Project Work Book: Programming Java for Longer Battery Life

Supervisor: Dr Alexander Brownlee

Anna Frances Rasburn - 2411187

2018/2019

|  |  |  |  |
| --- | --- | --- | --- |
| **Date** | **Brief Summary** | **Detailed Summary** | **To Do** |
| Thursday 19th April | Initial Meeting with Dr Brownlee | * Discussion of the project * We talked about the different ways we could do this project * Defined project definition, requirements, structure and timeline | * Read Resources about the topic * Start the preliminary report |
| Friday 18th May | Dissertation Proposal Submitted | * Dr Brownlee signed the document * Submitted to reception | * Do some reading and research over the summer |
| Sunday 15th July | Reading: Search-Based Energy Optimization of Some Ubiquitous Algorithms | * One of the first papers about applying search-based software engineering to minimize the energy consumption of JVM running programs. * Opacitor – a tool for measuring the energy consumption of JVM programs using a bytecode level model of energy cost. * Advantages of Opacitor:  1. Deterministic – ‘[something] in which no randomness is involved in the development of future states of the [thing]’. A determistic model should produce the same output from a given input or start. <https://en.wikipedia.org/wiki/Deterministic_system> 2. Unaffected by the rest of the computational environment 3. Can detect small changes in execution.  * Energy reduction up to 70% * 3 considerations for automatically reducing computational energy:  1. Tuning software to particular distributions of data 2. Trading off energy use against functional properties 3. Handling internal dependencies.  * Criticism: only experimented on a Samsung galaxy s3 * Criticism: only looks at CPU and does not consider display, WIFI and GPS * ‘Dependent of the total number of CPU cycles, there can be a large difference in the energy consumed by different opcodes’ * Opacitor ‘can detect small changes in execution profile, down to opcode level’ * Three SBSE approaches for optimisation  1. Genetic Programming – to obtain better pivot functions for quicksort 2. Hyperparameter – search over the efficiency trade-off of MLPs 3. Genetic Algorithms – OOP for guava and Apache commons collection classes  * Program behaviour: ‘Design by Contract’ – a software correctness methodology. Uses pre and post conditions to a document the change in state caused by a piece of a program. * Software development now considers small scale computing * CPU is often used as a proxy for energy use. These previous approaches are inaccurate as it omits CPU idle states. * A lot of research has measured power consumption directly from hardware * SBSE algorithm often repeats measurements, extends run-time and reduces confidence in the results. * Opacitor comparison: LLVM toolchain – measures energy consumption due to each bytecode at runtime. * Code perforation – finding parts of the code that can be skipped, such as loop iterations. * Data structures like lists, sets and maps tend to be less CPU heavy * Fitness function is the energy consumption measured by Opacitor. Fitness is determined deterministically * All non-deterministic features of the JVM are disabled like JIT and GC * Templar – a software framework for customising algorithms via the generative technique of the template method. * Multi-Layer Perceptron – a modification of the standard linear perceptron with ‘hidden’ layers of neurons with nonlinear activation functions. It can distinguish data that are not linearly separable and are often applied to classification tasks like image recognition. * Energy is reduced to around 20% of the worst-case solutions for all data sets. * SBSE allows the developer to easily explore the trade-off between functional and non-functional (accuracy and energy) * Open Closed Principle – both framework implementations and their client programmers can be assured that their code can interoperate without unexpected behaviour * Design by Contract is the explicit codification of the LSP via preconditions, postconditions and invariants. * Benefits to catch any situation where dependencies exist between variation points. * Criticism: programs failing the unit test were rejected by the search as invalid. * An SBSE approach is able to accommodate these interactions where they exist without the developer needing to specify where they are, achieving better performance without additional developer effort. * Criticism: No consideration for different versions of JVM | |
| Saturday 16th June | Reading: Reducing Energy Consumption using Genetic Improvement | * Big Idea:GI seeks to improve software’s non-functional properties by treating program as if it was genetic material * GI can successfully be used to reduce energy consumption by 25%. * The total ICT infrastructure generated 1.9% of global carbon dioxide consumption in 2010 * There is a big disconnect between a developer’s source code and the energy it needs when it is complied. * Petke et al – made MiniSAT which is a popular Boolean satisfiability solver can be optimised for execution time, so it is 17% faster than CIT * Fitness of a candidate solution is calculated my measuring the total energy consumed across all the tests in the training set. * A fitness greater than 1 indicates a solution that consumes less energy. * A fitness more than 1 indicates a solution that consumes more energy. * Crossover is done by selected one parent based on fitness and the other one is random. * Crossover continues till after selection, the population has doubled. * Intel power gadget tracks the OS a hardware and estimates the energy consumption of second and higher intel core processer. Doesn’t include energy consumed in main memory or in I/O tasks. * Combinational Interaction Testing (CIT) – Black box sampling technique to test highly configurable software. * Ensemble computation – the study of an NP complete variant of the Boolean circuit problem where one must find a smaller circuit that satisfies requirements. * GI is yet to applied to bigger hardware and applications. | |
