Appendix for Optimised fitness functions for automated runtime improvement of software

ANONYMOUS AUTHOR(S)

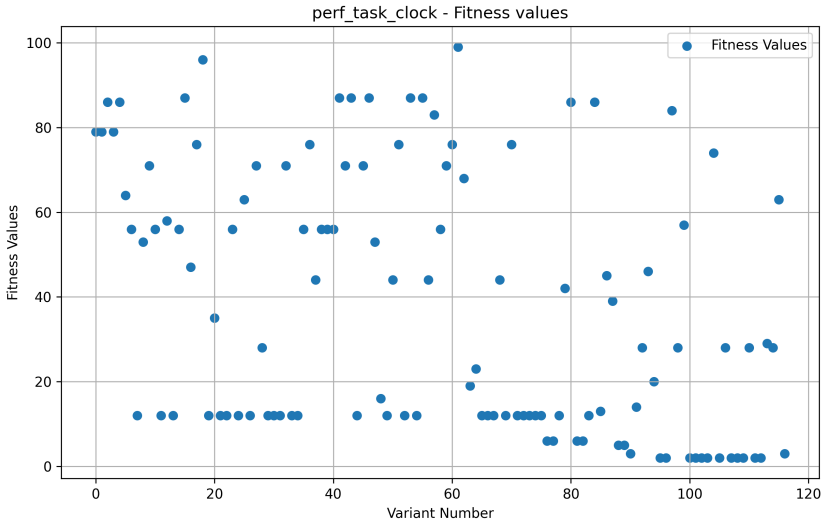
ACM Reference Format:
Anonymous Author(s). 2024. Appendix for Optimised fitness functions for automated runtime improvement of software. 1, 1 (September 2024), 5 pages. https://doi.org/10.1145/nnnnnnn.nnnnnnn

Appendix A

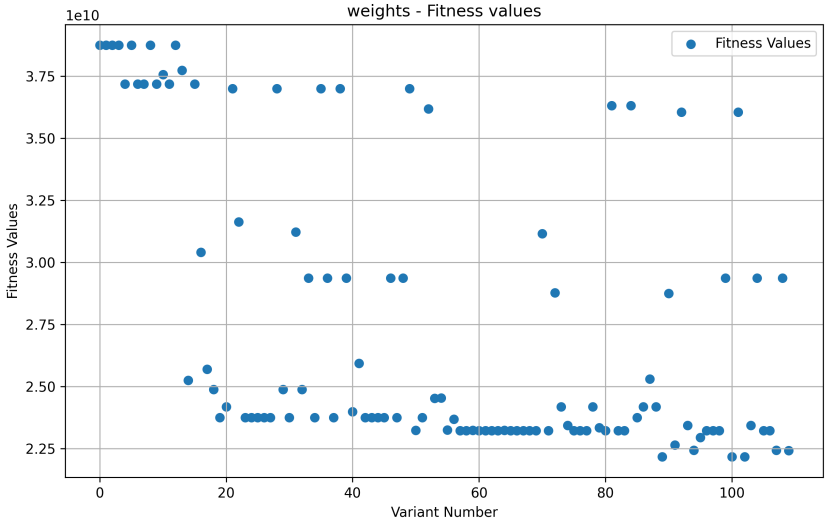
A visual representation of the variants examined and their fitness values in 2 different scenarios. One with the `task_clock` measure and a retry setting of 1 and the other with the `weights` measure and a retry of 2. We can observe how different measures and retry setting affect the exploration of the search space.

ACM XXXX-XXXX/2024/9-ART
https://doi.org/10.1145/nnnnnnn.nnnnnnn

Fig. 1. Comparative analysis of the search space for the LPG benchmark under different measures and retries.



(a) Search space exploration with task_clock and retry=1



(b) Search space exploration with ‘weights‘ measure and retry=2

Appendix B

In this section we present the results of running the same exact experiments as seen in Section ??, with the only difference being that MAGPIE was invoked with the Genetic Programming approach as the search strategy.

