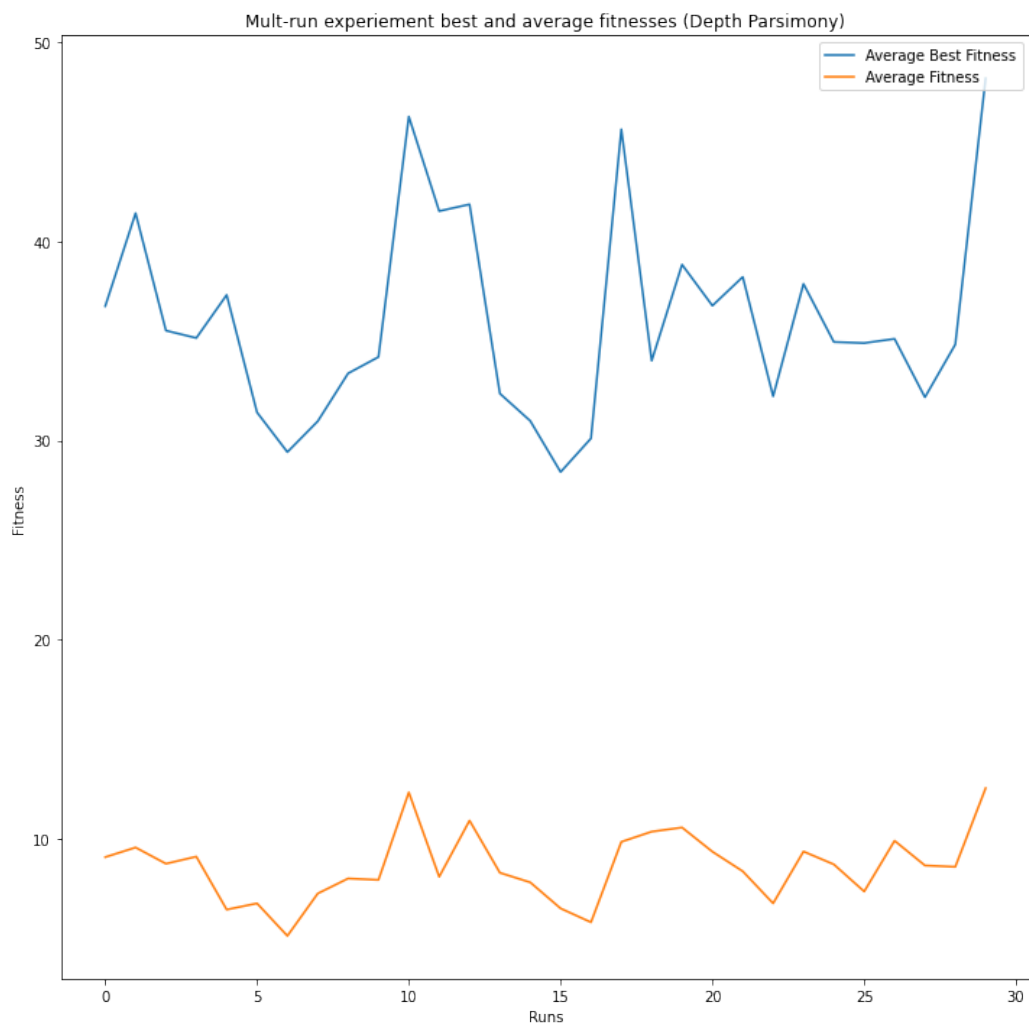


Figure 4: GPac Parsimony Investigation Generational Average Fitness & Best Fitness per Run

2.1 GPac Parsimony Investigation Statistical Analysis (Yellow)

In this analysis, the best fitness per run data from the GPac 2b green experiment, using node count as parsimony pressure, was compared against the best fitness per run data from the GPac 2b yellow experiment, using tree depth as parsimony pressure. The distribution of the data is not known to be normal and the sample size is greater than 29. Therefore, a two sample F-Test was performed for variances. The table showing the results of this test can be found in figure 5 on page 12. The chosen value for α was 0.025. The sample size for each sample was 30. The variance for sample 1 was 1063.97816091954 and the variance for sample 2 was 830.809195402299. The calculated test statistic F was valued at 1.2806528464148. The nearest critical value to F(F Critical one-tail) was 2.10099581728421. According to the test results, the value of F is greater than one and is less than F Critical one-tail. Therefore, there is no significant difference in the two variances.