| August | Reading: Specialising Guava’s Cache to Reduce Energy Consumption | * Big Idea: the use of Genetic Algorithms on parameter tuning with the Google Guava’s cache Library. * The creation of the tool Opacitor which is a tool that measures the energy consumed of a program. * Jalen – it uses time and CPU utilisation to calculate energy consumption. * Due to the numerous factors of parameter tuning regarding optimization, they used a SBSE to tune and specialise parameters because of the limited battery life of mobiles. * ECJ Toolkit is the most popular toolkit for evolutionary computation so it may be very useful in my project. * Previous work has included Hoffmann et al ‘s ‘Power Dial’ which is ‘a system for dynamically modifying trade-offs between accuracy in computation and use of system resources during load peaks’ * Power Dial reduces the amount of computing infrastructure instead of really focusing on energy consumption. * Wu et al used Genetic Algorithm’s for the tuning od deep and shallow parameters using dlmalloc memory allocator for execution time and memory consumption. * Mutation of source code – each variation point has a range of potential substitution values were selected till appropriate. In the template version of CacheBuilder the substitution values can be directly inserted. * Measuring Energy Consumption – Opacitor’s time to shine. It is designed to stop multiple runs and algorithms can be compared. * Using OpenJDK, Opacitor counts the number of times each java opcode is executed. Opacitor calculates the number of joules used. * Features of the JVM have to be disabled for example JIT and Garbage Collection. The initial memory allocated to the JVM is increased to the point, so the GC is not used. Once the evolution is done then then these features are re-enabled. * Opacitor can run concurrency with other programs so the fitness can be calculated in parallel on a multi-core system. * When JIT was enabled using the JVM’s default memory allocation, the results found that the noise was reduced. * Future work is to investigate the trade-off between energy/time and memory. | |
| Key Terms:   * Hyperparameter Optimization – the problem of choosing a set of optimal hyperparameters for a learning algorithm. * Hyperparameter – a parameter whose values are set before the learning process begins * Automated Parameter Tuning – automatic process of tweaking software parameters until optimal configurations. * Variation points – the declaration of default values for parameters. | |
| August | Reading: Object-oriented Genetic Improvement for Improved Energy Consumption in Google Guava | * Big Idea: using a metaheuristic search to find a semantically equivalent version of the immutable map that uses less energy. * Comparing the metaheuristic search to an independent exhaustive search at each variation point so the metaheuristic has a better performance. * Previous work – Sahin et al – measuring the effect of 6 refactoring techniques on Java programs and finding out that the energy usage is highly end-application dependent. * Semantics preserving are done through behavioural equivalence which is important to object-orientated principles. * Mantos et al’s ‘SEEDS framework’ which alternates subtypes container classes are substituted into bytecode, so it minimises power consumption. * DRAW IMPLEMEMTATION DIAGRAM here * Use a map instead of a collection as they have a more efficient Big O (o(1)). * Mutating source code – each variation point we need to determine the interface of the created object. * Opacitor is unaffected by anything else executing so it can be executed simultaneously with other programs without difficulty. Object-Oriented Genetic Improvement – a technique by which non-functional properties so time/energy consumption may be optimised. | |
| Key Terms:   * Metaheuristic – a high-level problem independent algorithmic framework that provides a set of guidelines to develop evolutionary algorithms. * Genetic Program – technique whereby computer program is encoded as a set of genes that are then modified using an evolutionary algorithm. * Grammatical Evolution – to find an executable program that will achieve a good fitness value for the given objective function. * Liskov Substitution Principle – if something expects a parent class but gets a child, the child should conform to what is expected of the parent. | |
| Wednesday 12th September | Initial welcoming meeting back with Supervisor | * High level introduction to generic algorithms and deep tuning * Talk about how to begin project and different ways we can go * IDEA: Does the version of java affect energy/execution time * We talked that we want to get the fundamentals right first and if we do go to Android Application ‘that will be the icing of the cake’ * Agreed that would take the initial route of * Switch circuit? – why if(a+b) is different to if(b+a). this is because if a is false then we won’t even check if b is false as we just give up. If we swap and b is true, then we still see a. It is useful in using methods. * Looked at Gin and Jalen * Looked whether to go creating my own GA or going via gin. * I’m going make my own GA and unit tests to test with. As gin is open source so it wants to change approach, I can do this. * REMEMBER TO EDIT CONFIGURATIONS TO RUN THE APPLICATION * LOOK AT THE GIN IF STATEMENTS FOR ‘SUCCESS’, FAILURE AS STARTING POINT * Gin has never been used with Mac before, so this may be interesting | * Try writing a piece of code that emulates the test framework of the gin code. * Research and read about Genetic Algorithms * Due Date: 19th September |
| 13th September | Reading: Artificial Intelligence – A Guide to Intelligent Systems  Page.222-232 | * Genetic algorithm – ‘a class of stochastic search algorithms based on biological evolution’ * Generation – The iterative process run of a GA * Mitchell – typical number of a simple GA can range from 50 to 500. * Run – the entire set of a generation, at the end of a run we find out of more highly fit chromosome. * DO DIAGRAM OF GENETIC ALGORITHM HERE * The fitness of a population may remain stable for a number of generations before a superior chromosome appears. | |
| Friday 14th September | Got GinFork code to work | * Met up with Sandy to ask for help with running the ginfork application * Problem seemed to be with the run configurations. Initially I used these configurations: gin.LocalSearch examples/simple/Triangle.java -cp=examples/simple/-className=Triangle. The run configurations I use now: examples/simple/Triangle.java -className=Triangle -cp=examples/simple/ -steps=10 |  |
