Programming for Longer Battery Life

Anna Frances Rasburn

**April 2019**

**Dissertation submitted in partial fulfilment for the degree of   
Bachelor of Science with Honours in Computing Sciences**

**Division of Computing Science and Mathematics  
University of Stirling**

Abstract

Problem

The surge in the dependence on mobile phones and portable devices raises questions regarding their environmental impact and in, in particular, their energy requirements. This dependence is due to the fast advancement in mobile technology and the convenient nature of mobiles but has also led to users wishing for a mobile battery to last longer. Mobile applications are a significant drain on battery life due to the higher demand on the device's CPU. This project's aim is to research and investigate ways in which Android Java applications can be modified to reduce CPU power and therefore require less battery life. If we can find out a way to save energy on small devices the same principles can be applied to larger devices such as large scale data centres which would lead to a big environmental benefit.

This project focuses on using optimisation techniques applied to Java, particularly those which use Genetic Improvement and small changes that developers can introduce within their source code in order to provide a significant change in the requirements of CPU power.

Methodology

This project has created a software application which applies edits to source code. The edits included the following operations: moving, deleting, swapping block statements, inserting lines and changing the use of If statements. The application uses artificial intelligence search techniques such as Hill Climbing and Genetic Algorithms to see if it is possible to find an optimised version of the source code by identifying the impact of the edits and iteratively amending the edits.

Achievements

Attestation

I understand the nature of plagiarism, and I am aware of the University’s policy on this.

I certify that this dissertation reports original work by me during my University project except for the following (*adjust according to the circumstances*):

* The technology review in Section 2.5 was largely taken from [17].
* The code discussed in Section 3.1 was created by Acme Corporation ([*www.acme-corp.com*](http://www.acme-corp.com)*/JavaExpert*) and was used in accordance with the licence supplied.
* The code discussed in Section 3.5 was written by my supervisor.
* The code discussed in Section 4.2 was developed by me during a vacation placement with the collaborating company. In addition, this used ideas I had already developed in my own time.

**Signature** *(you must delete this, then sign and date this page)* **Date**

Acknowledgements

Acknowledge anyone who has helped you in your work such as your supervisor, technical support staff, fellow students or external organisations. Acknowledge the source of any work that is not your own.

Table of Contents

The table of contents below is automatically generated from the paragraphs of style *Heading N* and *Unnumbered N*. To update this after revisions, right-click in the table and choose *Update Field* for the entire table.

Abstract i

Attestation ii

Acknowledgements iii

Table of Contents iv

List of Figures vii

List of Tables viii

1 Introduction 1

1.1 Scope and Objectives 1

1.1.1 Scope 1

1.1.2 Users 2

1.1.3 Objectives 2

1.2 Achievements 2

1.3 Overview of Dissertation 3

1.4 Glossary 3

2 State-of-The-Art 4

2.1 Genetic Improvement 4

2.2 GIN 5

2.3 GI In other Languages 7

2.4 TODO: Another GI example – Look at paper references 8

2.5 Opacitor 8

2.6 Jalen 9

2.7 Optimization Framework for Mobile Applications. 10

2.8 Code Refactoring’s effects on Energy Usage 11

2.9 Green Software Engineering 12

3 Background 14

3.1 Optimisation regarding Java 14

3.2 Genetic Improvement 14

3.3 Genetic Algorithms 14

3.4 Fitness 16

4 Methodology 17

4.1 First Section 18

4.1.1 First Subsection 18

4.1.1.1 First Subsubsection 18

4.1.2 Second Subsection 18

4.2 Second Section 18

4.3 JCodec 18

5 Creation of Toolkit 20

5.1 Tools Used 20

5.1.1 Opacitor 20

5.1.2 Fitness 20

5.1.3 Unit tests 20

5.1.4 Maven 20

5.1.5 Spring 20

5.2 Development of Toolkit 21

5.2.1 Gin-Main 22

5.2.1.1 Patch 22

5.2.1.2 Edits 22

5.2.2 Test-Runner 22

5.2.3 Opacitor 22

5.2.4 Genetic Algorithm 22

5.2.5 Hill- Climbing 22

5.2.6 Testing 22

5.2.6.1 Test Driven Development 22

5.3 Structure 22

5.3.1 Overview 22

6 Experiments 23

6.1 Recursive vs Iterative Loops 23

6.1.1 Introduction 23

6.1.2 Method 23

6.1.3 Results 23

6.1.4 Conclusion 23

6.1.4.1 Further Work 23

7 Use Cases 25

8 Results 26

9 Conclusion 27

9.1 Summary 27

9.2 Evaluation 27

9.3 Future Work 27

10 Glossary 29

References 30

Appendix 1 32

Appendix 2 – User guide 33

Appendix 3 – Installation guide 34

List of Figures

[**Figure 2.1** Flow Diagram of a Genetic Algorithm 5](#_Toc536712252)

[**Figure 3.1** Diagram of changing code lifecycle with Genetic Improvement 8](#_Toc536712253)

[**Figure 3.2** Example of GIN output 9](#_Toc536712254)

[**Figure 3.3** UML diagram of how GIN works 9](#_Toc536712255)

[**Figure 3.4** Architecture of Jalen 13](#_Toc536712256)

[**Figure 3.5** Ratings for Mobile applications 16](#_Toc536712257)

[**Figure 5.1** Example of a Bean for a removeLineEdit 21](#_Toc536712258)

# List of Tables

[**Table 1** Example to Demonstrate the levels of Fitness 3](#_Toc534297458)

# Introduction

For editorial consistency, it is important to use Word styles properly. Word 2003 onwards has so-called quick styles. If the styles referred to below are not visible on the *Home* ribbon in the *Styles* category, choose *Apply Styles* from the down arrow at the bottom right of the *Styles* category. Styles can then be applied from the drop-down box. To make a style visible as a quick style, choose *Apply Styles*, then click *Styles* (the *AA* icon), then click on the drop-down list for a style, and then *Add to Quick Style Gallery*.