With the variances of the data sets having no significant difference, a Two-tailed two-sample t-test assuming equal variances was used next for comparison. The table showing the results of this test can be found in figure 6 on page 12. The chosen value for α was 0.05. The sample size for each sample was 30. The sample mean for sample 1 was 79.4333333333333 and the sample mean for sample 2 was 130.466666666667. The sample variance for sample 1 was 1063.97816091954 and the sample variance for sample 2 was 830.809195402299. The calculated test statistic t was found to be -6.42146802394381. The upper critical value (t Critical two-tail) was 2.00171748414524.

According to the t Test results, t Stat was less than 0 and less than -t Critical two-tail. Therefore, the t-test found that the two experiments produced a significantly different mean fitness. The experiment with the highest sample mean can then be assumed to produce a significantly higher mean fitness than the other. This would be experiment 2 with depth parsimony pressure.

F-Test Two-Sample for Variances		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	79.43333333	130.4666667
Variance	1063.978161	830.8091954
Observations	30	30
df	29	29
F	1.280652846	
P(F<=f) one-tail	0.254782127	
F Critical one-tail	2.100995817	

Figure 5: GPac Parsimony Investigation (Yellow) F Test

t-Test: Two-Sample Assuming Equal Variances		
	<i>Variable 1</i>	<i>Variable 2</i>
Mean	79.43333333	130.4666667
Variance	1063.978161	830.8091954
Observations	30	30
Pooled Variance	947.3936782	
Hypothesized Mean Difference	0	
df	58	
t Stat	-6.421468024	
P(T<=t) one-tail	1.36289E-08	
t Critical one-tail	1.671552762	
P(T<=t) two-tail	2.72578E-08	
t Critical two-tail	2.001717484	

Figure 6: GPac Parsimony Investigation (Yellow) T Test

2.2 GPac Parsimony Investigation Informal Analysis (Yellow)

Compared to the green 2b experiment, the number of non-purposful moves was even less. In this experiment result, it seemed that the pac player showed a bit more planning, chaining decisions of where to move and what to go after without bouncing around aimlessly beforehand. Aside from completing the game, the pac player was able to collect more fruit than the green player, which lead to a higher score. The high score was despite the green (node count parsimony pressure) player completing the game faster than this yellow (depth parsimony pressure) agent.

3 Multiple Identical Pac-Man Controllers (Red One)

The experiment was run 30 times for 2000 evaluations per run. The config file used for the experiment can be found at `configs/red1_2b_config.txt`. Best Solution Raw Fitness Score per Run are given in table 3 on page 15. The best solution found was given a fitness score of **161**. This fitness was hit in the **23rd run**. A plot showing the average of the generational average fitness and average generational best fitness per run is shown in figure 7 on page 16. The tree for the best solution can be found at `/solutions/bestRed1Tree.txt`. The worldfile for the best run can be found at `/worldFiles/overallBestRed1.txt`.

Run Number	Best Raw Solution Fitness Score
1	80
2	96
3	86
4	96
5	103
6	106
7	96
8	96
9	100
10	90
11	116
12	103
13	96
14	96
15	146
16	103
17	96
18	110
19	106
20	126
21	100
22	113
23	161
24	113
25	106
26	60
27	123
28	113
29	100
30	106