| Saturday 15th and Sunday 16th | Implemented a basic genetic algorithm | * I created my own genetic algorithm simulation with 3 basic classes of individual, population and GA. * It runs okay, it generates a new population and calculates the fitness and the fitness improves as the generation count increases. * DO A UML DIAGRAM OF THIS |  |
| Friday 14th – Wednesday 19th September | Play around with GinFork | * Played with the. Firstly began by trying to get it to break and see how it runs. * Added breakpoints and ran for large number of steps to see differences. |  |
| Wednesday 19th September | Second meeting with Supervisor | * Presented to Supervisor my basic genetic algorithm. We discussed and improved it. * Asked questions I had about Gin code * Discussed how my GA and gin code is different as my GA has a population and gin doesn’t technically. * Gin has a ‘hill walker’ structure. * Discussed the interim report and ethics checklist. | * Ethics checklist * Interim report beginning * Play with Gin code * Make a GIN like edit myself * Get the jcodec code to work |
| Friday 21st September | Ethics Checklist submitted | * Marked no to everything as it is a research project therefore has no end users |  |
| Saturday 22nd September | Beginning of Interim Report | * Started writing the introduction and state of the art sections so I can show Sandy. * Added references too. |  |
| Sunday 23rd September | Started implementing my own edit in a similar vain to GIN | * Started with 2 edits that worked similar to gin. |  |
| Wednesday 26th September | Third Meeting with Supervisor | * Discussed how to run jcodec with sandy as I realised, I had been doing it wrong. Also developed a better understanding of how jcodec works. * Discussed the interim report and which examples we can talk about in state of the art * Discussed the ways we can go with the project, whether to use Opacitor at the end or introduce it from the beginning if the project. * Opacitor does not have binary code that is implemented to run on a MAC OS so that may be a challenge as the project progresses but will work out a solution. * Discussed the test runner areas of GIN code. This is the really important part of GIN. * Going to make my own code using test runner which may be quite challenging but going to try and get as much as I can done in a week. * Need to look at the article ‘Genetic Improvement of Software: A comprehensive survey’ <- look at types of changes and transformations. | * Run gin targeting jcodec * Extracting testrunner * Continues with interim report |
| Thursday 27th September |  | * Copied over the test runner classes as a starting point to make my own wrapper * Continued with interim report |  |
| Friday 28th September |  | * Continued with Interim report * Made an initial project with classes from the original gin code |  |
| Saturday 29th September |  | * Tried implementing a project similar to gin. * Continued with eduts |  |
| Sunday 30th September |  | * Realised the code I was writing over the weekend was practically a copy of gin so made a new project structure and started implementing a much simpler version. Not yet completed. * Changes I wish to make. I think the gin code is good but at the moment it ignores tests that aren’t unit, so I wish to make a better ground that is extensible and can be changed so more stuff can be added later on. * Also wish to simplify the structure by trying to see the edit as an individual and a patch as a population so in a similar vain to the genetic algorithm I made in the beginning. * Also changed to use maven and spring framework as wanted an excuse to practice these more. * Also realised the bad structure of my project and git repository so deleted it and started again with a better structure * I think I have jcodec to work will find out. Have a few questions though |  |
| Monday 1st October |  | * Met up with Sandy to ask some quick questions * He should me how to get jcodec to work * I played with some of the other jcodec arguments * Started implementing my own edit and patch in a new way. |  |
| Tuesday 2nd October |  | * Continued with my implementation. * Tested with the different jcodec methods. |  |
| Wednesday 3rd October | Fourth meeting with supervisor | * Asked my questions about jcodec and gin * Showed my current code * Discussed the poster and interim report and how to do this so I can get a first draft done soon. * Tested with the different jcodec methods. * Also got the refactored version of gin code. | * First interim report draft sends to sandy ASAP * Continue coding * Begin the Poster? |
| Friday 5th October | Implementation continues. | * Did a small amount of implementation into software. It can currently read in a class as inspired by GIN. However I used Java 8’s paths which I learnt about on my placement. * Also read about hot spotting as an alternative to java reflection. Thought it could be an alternative for GIN. Hoping to implement it in my own implementation. |  |
| Sunday 7th October | Sent Sandy my first draft of interim report | * Completed the first draft of interim report and sent it to sandy. |  |
| Wednesday 10th October | Attended WIT conference in Glasgow | * Attended WIT conference 2018 in SEC Glasgow. * Got to attend a lecture about AI and its future. This is directly related to my project but did |  |
| Thursday 11th October | Fifth Meeting with Supervisor | * Discussed changes needed for the interim report. * Clarified understanding of issues brought up in interim report * Received the Opacitor code * Discussed hotspots as Sandy is unaware and I know I need to improve my knowledge of them. |  |
| Friday 12th October – Saturday 13th October | Changed interim report and started poster | * Changed the interim report so it included improving my referencing and adding more state-in-the-art descriptions. * I included some details about green software engineering and genetic improvement. * Also made a first draft for the poster * Tried running the Opacitor code on my MAC and received an error. |  |
| Sunday 14th October | Send Sandy second draft | * Sent Sandy a second draft of the interim report. * Also looked over the Opacitor code base |  |
| Monday 15th October | Received feedback on second draft | * I received feedback on my second draft from sandy so included a slight restructure of the report and improved the referencing. |  |