Table 1. RQ2: Mean Median Execution Times, Median Ranks, and Number of Variants for Each Measure

Measure	Mean MET(s)	Median Rank	NoV
Original	9.65	13	-
Energy	6.21	11	257
L1_dcache_loads	5.99	7	397
branch_misses	5.75	5	431
branches	6.22	7	393
cache_misses	7.02	8	419
cache_references	4.64	5	451
cpu_clock	8.85	14	376
cycles	6.15	4	426
faults	7.28	12	381
instructions	6.48	7	384
task_clock	9.87	14	387
perf_time	5.66	7	409
Time	5.17	4	414
Weights	6.96	12	275

MET: Median Execution Time
NoV: Number of Variants examined

Table 2. RQ3: Statistical Analysis of Retry Rankings Across All Benchmarks

Retries	Median Rank	Mean Rank	STD	IQR
1	2.36	2.47	0.27	0.39
2	2.93	2.95	0.40	0.50
3	3.00	3.07	0.46	0.71
4	3.29	3.18	0.41	0.39
5	3.36	3.33	0.41	0.32

Table 3. Recommended Retry Numbers for Each Measure

Measure	Mean RN	Median RN	IQR	STD
Energy	3.0	3.0	1.0	1.10
L1_dcache_loads	2.64	2.0	2.5	1.57
branch_misses	2.64	2.0	2.0	1.50
branches	2.18	2.0	1.5	1.54
cache_misses	3.0	3.0	4.0	1.79
cache_references	2.18	2.0	1.5	1.33
cpu_clock	3.45	3.0	2.5	1.44
cycles	2.45	3.0	2.0	1.37
faults	2.09	2.0	1.5	1.22
instructions	2.55	2.0	3.0	1.51
task_clock	3.64	4.0	1.5	1.29
perf_time	2.82	3.0	2.5	1.54
Time	2.73	2.0	3.5	1.74
Weights	2.27	2.0	2.0	1.35

RN: Retry Number

Table 4. Percentage Decrease in Variants Explored by Number of Retries

Retries	Mean Decr.%	Median Decr. %	STD	IQR
1	0.0	0.0	0.0	0.0
2	31.87	34.84	24.64	38.11
3	52.86	60.07	22.30	25.65
4	64.73	69.93	19.58	21.00
5	72.40	75.58	15.14	15.63

Table 5. Median and Mean MAD for Each Measure

Measure	Med. MAD	Mean MAD	STD	IQR
Energy	0.085	0.097	0.230	0.234
L1_dcache_loads	0.140	0.129	0.224	0.229
branch_misses	0.126	0.121	0.190	0.176
branches	0.183	0.162	0.246	0.245
cache_misses	0.087	0.077	0.159	0.152
cache_refs	0.186	0.175	0.251	0.280
cpu_clock	0.461	0.533	0.150	0.123
cycles	0.170	0.145	0.210	0.191
faults	0.022	0.025	0.036	0.040
instructions	0.142	0.139	0.260	0.257
task_clock	0.428	0.498	0.165	0.131
perf_time	0.122	0.150	0.248	0.224
Time	0.210	0.169	0.252	0.257
Weights	0.153	0.159	0.340	0.350

Table 6. Median and Mean MAD for Each Retry Number

Retries	Med. MAD	Mean MAD	STD	IQR
1	0.261	0.311	0.206	0.249
2	0.197	0.242	0.189	0.192
3	0.176	0.223	0.194	0.207
4	0.146	0.188	0.168	0.207
5	0.141	0.193	0.169	0.242

Table 7. Analysis of Retry Parameters on Search Efficiency

Retries	TD	ADPS	POD
1	8.0	0.0500	0.0123
2	7.0	0.0563	0.0178
3	6.0	0.0521	0.0251
4	6.0	0.0497	0.0303
5	5.5	0.0642	0.0365

TD: Total decreases
ADPS: Average Decrease Percentage per Step
POD: Proportion of Decreases

Table 8. Performance of Each Measure in the Search Process

Measure	TD	ADPS	POD
Energy	6.0	0.0291	0.0383
L1_dcache_loads	8.0	0.0343	0.0253
branch_misses	8.0	0.0486	0.0215
branches	8.0	0.0429	0.0254
cache_misses	6.0	0.0324	0.0232
cache_references	7.0	0.0571	0.0209
cpu_clock	4.0	0.2337	0.0149
cycles	8.0	0.0497	0.0231
faults	7.0	0.0129	0.0255
instructions	7.0	0.0384	0.0296
task_clock	5.0	0.1791	0.0153
perf_time	7.0	0.0439	0.0254
Time	7.0	0.0564	0.0229
Weights	5.0	0.0792	0.0329

TD: Total decreases
ADPS: Average Decrease Percentage per Step
POD: Proportion of Decreases