Chapters are entered using the Heading 1 paragraph style. The Heading 1 style automatically moves to the start of a new page and supplies the next chapter number. The new paragraph when you press Return after a heading automatically uses the *Body First* paragraph style (like this one, with no indent on the first line).

However, the most text uses the Body Text paragraph style (like this one, with 11 point Times New Roman, 1.5 line spacing, single-sided pages). Enter most text using the Body Text paragraph style. The new paragraph when you press Return after a Body First paragraph automatically uses the Body Text paragraph style.

In general, use the default spacing that headings and paragraphs give you. Avoid using new-lines or spaces to format text. If you need to use quotes, preferably use single curly quotes ‘…’. If you wish to emphasise something, usually use *italic font*.

**Remember to Save frequently while you are working!**

Give the background to your project and context of what you have done. Sections are entered using the *Heading 2* paragraph style – th*e Heading 2* style automatically supplies the next section number.



## [Scope](http://www.cs.stir.ac.uk/~kjt/research/conformed.html) and Objectives

### Scope

The scope of this research is to explore different ways by which developers can change their source code in order to make more-optimised applications that firstly have a shorter execution time but also achieve the same functionality. This research will be followed by exploiting the same principles and seeing whether we can optimise applications to require less energy. Furthermore, it is hoped that this research will lead to the development of a toolkit that uses elements of artificial intelligence methods, such as Genetic Algorithms and Local Search, incorporating a developer’s source code and well-written unit tests to create a more-optimised version of the original source code.

### Users

The users of this research would be Android application developers and Android application users. Many Android mobile users wish to have a device which does not ‘die' as often and therefore users are likely to prefer applications that do not ‘drain’ the battery rapidly. This research can also apply to the work of Java developers more generally as this research should also be applicable to larger applications and devices such as a computer where the lower energy consumption will have a greater benefit for the environment.

### Objectives

The objective for the project is:

1. To research the different ways we can optimise Java, specifically looking at changing operators in the code to see its effect What kind of changes can be made to Java code in general?
2. Applying the research from objective 1 to create some software in a similar vein to the Gin code framework [2] [4] that uses genetic algorithms to change the source code, so it is less execution and energy intensive.
3. To learn about and use the existing Opacitor [1] [5] and Jalen tools [6] to approximate the energy consumption of the changes Java code and prove that the experiments have been successful.
4. Applying GI techniques to software so it has a shorter execution time.
5. Applying GI techniques to software so it requires less energy to run.

## Achievements

Summarise what you have achieved.

Toolkit

I began the project by creating a toolkit in a similar vein to the GIN code base [REFERENCE]. The toolkit had additional functionality of being able to take an entire project at once.

I used the GIN code as an inspiration.

The toolkit also has the test runner and Opacitor running in one movement. The Gin code had no Opacitor connecting.

Genetic Algorithm

## Overview of Dissertation

Briefly overview the contents of what follows in the dissertation.

## Glossary

# State-of-The-Art

Summarise current knowledge and what others have done in the various topics of your dissertation – in the application area and in the various technologies that you might have used or did use. Write for someone familiar with computing, but not necessarily expert in the particular topics of your project. Give references to other work by using *cross-references* to entries in the References section, like this [2] (see the References tab in Microsoft Word's ribbon).

It is important to write a *critical* literature review that identifies gaps in current solutions and that clearly shows how the project was driven to address these gaps. This chapter should, therefore, feed into well-defined requirements for the project. Avoid a superficial description of related work that does not carefully analyse its strengths and weaknesses.

State of art split into 3 sections:

Genetic Improvement/Changing Software

Energy Consumption

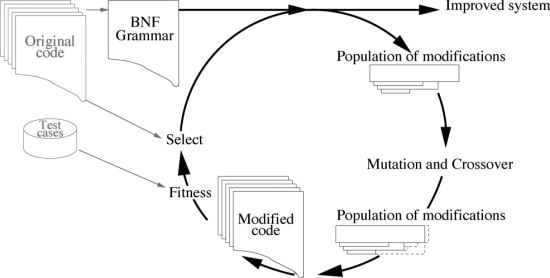
Green Software Engineering

## Genetic Improvement

Genetic Improvement is used to improve pre-existing software by using an automated search to find improved software. These improvements include improvements in energy consumption, reductions in execution time and reductions in memory consumption. In a survey of 3132 distinct papers that were published between 1995 and 2015 [7], the most popular search technique is evolutionary computing as 96% of these papers referenced or used these techniques. This demonstrates the increase in popularity of evolutionary computing.

This survey is very relevant to this study as it is one of the largest and most comprehensive surveys on computer science papers. The authors concluded that there were 82 papers out of 1131 that was about genetic improvement [7].

However, in order to cover a wide range of topics that fall within the definition of genetic improvement, the number of papers analysed in the survey had to be restricted. Hence, owing to the survey's strict definition, this may have led to the exclusion of many papers which may have had key information about Genetic Improvement.