| Wednesday 17th October | Opacitor code and implementation | * Borrowed a friend’s Linux machine to run Opacitor and it worked. * Continued implementation of software execution software. Tried to look at making a simple gin that currently only does one type of edit (delete) without the genetic algorithm. |  |
| Friday 19th October | Interim report submitted | * Finally submitted interim report |  |
| Saturday 20th October | Worked on implementation | * After finishing creating a delete line edit in a slightly different way to GIN code, I made some unit tests for this code. * I then did a move, swap and insert line edits using Test Driven Development and in a similar way to the other line edits * Changed edits so they are more functional interfaces rather than do everything in source file |  |
| Sunday 21st October |  |
| Monday 22nd October |  | * Set up a remote router on Linux machine so can ‘sort of’ run Opacitor on a mac. * Started a remove statement edit. |  |
| Tuesday 23rd October |  | * Made a simple hot swapping example. Need to consider whether to use this as it uses the JVM in a slightly different way. |  |
| Wednesday 24th October | Meeting with Sandy | * Opacitor – goalDirection.minimisng – we want to use less energy * Opcode.put(opcode, number of joules) * Bytecode histogram – the Opacitor is a histogram of joules to bytecode * Energy consumption can change over time doing the same instruction as it does it in different ways * Simple Jalen – looks at joules and super simple. These don’t use the fancy JVM so can run on mac. | * POSTER * Look at SB test * Generate 100 versions of altered code * System.time.realsea – time the programs |
| Thursday 25th October |  | * Received Opacitor code * Added to my gin code |  |
| Tuesday 30th October |  | * Looked through that Opacitor code * Added to my current code |  |
| Wednesday 31st October |  | * Made some new edits so it included blocks rather than just lines * Looked at having a multiple class edit |  |
| Thursday 1st November |  | * Started Patch * Made a combination of edits |  |
| Friday 2nd November | Code Work | * Realised problems with my lines so changed * Did a new edit that changes the operators in if statements | * Make own compilation unit * Make a patch (stop edits for right get a full of start to end program) * Work out a way to compile new formed code. * Made a random search * Save different versions to a file * RUN OPACITOR |
| Meeting with Sandy | * Presented my code with a particular focus on edits * Opacitor may not pick up changes of operations (but may improve runtime) * Consider using javapaser to swap blocks and block edits rather than currently just using for loops. Example: SourceFile Line 335 * Principle Opacitor\_sb.jar file should be in the build path and just run anywhere * 90% of the work in optimisation is defining the problem * Integrate the current GIN test runner into my system * Also look at how GIN compiles code – current system is time consuming * DO NOT try to integrate test runner and Opacitor at the same time.   Things to consider/write about in final thing   * Random deleting of things can be dangerous when regarding threads and deleting/altering * Android studio java methods/activities have correlating XML files for the setup so we need to consider how do you do this in the GIN code. |
| Saturday 3rd November |  | * Changed Move Block edits so it uses java parser to get a block rather than current system. |  |
| Sunday 4th November |  | * Continued with patch and realised I’d need to change an element in all previous edits to get patch to work so started doing that * Ran Opacitor test2 on windows machine |  |
| Tuesday 6th November |  | * Made a compile test using InJavaMemoryCompiler |  |
| Wednesday 7th November | Meeting with Sandy | * Showed current code to sandy * Discussed that I need to make an end-to-end system. * Discussion of poster session |  |
| Thursday 8th November |  | * Modified my git repo so it now works yay! |  |
| Friday 9th November |  | * Realised a better way of doing a patch so scrapped some of previous work * Made an Edit Factory and Patch factory |  |
| Saturday 10th November |  | * Changed the formentioned factories so they are implemented with spring so they could be easier to import as dependencies * Made a random edit file |  |
| Sunday 11th November |  | * Finished the patch so it applies a bunch of edits to a source. * This new source is outputted so you can see what has changed * Throws some errors as some edits remove lines so it can swap lines that aren’t there. Need some validation to see whether lines are there etc. | * Add it so the output file compiles using in java compiler. |
| Monday 12th November |  | * Added a basic in memory java compiler so it can compile new code. * Have a few problems as it does not entirely compile – could be due to the way that I’ve made the source into a string * Also thought about how I could do reps and did tests to prove it |  |
| Wednesday 14th November | Meeting with Sandy | * In order to get a first do some more wider reading for the actually 10,000 words. * Need to do more experiments. Since I can use random I have a baseline * Need to plan what questions I want to be asking? * Way to fix my edits problem: gin used a map instead of a linked list. Increases the predictability . maybe a tree map? * Change the insert line edit so it inserts earlier on * Why I used spring: * Have previous experience with spring * Personally found it easier to insert edits and make the code more flexible * Need to get a baseline and sanity check * Look at line 75 * Use the test runner from gin | * Get it to compile * Get it to repeat compile * Measure time and energy * Get Opacitor to run on mac * Running junit tests? |
| Poster Session | * Things to look into/Think about * Unwrapping (maybe unmapping) loops * Edit idea: to look at recursive loops * Do the loops unwire the bytecode? * Can you measure clock cycles as a measure of energy? * Compressing strings before pushing to console can have an effect on execution time |  |
| Work | * Configured the compile so it can run the hello world example |  |
| Thursday 15th November |  | * Inserted a system.currenttime to measure time. |  |