Table 3: Multiple Identical Pac-Man Controllers Red One Best Solution Per Run

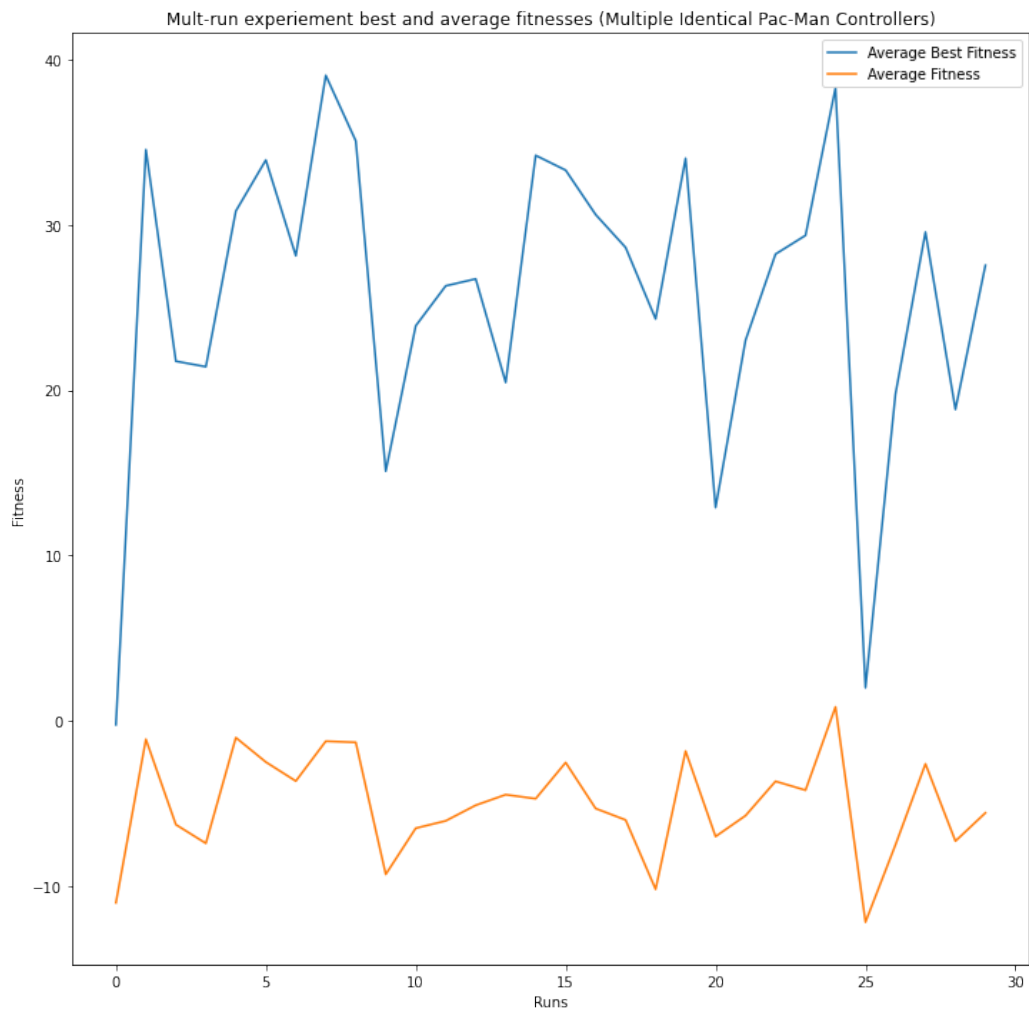


Figure 7: GPac Multiple Identical Pac-Man Controllers Generational Average Fitness & Best Fitness per Run

3.1 GPac Multiple Identical Pac-Man Controllers Informal Analysis (Red One)

My expectations for this experiment was that the pac player agents would complete the game the fastest of all previous experiment, due to there being more players to simultaneously go after the pills. However, the opposite happened, with this experiment finishing the game in the slowest time. I believe this is due to all of the pac players using the same controller and focusing on the same targets. If one pac player gets a pill that the other pac players were focused on then they will have to reevaluate and find a new target. This wastes time results in a longer time to complete the game. Additionally, this experiment had the lowest score compared to yellow and green experiments.

4 Multiple Identical Ghost Controllers (Red Three)

The experiment was run 30 times for 2000 evaluations per run. The config file used for the experiment can be found at `configs/red3_2b_config.txt`. Best Solution Raw Fitness Score per Run are given in table 4 on page 19. The best solution found was given a fitness score of **1353**. This fitness was hit in the **14th run**. A plot showing the average of the generational average fitness and average generational best fitness per run is shown in figure 8 on page 20. The tree for the best solution can be found at `/solutions/bestRed3Tree.txt`. The worldfile for the best run can be found at `/worldFiles/overallBestRed3.txt`.