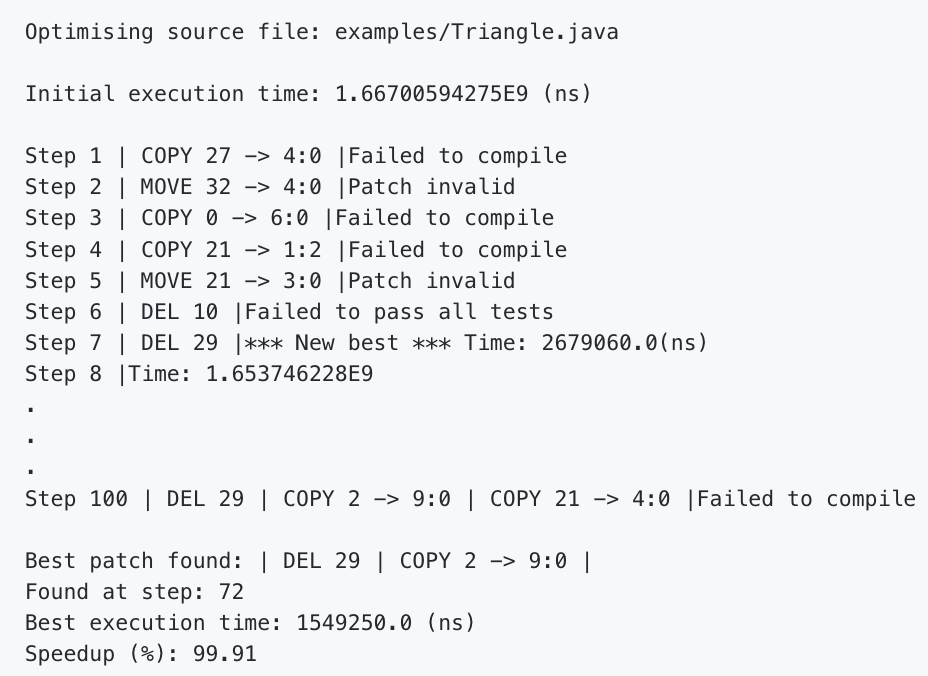


**Figure 3.1** Diagram of changing code lifecycle with Genetic Improvement

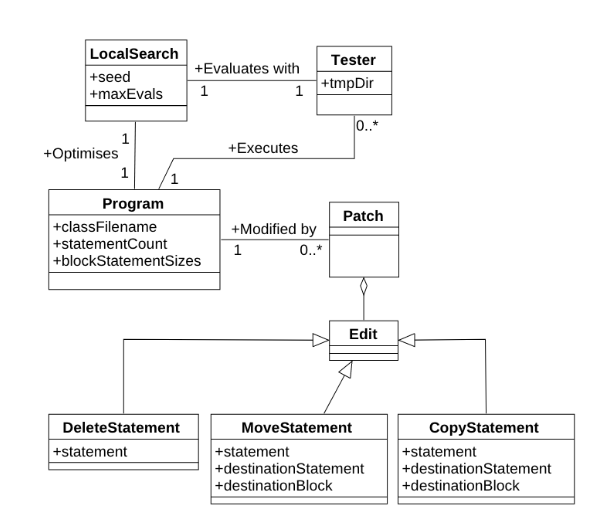
A lifecycle of improving software using genetic improvement can be seen in **Figure 2**. Langdon and Harman [8] [9] [10] state that the ‘original code is a state oracle’, therefore giving a programmer a benchmark for comparison and provide reassurance that it delivers the expected output. Hence, it may be noticed in **Figure 2** that the original code is included in the Select step after the Fitness assessment of the modified code. Genetic Improvement tries to keep the same functionality as the original code in order to improve the outputs generated as well as optimising the code. However, Langdon and Harman’s work [10] [8] assumes that the original code is correct and is passing unit tests. Therefore, this ignores integration testing and only focuses on genetic improvement for individual sections of code rather than considering a complete system.

## GIN

GIN, Genetic Improvement in no time, is a Java tool that makes it easier to implement Genetic Improvement and has a goal to “to stimulate development in GI tooling” [4] [2]. The Gin application itself optimises a Java class, using a local search to add “Edits” to the source code until a more-optimised version of the code is found. As seen in Figure 2 and 3, GIN manipulates code in three principal ways : deletion, copying and replacement. For instance, particular lines of code or blocks of code could be deleted such as in the ‘Triangle.java’, where the removal of the reference to the ‘delay()’ method creates a version of the code that passes all the unit tests and has a lower execution time. The possible changes are Line Edits, Statement Edits and Constrained Statement Edits. The inspirations for the ‘Delete, Replace, Copy, Swap’ [REFERENCE] of line edits came from research conducted by Petke and Langdon [9] with their GISMOE tool [10] [8] [2] . The Statement edits changes of ‘Delete, Replace, Copy, Swap’ [REERENCE] come from the GenProg [11] automated program repair tool. All Constrained statement edits are Gin’s own creation which presents the innovation of the Gin tool and its exploration of [9] the optimisation field. An example of the Gin code output is seen below.



**Figure 3.2** Example of GIN output



**Figure 3.3** UML diagram of how GIN works

The GIN toolkit is experimental and allows for the development of skills and knowledge about optimised code and GI techniques that could later be integrated into the JVM and compiler.

The GIN toolkit is one of the very few tools looking at integrating GI within a Java project. It applies a variety of edits to source code and concludes in the most optimised version of the source code. The toolkit creates optimised source code with shorter execution time. Its use of edits may seem to be ‘small changes' to the developer but are shown to have a large impact on execution time and optimisation.

The GIN toolkit uses a local search to make decisions on how to apply edits to the source code. The use of a local search affords the ability to find reasonable solutions in continuous state space. In the gin example, there are many neighbouring solutions made quickly that lead to an optimised version of the example source code “Triangle”. Figure 4 presents a UML diagram of the different class that make GIN work where a key component is the Local Search Class. The Local Search class is used to evaluate the tests and to optimise the program within the maximum number of maxEvals.

The fitness function of the GIN toolkit is based on whether the Junit tests are successful as it is a crucial criterion that the edited source code must have the same purpose as the original source code so that the unit tests are still passed. However, this fittest function implies that the developer has created well-written unit tests and source code. Therefore, finding an optimised version depends largely on the developers' initial code.

The GIN toolkit is limited to only optimise Java code. One can imagine many of the principles that are learnt from this software could be applied to other object-oriented languages such as C. Some of the features, such as moving edits, can have a major impact on interpreted languages, such as python, which have every line of the script interpreted individually and where the position of lines are important when optimising source code.

## GI In other Languages

A large proportion of research about Genetic Improvement focuses on object-oriented languages such as Java, C and C++, therefore ignoring dynamically typed languages such as Python. Research conducted by Haraldsson [14] focuses on genetic improvement in Python who showed that it can be used to fix multiple bugs within the Python script at the same time and in a reasonable time frame.

Haraldsson's research works in a similar way to GIN where a series of edits are applied to source code. This research applies experiments to small python scripts to see whether solutions created by genetic improvement have been successful. A random walk analysis is used, where the start state is assuming that the program is correct and the fitness is monitored as the edit list increases in size one edit at a time. The other stage of the experiment is the exhaustive neighbourhood analysis where the fitness is evaluated for every mutant of the program.