| Saturday 17th November |  | * Added the GIN test runner to my code * Needs a lot of modification to work in the new end to end system * Started that modification. * Modifications as I made a patch differently * Also looked at possibility of making my own with hot swapping – might be a over holiday task |  |
| Sunday 18th November |  | * Added Opacitor to my code * Not quite got running on Mac so need to see Sandy about this. |  |
| Tuesday 20th November |  | * Put all current work code onto the windows machine. * Had to change structure of code to get to work – need to fix this * Continues with test runner |  |
| Wednesday 21st November | Meeting with Sandy | * Got Opacitor to work on the windows machine after a lot of trial and error * Discussed performing experiments on loops e.g. iterative vs recursive sort algorithms * Discussed poster session. * Need to sort out the structure of the project on the windows machine | * Decide on a couple of started experiments * Continue with test runner * Play with Opacitor |
| Friday 23rd November |  | * Wrote a word document outlining an experiment about a random patch * I think I have finished altering test runner so it should fit in with my code. Need to add this to the ApplicationMain so it compiles -> Test Runner -> Opacitor. | * Make it go from compiled code to test runner * Go from compiled code to opacitor |
| Monday 26th November |  | * Realised a problem with block edits so fixed remove and move (hopefully) * Need to modify the if statement edit so it works better |  |
| Tuesday 27th November |  | * Tried adding the test runner and Opacitor to the application main but currently having a few problems to connect. Need to ask sandy about this | * Get test runner and Opacitor included |
| Wednesday 28th November |  | * Made 2 basic bubble sorts in a iterative and recursive. This is for looking at starting experiments about energy consumption when it comes to unwrapping loops. * Wrote an experiment for this * Tried putting this to Opacitor but I may have the Opacitor wrong or my understanding how to use it wrong. |  |
| Thursday 29th November | Meeting with sandy | * Test runner is having a class not found exception * Might not be getting the Junit library * Compiled triangle – not being found?   Bubble   * Generate 30 different arrays. Seed it. 1/3 random and 30 reps of algorithm,. 1/3 ascending order and 30 reps. 30 descending order and 30 reps. * This helps gets rid of (a majority of) noise * Keep experiments simple so change from 4 edits to 1 and increase over time. * Write what edit type and what is the edit. * WRITE OUT AS MUCH DATA AS YOU CAN. | * Write results to .csv file * Get testrunner to find the classes * Run with testrunner to find the classes |
| Friday 30th November |  | * I made a basic POJO which can be used to write the results to a .csv file. * Corrected errors * Got test runner to completely run after seeing sandy * Need to get Opacitor to run for the bubble sort |  |
| Saturday 1st December |  | * Ran some experiments and got it to print to the .csv file in a slightly better format * Realised need to change project structure to get Opacitor and full test runner on .csv * Tried adding the test runner and Opacitor in a (have to admit hacky way) to run the current code but having a problem with the “.compileToRawBytes” method and am struggling to understand why * Also test runner test is now failing and I have no idea why |  |
| Sunday 2nd December |  | * Removed the above test runner from that structure as I realised the main needs to be outside all of the modules to use them all. Think I need to change the structure so there are 4 modules: Opacitor, test-runner, patch, edits and one main module that uses them all – this will take time waiting till Christmas hols |  |
| Wednesday 5th December |  | * Worked out why the test runner wasn’t working as there was a problem in the anna path * Reformatted project structure so modules were more connected - this needs work and a new structure * Reprogrammed the bubble sorts so they actually worked, still not got them running so well in Opacitor – need to ask Sandy | * Bubble sort and Opacitor * Project structor * Test runner to file |
| Thursday 6th December |  | * Finally got the test runner working. * Write test runner results to file * Refigured results to a ResultsFileWriter and made it into a builder * Changed the remove line edits so it doesn’t ‘remove’ in the same sense just changes that particular line to null and then streams so it does not consider that line, * Need to do this with all the other edits to avoid an index out of bounds exception – gin used a map to do this * Have an idea for the insert line edit, what if used the JAVADOCs to get a random line and insert that rather than the //this is a comment line. Could maybe use a GA to decide on the line. Could we have a GA write code? |  |
| Friday 7th December | Meeting with Sandy | * Needle in a haystack problem = the insert line edit were it just inserts a comment * GI – the code that is already there should be enough/changed * GI – insert lines increases the chance of it not working * Ideal endpoint of GI: code making code * GA should be easy. * The fitness is already there as (compile, test runner, Opacitor) * The making children is already there with making new versions of code * Opacitor.updateCode <- the tuple consists of (replacement, package, class) * Once test runner passes then it goes to Opacitor. * For the bubble sort you can feed in a unsorted list as the parameter * Opacitor runs with a main method so Triangle needs one * Need to change javac on mac to 8 so Opacitor is happy * 10k words: * Literacy report is 20% get done by end of the year * Need more iterations than in interim report. Need to look at more alternatives than just what’s directly related * Recommend: 30-40 citations * Demo: * Print off a copy of the project log * Few minutes of intro * Explain how new variation (edits) and fitness (compile, test runner and Opacitor) | * Add a main method to the Triangle class |
| Friday 7th December |  | * Added Opacitor to the main method after a lot of work. * It adds the measurement to the .csv file * The measurement is always 1179.0 though so I question how accurate this is. * Finished enough to demo will focus on exams for the next few days |  |