Run Number	Best Raw Solution Fitness Score
1	1317
2	1331
3	1336
4	1333
5	1318
6	1326
7	1327
8	1314
9	1328
10	1293
11	1330
12	1323
13	1338
14	1353
15	1335
16	1326
17	1308
18	1332
19	1335
20	1293
21	1324
22	1339
23	1335
24	1327
25	1334
26	1318
27	1321
28	1319
29	1241
30	1301

Table 4: Multiple Identical Ghost Controllers Red Three Best Solution Per Run

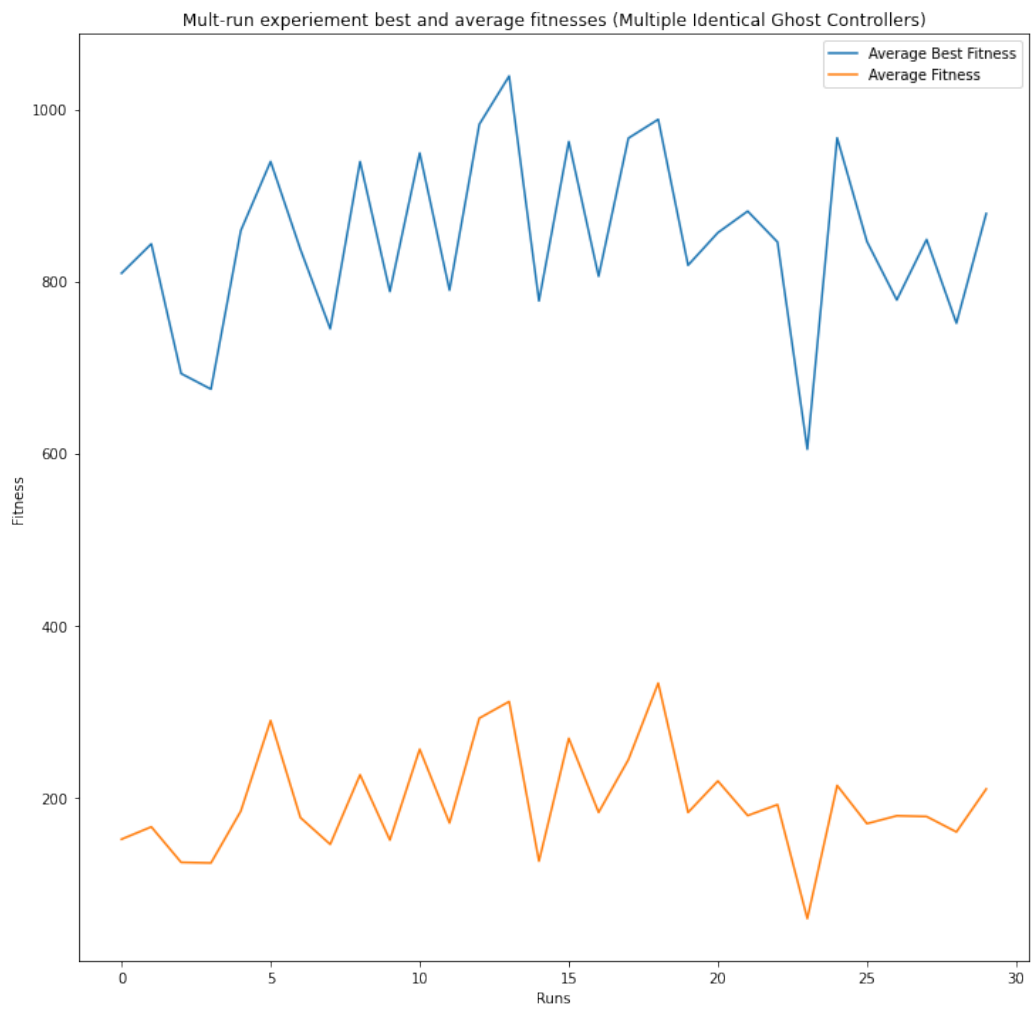


Figure 8: GPac Multiple Identical Ghost Controllers Generational Average Fitness & Best Fitness per Run

4.1 GPac Multiple Identical Ghost Controllers Informal Analysis (Red Three)

I was very surprised how effective the ghost controller was. It seemed to go straight for the pac player. I am guessing that this was because evolution decided that the only thing the ghosts needed to be concerned about was the location of the pac player. This is reflected in the solution tree for the ghost controller which only has the W and M tokens in it.