

Instructions for replication of results

The Dana code required to replicate the results seen in the paper is included in the folder “Genetic Improvement Code”.

The code requires Dana version 204 to run. This is due to improvements to the Dana language which have occurred since the code was written, which render it non-functional on later versions.

While this code has already been compiled it is possible it will need to be recompiled for your Dana Virtual Machine version, especially if using a non-Windows OS. To do so, navigate to the folder (Genetic Improvement Code) in a command terminal, and use the command “dnc .”. This will recompile all files. If you have already run the code you may need to delete all files contained within the “code/” folder, as it will have the genetic generated code, which may be uncompileable.

Altering settings

Settings are read from the file settings.txt, located within the main folder. Altering values in this file does not require re-compiling existing code. The file is set to default values, with a synthesis phase (Referred to in file as “explore length”) of 6, the default.

Extracting and processing results

When run, the code will log its output to the folder it is executing in. The primary results file will be of the form “Fitness_***.log”. The precise name is programmatically generated, to allow multiple runs to be performed and the results collected in bulk. The first value in the file name (“_xN_”) is the synthesis length used for that particular run. This is to allow synthesis phase lengths to be easily compared, by altering the settings file, running batches of experiments, then copying the results to a second folder for analysis.

To process these files, a short Java program has been included, in the folder “javaDataProcessingScripts”. This program reads the fitnesses logs, and has a number of methods available which can return different values of interest. It is this script which was used to extract and process the results used in this paper. The script reads any set of fitness files placed within the “RawGIFolder”.

Changing Benchmarks

Altering the benchmarks used by the program requires replacement of files within the “testing_files” folder, found within the “GeneticImprovementCode” folder. The two types of file needing replacement are the “CacheHandlerUnitTest.txt” file, and the sequential files of the form “CacheBenchmarks*.txt”.

By default the code attempts to read for Benchmark files, named CacheBenchmarks1.txt to CacheBenchmarks4.txt. A set of alternative Benchmarks have been provided, which can be swapped into this set, by replacing an existing file with the new choice.

Alternatively, the number of benchmarks read can be altered in the code. This requires altering of lines 25, 26 and 29 in the Dana file “CacheHandlerBenchmarkTesting.dn”. The digit “4” should be altered to the number of benchmark files desired. This will then require re-compiling the Dana source code.