| Tuesday 11st December | Demo | * Demo went well (I think?), lasted around 10 minutes. * Got ideas of inserting break/return statements as possible edits |  |
| Wednesday 12nd December | Meeting with Sandy |  |  |
| A few days break for exams | | | |
| Monday 17th December | Final report | * I wrote the abstraction (one page) of the dissertation. * Needs improvement |  |
| Tuesday 18th December |  | * I restructured the project’s structure so there are multiple modules that take in each other as maven dependencies. * Now has a main method * Now has an algorithms module that can allow for Gas and local search |  |
| Wednesday 19th December |  | * Worked on the introduction of the final report needs improving |  |
| Thursday 20th December |  | * Worked more on introduction * Made a TOC for the state of the art section |  |
| Friday 21st December |  | * Worked on the state of the art section. Added new reference |  |
| Wednesday 26th December |  | * Fixed the swap block edit |  |
| Thursday 27th December |  | * Added a lot more papers to state of the art * One of the papers needed for my state of the art required emailing the authors in order to receive so this has been delayed completion. |  |
| Monday 31st December |  | * Read over state of art and hopefully fixed grammar errors * Still need to get that paper. |  |
| Wednesday 2nd January |  | * Received the paper from Lori Pollock * Added to the state-of-the-art and I read over and currently have 4000 words of the writing completed. |  |
| Friday 4th January |  | * Checked the diss for grammer and sentences and stuff. Aim to send these 4000 words to sandy by monday |  |
| Saturday 5th January |  | * Refactored and fixed edits. * Created 30 tests for the bubble sorts. Compared execution time, recursion takes up to 2753 milliseconds longer. Need to put this through Opacitor to get difference in energy |  |
| Sunday 6th January |  | * I broke my dissertation. The test runner has stopped working and I have no idea why. * Continued on the 10000 words, hope to send to sandy tonight or tomorrow. |  |
| Tuesday 8th January |  | * I conducted the experiments on the bubble sorts and have collected the data, I need to tidy up the csv files |  |
| Thursday 10th January |  | * Cleared up the bubble sort files and noticed a significant energy usage between iterative and recursive sorts |  |
| Saturday 12th January |  | * Dad finished reading over my diss, I then took his modifications and approved/disapproved of some and sent the modified version to Sandy |  |
| Sunday 13th January |  | * Asked Morgan Stanley pals for help with why test runner was not working. It turns out the maven dependencies were not correctly done so they helped me sort it in half hour so code is back to working now. | * Implement a GA * Implement HC * Need to discuss with sandy the measure of Fitness |
| Monday 14th January |  | * Created a NoEdit Edit which runs successfully * Added method names to GA <- Needs implementing * Have confusion about wasteCPU method, Not making any changes to opacitor readings <- Need to ask Sandy |  |
| Tuesday 15th January |  | * Added current code to the different sections of a GA class. It worked better than in random. Need help with the full implementation |  |
| Wednesday 16th January |  | * Added 2 small methods to make a ‘weighted fitness’ need to check with sandy that I am right. Doing it as a fitness score |  |
| Friday 18th January |  | Multiple Fitness   * 2 approaches:  1. Weighted 2. Energy and unit test are separate  * 1. Weights * Can do the number of unit tests failed added to the current energy and this equals the fitness score. * This number is wanted to be as low as possible * 2. Energy and unit test are separate * Up to interpretation but the code should be working before it is energy efficient/ optimisation. Can be interpreted vice versa * So you compare fitness score 1 (number of failed unit tests) first and then if the different versions have the same fitness score 1, you use fitness score 2 (energy) to work out which of the solutions is the most optimised. * Fitness can be fuzzy – are all the unit tests equal (maybe ignore for now – say it in writing that all the unit tests are equally weighted.   Writing   * Don’t go too far above or too few below the 15,000 words * Look at previous ‘research’ dissertations * Fitness stuff (currently in intro) could be moved to methodology * Have a section maybe about what the optimisation problem looks. Search for edits through the source code. Gas are good for this purpose because… * Refine the objectives so fewer words. Lose in objective 1 from “use of GI to java” * Objective 2, blatantly reference objective one where you have said ‘earlier’. And add the word ‘framework’ to the gin code. * Objective 4 and 5 remove the ‘acquire to apply’ * In achievements: can say I now understand yhe affects on energy and time code has and this knowledge can be applied to my future career * Give examples in GI section of state-of-the-art. Typical approach of moving lines of the code etc * Can’t cite paper which are not published yet but can use quotes. Use 5.4 operator set   of new paper. This has citations SO USE   * Jcodec goes into a different chapter * Tools used like opacitor maybe should be more in methodology * Methodology – my toolkit. How I made the toolkit. Include class diagrams and code snippets * The results from bubble sorts could be influence the research   Code   * 4 different measurements of opacitor (energy). Currently using the wrong one which only counts the number of lines. * Opacitor – can switch off GC method and see different results * Do bubble sort tests again with other opacitor measurements | * Re-preform the bubble sort experiments with Jalen * Continue with GA |
| Saturday 19th January |  | Writing   * Added references and in cite notations * Added the comments about the what used in gin from the paper sandy sent me   Code   * Having problems getting Jalen to work <- Need to see sandy about this * So I can focus on continuing with finishing a genetic algorithm, I am making my second unit fitness score to be time then once I have the GA finished, I can insert the energy measurement as a the second fitness |  |
| Sunday 20th January |  | * Got simple Jalen to work so going to use this for a while till sandy meeting. * Worked on bubble sort results with new opacitor |  |