Genetic Improvement in Python has the advantage that the search space is less restricted than the search space for Java programs. This allows for more changes to the source code in a Python Script in comparison to small Java programs. Optimising Python is important because Python is ‘becoming the world’s most popular coding language’ [15] and therefore the increasing number of python developers can code an optimised version of the language with bugs fixed quickly in a reasonable time.

The larger search space of optimising Python requires a longer time to achieve results due to fewer restrictions. The results produced are also difficult to verify. For instance, it can be difficult to prove that the optimised script has the same functionality as the original script as there are no unit tests which are used a measure of fitness in Java and OOP languages. This research in Python programs is instructive as many of the concepts about genetic improvement can be used in my research but Haraldsson's research does not fully apply as it focuses on Python rather than Java.

## TODO: Another GI example – Look at paper references

<https://link.springer.com/chapter/10.1007/978-3-319-22183-0_23>

<https://link.springer.com/chapter/10.1007/978-3-319-22183-0_20>

## Opacitor

Opacitor [15] [16] is a tool which measures the energy consumption of JVM programs. It does this by using a bytecode level model of energy cost; counting the bytecodes executed by a running program and matching these to previously measured quantities of energy consumption. Opacitor ‘can detect small changes in execution profile, down to opcode level’ [15] therefore it records the energy used by a program down to its lowest level. The Opacitor measures energy usage in joules and, unlike many similar tools, does not use a CPU as a proxy for energy use. Therefore, using joules and bytecode increases the accuracy of the readings.

Opacitor [1] [5] is one of the first tools of its kind to measure the energy produced by an executing program using Search-Based Software Engineering. Due to the implementation of the Opacitor it can detect the smallest changes in execution, thereby recording all the possible energy usage from the entire execution of the program. The Opacitor is unaffected by the rest of the computational environment so that when it examines the smallest changes these are the only changes in the program and does not include any other environmental factors.

The Opacitor is deterministic, i.e. it will produce the same output from a given input. The Opacitor’s results have helped reduce a program’s energy usage by up to 70% in the best case and 20% in the worst case [1]. The Opacitor is one of the few tools that can detect small changes in the execution and is responsive to metaheuristic searches [1].

The Opacitor's only focus is on the CPU and it does not consider other features of a mobile phone that use a large amount of battery power such as the display, WI-FI and GPS. The Opacitor is exclusive to applications and the CPU does not consider other features or how they work with applications and the effect that they may have on the battery.

## Jalen

Jalen is a ‘runtime measurement software for estimating the energy consumption at code level for Java applications’ [6]. Jalen uses information provided through application-level monitoring tools to calculate the energy used by an application. The energy information is estimations from the levels of threads and methods. The information collected is method durations, CPU time and so on. The information gathered is used to create statistics which is used to calculate the energy consumption as well as using energy models seen in Figure 5. The energy consumption is calculated per method and exposed as a service. Jalen can provide information about Net Energy, Net Library energy and all energy usage which includes energy consumption of all methods running in the JVM.



**Figure 3.4** Architecture of Jalen

Jalen offers a valuable tool that provides an energy count on a code level so that there is an increased understanding at the energy consumption and distribution in software. Jalen is also one of the very few toolkits that detect energy hotspots in code and knows what particular areas of software are energy intensive. Jalen has been tested with experiments on a range of different programs from simple algorithms to real-world scenarios demonstrating its versatility and the potential of energy code profiling. The information gathered from Jalen will help developers find out which methods are energy intensive so they can change the ways they code to ‘reduce the energy footprint of their application' [6].

Unlike Opacitor, Jalen prefers statistical sampling over bytecode instrumentation. It also presents data as percentages rather than raw Joules values. This is because different hardware changes the determined results which is not a problem that Opacitor finds. Offering the results as a percentage means that results do not change drastically when the hardware changes as they do when the data are presented as a Joules count.

## Optimization Framework for Mobile Applications.

<https://arxiv.org/pdf/1608.05248.pdf>

<https://ieeexplore.ieee.org/document/7935484/references#references>

Li et al created a framework which looks at Source-level energy optimization with a particular mobile application [17]. This framework proposed an energy-optimization framework with a source code energy model that allows developers to know about energy usage. This framework changes and improves the code to save CPU energy by analysing energy features of the source code before optimisation. The model maps the energy use to the basic operation within the source code which works in a similar way to Jalen as it focuses on the energy impact on individual operations within the source code.

This framework was a success as it was able to save between 6.4% to 50.2% of the CPU energy in various mobile applications. These toolkits and frameworks highlight to developers which operators are energy intensive, so they may make better choices when developing an application. This framework focuses on mobile applications which are different to Jalen, the significance of which is that the way mobile applications interact with mobile hardware is very different from the way software interacts with other machines.

Unlike the Opacitor, this framework does not look at the bytecode level and therefore its results may be affected by features of a mobile device such as WI-FI and GPS. These features would have a major impact on the results because they are energy intensive and will increase energy usage results markedly. The reliance of highlighting to a developer their energy usage means that for there to be any saving in energy, this is entirely the developer's responsibility and then trusting them to be energy conscious.

## Code Refactoring’s effects on Energy Usage

Research conducted by Sahin et al [18] looked at energy impacts of 197 applications when 6 commonly-used refactoring techniques have been applied to those applications and the change in energy usage those applications have after those techniques applied. The results of this research concluded that there was an impact on energy usage and that a change in refactoring techniques can effectively increase or decrease the amount of energy used by an application. This study conducted 394 experiments on those 197 applications and found that 28% (109 experiments) indicated ‘a statistically significant difference in energy usage between the original and refactored versions.’ [18]. The refactoring technique of ‘Convert Local Variable to Field’ [18] had the largest positive impact on energy out of the six different techniques. This study is notable in that it presents the effects of different refactoring techniques and has a difference in results. However, this study could have been extended to include more applications and refactoring techniques as it makes it harder to generalise. This study uses only two JVMs: 6 and 7, which does not consider whether new versions of Java have integrated optimisation techniques that affect energy usage.

