SBSE report

Part 0:

Mitch:

Tony: I completed Part 1 and Part 2

Part 1:

GIN is a genetic algorithm that is designed to\ optimize and improve the efficient of the program passed as an argument. The fitness function used by this genetic algorithm is to optimize program’s speed. The aim of this experiment is to apply the genetic algorithm GIN to four functions SortBubbleDouble, SortBubbleLoops, SortInsertion and SortCocktail and evaluate their performance upon improvements by the tool. We will compile and run 15 iterations of each of the four sorting functions. The best execution speed from each iteration will be recorded and graphed, this will be completed for all four sorting functions. An evaluation of effectiveness and practicality of the patches will be supplied. Below are the graphs of results obtained from the best execution for all four functions.

The methodology used to obtained the results was done by first downloading the gin-master.tgz folder. Gradle 3.3 and JDK 1.8 was downloaded according to the README file. The build file was compiled by running the command “gradle build”. When built, to obtain the best execution speed the following command was ran inside gin-master folder “java -jar build/gin.jar examples/locoGP/<name of program>, where name of program is the name of functions with filename .java at the end. The program will then compile and applies numerous patches to the function to try and optimize it. Upon completion, the program will state where the best patch was found, best execution time and achieved speedup. Fifteen iterations of the described method were executed for all four sorting functions, during which the best patches applied and corresponding execution times were recorded. The best execution speeds of each function were graphed separately, the produced results can be seen below.

Fig 1. Optimized functions best execution speeds

Fig 2. Optimized Sort Cocktail best execution speed

From a surface level we can see that the genetic algorithm produces low variance results, having similar best execution speeds. However, the flaw of the GIN genetic algorithm lies in the details. The best patches of each algorithm were recorded and observed. Upon observation we can see that some best patches applied made the algorithm efficient in terms of speed but were not programmable correct. We can see this in four separate cases, the first case is the function Sort Bubble Double the initial nested loop was kept the same while the second nested loop did not contain the if statement inside. The code still compiled and was more efficient than the original algorithm but the empty nested for loop was redundant and computed nothing. In a similar fashion for the Sort Bubble Loops, the first loop in the three nested loop was moved outside the nested loop, creating a redundant for loop. The algorithm still compiled and passed all test cases and provided the best execution speed. In case of the Sort Insertion there was a redundant assignment a[j] = index. Finally, in the Sort Cocktail function the best patch copied the swapped = false line and subsequent for loop and added later into algorithm creating redundant code. The reason for this is due to the fitness function. The fitness function of the Genetic Algorithm GIN only accounts for the execution speed of the optimized function it is not trying to optimized how the correct the code is and as such as long as it provides the best execution speed the correctness of the code is irrelevant.

Part 2:

The benchmark program chosen for this part was the Sort Cocktail function. The benchmark provides patches more than one edit long. By running this benchmark and recording the patches we can analysis the effect of each patch on multiple runs.

A similar methodology was

Part 3:

Part 4:

Part 5:

From my understanding, we need to create a benchmark program that is inefficient. We then run GIN against this benchmark program to improves it efficiency. We can alter a couple of lines in the program to make it efficient and see if GIN can pick up on it. In the report we write where the inefficiency comes from and run against GIN and presents the results we obtain.