| Tuesday 22nd January |  | * Refactored GA * Did bubble sort tests * Got opacitor with the bytecode\_histogram working. * Still need selection, crossover, mutation sections done |  |
| Wednesday 23rd January |  | * Refactored the GA more * Added more features to patch. Established a current of population of 2 patches. One patch is an individual * Researched crossover – need to talk to sandy asking about which route to go with this method * Did the selection method and currently have the different patches printing the scores? * I think also I need to check the test runner as they all seen to pass the unit tests when thinks like lines are removed and they don’t compile but still pass the tests? |  |
| Thursday 24th January |  | Writing   * I configured my referencing so hopefully its better * I added a couple of hundred words to potential further work   Code   * I need to ask Sandy about how to finish making the GA. |  |
| Friday 25th January | Meeting With Sandy | Writing   * Writing is not necessarily passive – so can use ‘I’ and royal ‘we’ pronouns. Need to use I in stuff you learned.   Code   * Selection is choosing a couple of patches from a population of them. Depends on the type of selection. RECOMMEND TORNASEL (population) return solution-pair * Look up elitism in a GA * My GA is a messy GA as patchs have variation in length of edits * Random number for number of edits and patches * Look in intro of GA book for stuff about messy GAs * Crossover – can be a random number of edits from parent1 combined with random number of edits from parent 2 * Might want off spring with the same length as parent. * Offspring could be you get the list of edits from one parent (the smaller edit list parent) and attach some of the edits from the other parent * RECOMMEND One POINT CROSSOVER * Mutation – are you going to delete or keep. * Compare GA to hill climbing * HC and GA have the same mutation, but GA has crossover * Extra Extra: Hybrid of HC and GA – possible do a GA to generate everything and then do a HC on the offspring | * Do the selection method * Do the mutation method – Look at equivalent in Gin’s local search * BUBBLE SORTS * FIND A LOOP TO GENERATE THE ARRAY * MAKE THE PROGRAM INTO A JAR FILE * HOT WIRE THE .JAR TO USE SANDYS 32 CORES. |
| Friday 25th January | Work I did | Code   * Implemented a selection method in GA * Implemented a mutation method in GA in the same way as the equivalent method as the gin code. Its odd so probs needs improving * Did some unit testing |  |
| Saturday 26th January |  | * Changed all bubble sort main so more efficient * Converted to a jar in order to but on unix machine * Had major problems getting it to work on unix * I installed UNIX and then sent the files over and converted it to a jar file on linux machine |  |
| Sunday 27th January |  | * Finally, I was able to run on UNIX machine * Found out that it can’t do 1 million size arrays * Ran a 100,000-size array * Made the main class concurrent on my computer need to move this over to unix machine as even with 100,000 size array it takes a long time to process. So if wish to do this better It needs to be more concurrent so takes less time. |  |
| Monday 28th January | Quick meeting with Sandy | * Explained concurrency problems with sandy * He said to do a bash loop that runs multiple threads |  |
| Tuesday 29th January |  | Bubble Sorts   * Ran 16 Iterative Bubble sort experiments – random experiments * Ran 10 Recursive Bubble sort experiments – random experiments * Had some problems with running a 10,000-size array on recursive   Code   * Played around with crossover other method   Reading   * I read some old dissertations and got a few more ideas how to structure the technical chapters   Writing   * I roughly wrote about 400 words |  |
| Wednesday 30th January |  | * I worked more on GA * Started implementing a Hill climbing |  |
| Thursday 31st January |  | * Changed the GA so it ideally copies the population * Need to ask Sandy about this * Worked on my writing and improved based on SB comments | * Redo the results class so it prints to .csv better |
| Friday 1st February | Meeting with Sandy | * Genetic programming – GA for program trees * Death penality function – Has to pass unit tests. Adding a bogus fitness score * Add flag to offspring to show which parents happen * Always use a seed in random * Use the same random for the whole program | * Add flags to offspring * Add an argument to random in selection * Improve resultwriter |
| Friday 1st February |  | * I ran all the ascending bubble sort tests. * Averaged the data * 250 words |  |
| Saturday 2nd February |  | * I ran all the descending bubble sort tests. * Averaged the data * 250 words |  |
| Sunday 3rd February |  | * Made graphs for all the data and added to dissertation. |  |
| Monday 4th February |  | * 250 words * I changed the seed * My mutation method needs redoing |  |
| Tuesday 5th February |  | * I changed the result writer so it prints to a csv file better. It is not pretty but currently works the best. * Added a off spring and neighbour class that extend patch * I changed the edit config as a test so that the remove line edit only removes the 9th line which has a direct call to the wasteCPU method so in theory that would affect energy but it has a compilation error which to me makes no sense. |  |
| Wednesday 6th February |  | * 250 words – currently on roughly 7500 words a day |  |
| Thursday 7th February |  | * 250 words |  |
| Friday 8th February | Meeting with Sandy | * Arrays.sort() – it uses a tin(tim)sort. For small arrays it uses an insertion sort and a merge sort for large arrays * Can mention that I used this arrays.sort for ascending arrays bubble sort data. This is theory should not influence memory * Graphs – Add error bars. Energy is statctic so on a live running machine so maybe some variation. Use standard deviation for the error bars. Check if the data is normalised – do a histogram * Try writing the source out after the remove line edit. Is this the same for Opacitor and in Memory Compiler * Run a new program 100 times and average | * Do a histogram for bubble sort data |
| Friday 8th February |  | * I broke my entire dissertation code. Life has ended |  |