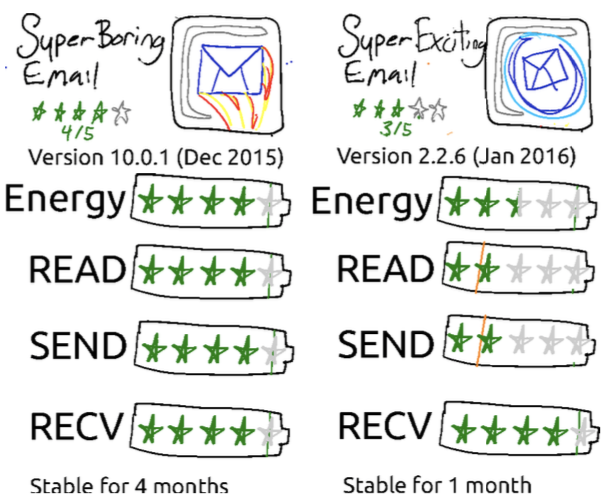
This research supported by research conducted by Morales et al [19] who focuses on the effects of antipatterns in mobile applications and its effect of energy efficiency. This research found that energy consumption of apps with anti-patterns are not statistically different to apps which do not have these anti-patterns. The research did conclude that refactoring anti-patterns can be positive or negative therefore developers need to consider the energy impact of refactoring when creating applications. The research was conducted with 59 android applications and showed that applications containing anti-patterns consume more energy than those which do not have them, therefore refactoring and removing these antipatterns has a major effect on energy usage. Similar to the Sahin's study, this study does not use many programs and this, therefore, makes it harder to generalise. This study is very focused on Android applications and considers elements of applications that only affect Android.

## Green Software Engineering

Green Software engineering, a term coined by Hindle [20] is a type of software engineering which is more environmentally sustainable so it takes into consideration the power usage of creating software. There are very few uses of physical measurements in software engineering even though this is critical for measuring energy. This lack of physical measurement creates limitations in regard to experimentation and analysing energy usage.

Green Software Engineering is an important yet emerging field. Because this field is at an early stage of development, it has been marked by a low level of resources , a small coherent community and an uncertain methodology. This field needs an increase in support and research which can pose a negative image of making the field an ‘impossible bar' [20]that may not be achieved. An increase in support and research may come with the on climate change [21]. This issue has only amplified the importance and necessity of this research if the desire to keep our current technology-heavy lifestyles as well as to ‘limit climate change'.

Research conducted by Hindle as it focuses on software’s interaction in other contexts other than mobile applications. Hindle researches in mobile applications, Data Centres and embedded sensors. Hindle has a proposal to have an energy rating on mobile applications included with the mobile application other functions, an example of this can be seen in Figure 6. Therefore, the user has the responsibility of choosing a more sustainable application which offers an interesting proposal as it gives the user more responsibility and can be used an extra selling point for a mobile application.



**Figure 3.5** Ratings for Mobile applications

All of the previous research mentioned is useful in regard to green software engineering. As the combination of changing source code using Genetic Improvement to reduce energy consumption is the focus and aims of this project.

# Background

A large proportion of this research is building on research previously conducted by Dr Alexander Brownlee [1] and Dr David White [2]which focuses on methods used to optimise source code without losing any functionality through Genetic Improvement. This research also builds upon the research by Bruce et al [3], which considered methods to automatically change source code to reduce energy usage.

## Optimisation regarding Java

Optimisation is ‘the act of making the best or most effective use of a situation or resource'. Optimisation in Java focuses on goals within or related to the source code that we can optimise [REFERENCE]. This research focuses on changing Java code and Java Projects that can be used to improve an application’s execution time and reduce an application’s requirement for the energy it needs to run.

To find solutions to optimisation problems, the technique chosen for the current work is based on Genetic Algorithms seen in *Section 3.3*, which find solutions using the notion of fitness seen in section 3.4.

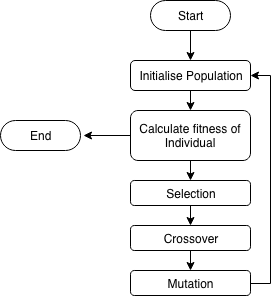
## Genetic Improvement

Genetic Improvement(GI) improves software by using machine learning and optimisation techniques. These techniques include Evolutionary Algorithms of which Genetic Algorithms are a sub-category. Genetic Improvement can be used to improve pre-existing software with improvements such as reducing energy consumption, execution time and memory consumption.

## Genetic Algorithms

Genetic Algorithms(GA) are algorithms based on the biological notion of evolution and natural selection as conceived by Charles Darwin. Genetic algorithms use the principle of survival of the fittest and inheritance. The fittest traits of the parent are passed onto the children who prevail. Throughout the duration of the algorithm, many possible solutions to a problem are found; these solutions are called ‘individuals'. Every iteration of the algorithm selects a solution with high fitness from a collection of individuals. Subsequent iterations generate further solutions from these individuals. Some of these solutions are closer at achieving the most optimised solution to the problem and some solutions are not. The selection stage decides what traits are to be kept for the future generations. Traits that do not confer improved fitness are disregarded by the selection in the next iteration.

In standard Genetic Algorithms, there are five main steps that are followed, this is depicted in Figure 1. The steps are: initialising population, fitness, selection, crossover and mutation.



**Figure 2.1** Flow Diagram of a Genetic Algorithm

An “Individual” or “chromosome” is a possible solution to the problem. In a classic GA, a chromosome is a string of bits and each individual bit is a gene. The selection stage ‘selects’ individuals with high fitness that will progress to the next population. In the mutation stage, individual genes of the highly fit parent chromosome are flipped from one to zero or vice versa to create a new mutated chromosome to pass onto further populations. The crossover stage requires taking genes from the two parents and combining them to make a new individual who may be a more optimised solution to the problem than their parents.

A “Population” is a set of individuals. The initialise population stage involves gathering a collection of individuals together to form a population. As individuals increase their fitness throughout the lifetime of the GA, the overall population has a higher fitness. It therefore has a higher possibility of finding a solution to the problem.

In the problems that form the focus of this project, Genetic improvement problems, the individuals can be lists of changes in source code or new versions of the source code. In the GI tool Gin [4] [2], an individual is a list of "Edits". Each "Edit" is a single change to the source code such as moving a line from one place to another. The mutation stage adds or removes edits to source code and in the crossover stage, a combination of edits is combined together into a new version of the source code. An edit is similar to a gene but not a direct parallel because an edit is what has changed in the code. A "Patch" is similar to a chromosome but not by the traditional definition because a "Patch" may comprise of no edits if it happens that in the current version of source code in also the most optimised source code. A patch could also contain a collection of edits to create the optimised source code. In the Gin Code [2] [4], a collection of patches is similar to a classic Genetic algorithm’s definition of a population.

[REFERENCE THIS- USE THAT GA BOOK]

## Fitness

As mentioned previously, the fitness of an Individual is how well that particular Individual performs the task. In this instance, there are several potential measures of fitness, for example, one can ask: Does the new source code compile? Does the new source code pass the unit tests? What is the energy consumption of the application? This has meant that each element of fitness needs a weight which is how important that element is in the fitness function. For example, whether the source code can compile has a larger weight than the unit tests because if a program does not compile then the unit tests are irrelevant as the program has lost all functionality. Table 1 presents the different elements of the fitness and a way that the different weights can be used together. In the example below, Individual 3 would be the most optimised because it has passed the most weighted elements such as it can compile and energy consumption although it failed two-unit tests.

**Table 1** Example to Demonstrate the levels of Fitness. P denotes a pass, F a failure.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Did the source code compile?** | **Unit tests** | | | **Energy**  **Consumption** | **Use this**  **individual?** |
| **Unit test 1** | **Unit test 2** | **Unit test 3** |
| *Individual 1* | Yes | P | P | P | 3000J | Yes |
| *Individual 2* | No | P | P | F | N/A | No |
| *Individual 3* | Yes | P | F | F | 2100J | Yes |
| *Individual 4* | Yes | F | P | P | 4000J | Yes |

For any optimisation problem, the fitness function needs to be defined before a solution is executed. Changing the fitness function can present new ways to optimise Java. For example, by changing the fitness to focus on execution time or energy consumption, the Java application can be optimised in different ways.

In this research I decided to use a lexicographical fitness. In the research this meant that the first measure of fitness is number of unit tests passed. Once the solution has passed a high number of the unit tests, we use a second measure of fitness which is the energy measurement. The solutions are compared on which has the lowest energy and that solution is seen as the most optimised solution of that generation. This was done because I thought it was more important to prioritise the functionality of the application over the energy because an application that does not work as intended will not be used so the energy measurement would be irrelevant.

# Methodology

**Note: This part of the dissertation will normally be expanded to be *a series of chapters*.**

The technical body of the dissertation consists of a number of chapters (just one here, but there will usually be more). Follow a logical structure in how you present your work. This may be the phases of the software development cycle, or different aspects of the system's design, etc. ***However, please do not write your dissertation to read like a diary.***

The emphasis should be on requirements, design and evaluation, with implementation details being of lesser importance. The requirements should be clearly stated, following from the client needs and weaknesses identified in the state-of-the-art review. The design should include discussion of the choices that were available and why particular decisions were made.

Include a chapter demonstrating what you have achieved and how your system is used in practice – for example showing a typical session as a series of pasted in screenshots, with an accompanying commentary.

If relevant, you ***must*** also include a chapter explaining how you obtained feedback from your “customer” or potential users of your system, what feedback you actually obtained, and your analysis and comments. The evaluation should relate back to the requirements, and discuss the extent to which these were met.

Include TTD

Include spring

Toolkit then GA etc

To begin with, I created a software application that was similar to the GIN code []. However, I created the application using a new structure and different edits.

The edits I created were looking at the impact of moving large chunks of the source code and seeing if that had an impact. Other edits I created were inserting lines particularly looking at return statements and break statements. The other edits I created were using if statements and changing those.

When I created the gin code, I used Test Driven Development [REFERENCE] because …. I also used spring framework as it is a framework that I have used before and I personally found it easier to install edits.

I started with a random search then moved to local search/hill-climbing/ GA.

I used the test runner and Opacitor as a measure of fitness.

I also looked at the difference in recursive and iterative loops.

## First Section

Subdivide your text into sections.

### First Subsection

If necessary, also use subsections. Subsections are entered using the *Heading 3* paragraph style (all these heading styles are self-numbering).

#### First Subsubsection

If you really need subsubsections, enter these using the *Heading 4* paragraph style.

### Second Subsection

And, as required, more subsections.

## Second Section

As an example of a figure, consider Figure 1. Captions are entered using the *Figure* paragraph style. The figure below is placed in a *Body Centre* paragraph, which is set up in this document to have an automatic *Figure* paragraph following it. *Figure* has automatic figure numbering, and it is possible to make *cross-references* to figures. Move large figures to the top of the next page, *past any other text,* rather than having a big gap in the text.

diss-fig

1. Highly Technical Diagram

## JCodec

JCodec is ‘a pure implementation of video and audio codecs’ [12]. JCodec is currently being used by the GIN toolkit as a set of programs to improve. There are 203 different programs within JCodec. JCodec is a good library for GIN to optimise because they are written in Java and they contain more than 10-unit tests per program.

The high number of unit tests allow for the functionality not to be compromised significantly so that the programs still perform in the same way after optimisation using GIN. The JCodec implementation is used in a large number of programs and large numbers of developers. This has created much documentation and if GIN finds a way of improving the library, there can be large saving in energy if all of those programs use the optimised versions.

JCodec presents its own problems when optimising because many programs use media elements which require much energy to work and cannot have too much content removed otherwise the functionality of the program is adversely affected, causing a wide range of errors. This library of GIN applications to JCodec applications on this library does not work on other multimedia techniques in Java such as the Oracle multimedia API [13] as they work in different ways in comparison to JCodec.