| Saturday 9th February |  | * Fixed my dissertation, I realised I was using the wrong Opacitor * I fixed my remove line edit so it is more accurate and removes the line calling the delay method, this has no effect on memory |  |
| Sunday 10th February |  | * 250 words * Played with wasteCPU method, if you enter too large parameters i.e. 1000+ the code breaks. * Also all the Opacitor measurements are the exact same, need to ask sandy about this |  |
| Monday 11th February | Meeting with Dyslexia Supervisor | * Referencing – Having a few problems with missing names and has a lot of lowercase letters for some reason * Plan with Jackie: We are going to have a 2 hour session where we go over the first 5 chapters so they are completely finished. | * Finish introduction chapter * Finish chapter 4: creation of toolkit/methodology * Get sandy to read over chapter 4 * Once read over, book appointment with Jackie. |
| Tuesday 12th February |  | * Created a reversestring program to test toolkit with other programs * Added this project to the code, having a class not found error * Fixed that * Not have: Unable to find class file for test [Class: ReverseStringTest Method: test1 Timeout: 10000]. Ensure that test classes can be found on the supplied classpath * Changed it so there is a better structure and reverseString file works | * Change test runner so it takes the unit test |
| Wednesday 13th February |  | * I started creating a userInput object, so people can enter in their what files they want to optimise. * So far working |  |
| Thursday 14th February |  | * Worked on improving the userInput object <- some problems most of the time it compiles but sometimes it does not and I have no clue why however when it is hardcoded, it works every time, I have no clue why this has such an effect. |  |
| Friday 15th February |  | * Take a population that you have made and put it back through the Genetic Algorithm * Look at Jcodec as other source code files to optimise * Don’t worry about having positive results, just have some sort of results. | * Complete the GA loop so the new population goes back into the population * Experiment using JCodec code |
| Sunday 17th February |  | * Tried closing the GA loop <- seemed to do it in a scripty way, this obvs didn’t work so I need to redo this |  |
| Monday 18th February |  | * Added jcodec to the project path |  |
| Wednesday 20th February |  | * 250 words |  |
| Friday 22nd February |  | * Improve userinput by adding a builder * Closed the GA loop * Thinking of implement a selection interface |  |
| Saturday 23rd February |  | * Changed unit tests and added Mockito * Changed block statements |  |
| Sunday 24th February |  | * Need to changed the CSV file * Comment * Unit tests * Implemented a Hill climber |  |
| Monday 25th February |  | * Sandy gives back his feedback on the writing * Made some alterations |  |
| Tuesday 26th February |  | * Added my old CSCU9A2 assignment to optimise. Opacitor does not like jframe |  |
| Wednesday 27th February |  | * Added complex program to optimise |  |
| Thursday 28th February |  | * Modified writing |  |
| Friday 1st march | Meeting with Sandy | * Needle in a haystack problem – not getting the right remove line edit * Change the seed each time * Search space is a combo of all the things. 10 billion possible edits. Currently sampling 80 edits * Solution: more initiale population and more loops and more offspring and neighbours * Multiple ways to edit triangle * Try running eith only delete, move or swap line * In writing: add error bars to graphs. Add more stuff to chapter 5 and |  |
| Saturday 2nd March |  | * Developing a genetic bottleneck. * Made so the second population must be at least the same size as initial population * Ran triangle with 100 iterations and 100 population size and got results. |  |
| Sunday 3rd March |  | * To fix the genetic bottleneck, solutions maybe: increase the number of edits that a patch has from 1 to 2. Add another offspring that is random . Only send back one offspring * Change number of edits to 2 * Added a new form of crossover that returns 1 offspring |  |
| Tuesday 5th March |  | * Wrote about 1000 words focused mainly on chapter 6 * Implemented a crossover with 5 offspring * Ran the code several time with both complex and triangle program |  |
| Wednesday 6th March |  | * Ran through the GA a few more times * Changed mutation method * Swapped the unit test scores so if passes its 0.0 and 1.0 when failed |  |
| Thursday 7th March | Meeting with Sandy | * Found out that the Opacitor JVM does not work with JFrame so only code length Opacitor works with JFrame. * Discussed causes of genetic bottleneck * My GA is more likely to be random search with some extras | * Change rng to random * Add new patches to second population |
| Thursday 7th March |  | * Added 4 new patches to every called of the second population function * Fixed CSV writer as there was problems with fitness * GA acts like a framework – on larger programs it is likely to find better solutions. |  |
| Friday 8th March |  | * Wrote about 500 words for conclusion chapter of report |  |
| Saturday 9th March |  | * Did error bars on the iterative and recursive graphs |  |
| Monday 11th March |  | * Changed the percentage in the mutation method to 70% means more edits are removed |  |
| Tuesday 12th March |  | * Edit experiments for triangle program conducted * Write up and graphs |  |
| Wednesday 13th March |  | * Edit experiments for complex program conducted * Write up and graphs |  |
| Thursday 14th March |  | * GA experiments conducted * Write up and graphs |  |
| Friday 15th March | Meeting with Sandy | * Have multiple graphs of averages with error bars in edit experiments. Ask the questions: How successful were the edits applied?how often applied? * Compare random search to GA * Improve figures * Improve glossary | * Add averages edit scores to results |
| Saturday 16th March |  | * Graphs modified and more experiments conducted as based on sandys opinions from the meeting |  |
| Sunday 17th March |  | * Writing up of the experiments and modifications according to sandy’s meeting |  |
| Monday 18th March |  | * Rewriting and referencing |  |
| Tuesday 19th March |  | * Modifications of writing |  |
| Wednesday 20th March |  | * Submiited draft to sandy for final revisions |  |
| 21st March till 4th April |  | * Worked solely on the report to get finished and improved |  |