# Creation of Toolkit

## Tools Used

### Opacitor

### Fitness

### Unit tests

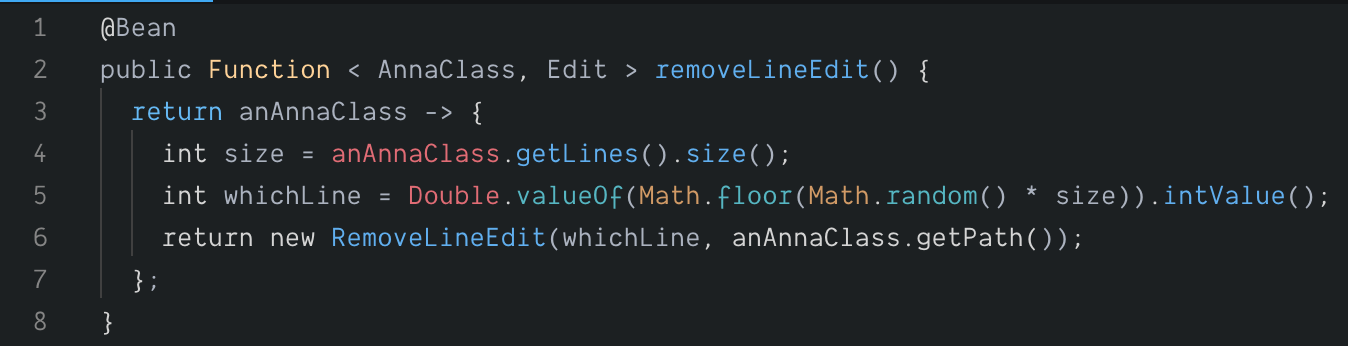
### Maven

Maven [22]is a ‘build automation tool’ with a focus for java projects. I used Maven to bring in java libraries and dependencies. Maven also allowed me to structure the project in a way that allowed for different sections of the project to be its own module and I could bring in that module as a dependency and use it as if it were a library. For instance, I had the opacitor as one module and in the main module (were the main processes begin) would bring in the opacitor as a dependency. The reason I used Maven is because I find it easier to use than other dependency interjection frameworks such as gradle, this is because the dependencies are automatically ‘brought in’ to the project. Maven is useful as it allows to structure the project, for instance each module has its own ‘pom.xml’ with their dependencies as well as setting which version of Java each module has, for instance the Opacitor needed to be Java 7 or less to run which is different from the main module which I set to be Java 8 so maven allowed different modules to have certain dependencies and Java version.

### Spring

(INCLUDE SCREENSHOTS OF CODE)

Throughout development of the Anna Gin Toolkit, the use of the spring framework allowed for a ease of inserting particular features into the application. The spring framework is an ‘application framework and inversion of control container for the Java platform’ [16].The use of spring framework was decided because of previous experience with the framework and a personal preference of wishing to avoid direct association with reflection. From a personal experience, the spring framework allowed for easier insertion of different types of edit, an example of an edit bean is seen in Figure 5.1. As you can see in figure 5.1, there is a @Bean annotation to denote the edit and the use of a lambda and function to create the edit. In the project, I created a bean for each of the different types of edit, this allowed for the ApplicationContext to bring all the beans in when ran.



**Figure 5.1** Example of a Bean for a removeLineEdit

This gave for increased ease as if a developer wished to add another type of edit, they would just need to create a new bean in ‘EditConfig’.java file for that edit to be brought into the program.

Include I used java 8 features such as lamda

## Development of Toolkit

The different steps of developing the AnnaGin Toolkit are:

1. Gin-Main
2. Test-Runner
3. Opacitor
4. Genetic Algorithm
5. Hill-Climbing
6. Testing

### Gin-Main

#### Patch

#### Edits

### Test-Runner

### Opacitor

### Genetic Algorithm

### Hill- Climbing

### Testing

#### Test Driven Development

## Structure

### Overview

INSERT UML DIAGRAM OF THE TOOLKIT HERE

TALK ABOUT USING MODULES IN INTELLIJ. USING MAVEN DEPENDENCYS TO BRING IN DIFFERENT MODULES INTO EACH OTHER

# Experiments

## Recursive vs Iterative Loops

### Introduction

After learning about the Opacitor and thinking of different ways that code can be changed for energy optimisation and a suggestion from another student, I thought of the effects of different loop types on energy. I specially focused on recursive loop in comparison to iterative loops and whether calling a method repeatedly in a loop has an effect on energy.

I hypothesized that recursive loops take more energy to conduct the same functionality as an iterative loop, I do not expect the energy measurements for recursive loops to be a major increase in comparison to an iterative loop.

### Method

In this series of experiments, I created an array of a certain length with a certain seed. I created in total of 48 (roughly) arrays with seeds ranging from 10 to 10,000 and array sizes ranging from 10 to 10,000. The 48 arrays are split into 3 subgroups: 16 totally random arrays, 16 ascending arrays and 16 descending arrays. I did conduct some experiments with an array length of 100,000 but conducting these experiments took a long time for up to 4 hours. For each individual array, I ran the array through the complication unit and the Opacitor Bytecode histogram thirty times. Repeating the array this way helped reduce noise in the experimentation.

There is an element that the energy measurements are obscured by other elements of the program such as generating the random arrays or sorting the arrays into ascending array for the bubble sort. IS IT THO THE OPACITOR JUST LOOKS AT THE CODE

### Results

### Conclusion

#### Further Work

The research gathered in this experiment could be conducted again with a different sort algorithms such as a Insertion Sorts or a Merge sort. This would lead to some potentially interesting results in looking at difference in energy for different sorts. Some sorts may be require more energy to sort the algorithm. This could lead to interesting questions about whether an algorithm’s Big O has an impact on energy so for instance does if an algorithm such as Merge sort (which is a faster algorithm than a Bubble Sort), has a smaller energy measurement than slower algorithms like Bubble Sorts.

This research can also be implemented into the AnnaGin toolkit. This could be done as implementing an Edit that changes a recursive loop to an iterative loop because largely iterative loops have lower energy measurements so potentially changing the loops can reduce the programs overall energy usage. This could led to interesting problems because recursive loops are normally implemented for certain reasons so there needs to be verification that the program’s functionality is not compromised by making the program not recursive, this can be conducted by implementing good unit tests that check for compromises in functionality.

# Use Cases

# Results

# Conclusion

## Summary

Summarise what you have achieved.

## Evaluation

Stand back and evaluate what you have achieved and how well you have met the objectives. Evaluate your achievements against your objectives in section 1.2. Demonstrate that you have tackled the project in a professional manner.

(The previous paragraph demonstrates the use of automatic cross-references: The “1.2” is a *Cross-reference* to the text in a numbered item of the document, it is *not* literal text but a *field.* The number that appears here will change automatically if the number on the referred-to section is altered, for example, if a chapter or section is added or deleted before it. Cross-references are entered using Word's **Insert** menu. Cross-references are set to update automatically when printed, but may not do so on-screen beforehand; you can update a field manually on-screen by right-clicking on it and selecting Update field from the pop-up menu.)

## Future Work

Explain any limitations in your results and how things might be improved. Discuss how your work might be developed further. Reflect on your results in isolation and in relation to what others have achieved in the same field. This self-analysis is particularly important. You should give a critical evaluation of what went well, and what might be improved.

Future work can include adapting the toolkit to analyse mobile applications with a focus on the Android Framework such as the Android Pie 9 Operating System [23]. This offers the opportunity to reprogram Java mobile applications, so they require less battery of the mobile device to achieve the same functionality. Implementing this feature would offer an interesting set of problems for instance, most android applications have a ‘AndroidManifest.xml’ file which is ‘describes essential information about your app to the Android build tools, the Android operating system’ [24]. This android manifest file contains many of the application’s components so if the toolkit removes these components, the toolkit needs to have the ability to change this android manifest file, so it adapts to the modified application.

The current toolkit only concentrates on Java programs. It is thought that a lot of the research done about the changing of code and Genetic Improvement could be applied to other languages to optimise them in a similar way. Other languages such as Python are gathering in popularity and the current toolkit would not be able to adapt to these applications. Adapting the toolkit would also be useful for large programs that require a lot of energy to run and optimising these large programs would save last amounts of energy. Many of such large programs use multiple languages that work in tandem, so the toolkit would need to be able to optimise multiple languages as well as ensure that the different areas of the program still work together without compromising the applications functionality.

# Glossary

GI – Genetic Improvement

# References

|  |  |
| --- | --- |
| [1] | N. B. a. J. S. A. Brownlee, “Search-Based Energy Optimization of Some Ubiquitous Algorithms,” *IEEE Transactions on Emerging Topics in Computational Intelligence,* vol. 1, no. 3, pp. 188-201, 2017. |
| [2] | D. White, “GI in no time,” in *GECCO '17 Proceedings of the Genetic and Evolutionary Computation Conference Companion*, 2017. |
| [3] | J. P. a. M. H. B. Bruce, “Reducing Energy Consumption Using Genetic Improvement,” in *GECCO*, 2015. |
| [4] | White, “gintool/gin,” Github, 2018. [Online]. Available: https://github.com/gintool/gin. [Accessed 07 October 2018]. |
| [5] | S. Black, “Opacitor: Search-Based Energy Optimization of some ubiquitous algorithms,” DAASE Project Blog, 01 June 201. [Online]. Available: https://daaseblog.wordpress.com/2017/06/01/opacitor-search-based-energy-optimization-of-some-ubiquitous-algorithms-sbse/. . [Accessed 07 October 2018]. |
| [6] | A. Noureddine, “Towards a Better Understanding of the Energy,” *Université des Sciences et Technologie de Lille,* 2014. |
| [7] | J. Petke, S. O. Haraldsson, M. Harman, W. B. Langdon, D. R. White and J. R. Woodward, “Genetic Improvement of Software:,” *IEEE Transactions on Evolutionary Computation,* vol. 22, no. 3, pp. 415 - 432, 2018. |
| [8] | M. H. W. B. L. a. W. W. J. Petke, Using Genetic Improvement and Code Transplants to Specialise a C++ Program to a Problem Class, Granada, Spainnnnnn: Springer, 2014. |
| [9] | B. Y. L. J. P. a. M. H. W. B. Langdon, “Improving CUDA DNA analysis software with genetic programming,” in *Proceedings of the Genetoc and Evolutionary Computation Conference, GECCO 2015*, Madrid, Spain, Jully 11-15, 2015. |
| [10] | W. B. L. a. M. H. J. Petke, “Applying genetic improvement to minisat,” *International Symposium on Seach Based Software Engineering,* pp. 257-262, 2013. |
| [11] | T. N. S. F. a. W. W. C. Le Goues, “Genprog: A generic method for automatic software repair,” *IEEE Trans. Software Eng,* vol. 38, no. 1, pp. 54-72, 2012. |
| [12] | JCodec, “Jcodec.org,” [Online]. Available: http://jcodec.org/. [Accessed 02 January 2019]. |
| [13] | Oracle, “Oracle Multimedia Documentation,” Oracle.com, 2019. [Online]. Available: https://www.oracle.com/technetwork/database/database-technologies/multimedia/documentation/index.html. [Accessed 02 January 2019]. |
| [14] | S. Haraldsson, “Genetic Improvement of Software: From Program Landscapes to the Automatic Improvement of a Live System,” PhD, University of Stirling, 2019. |
| [15] | P. h. b. c. p. t. a. v. n. audience, “Python has brought computer programming to a vast new audience,” The Economist, 19 July 2018. [Online]. Available: https://www.economist.com/science-and-technology/2018/07/19/python-has-brought-computer-programming-to-a-vast-new-audience.. [Accessed 02 January 2019]. |
| [16] | Spring, “Guides,” Spring, [Online]. Available: https://spring.io/guides. [Accessed 24 01 2019]. |
| [17] | X. L. a. J. Gallagher, “A source-Level Energy Optimization Framework for Mobile Applications,” 2019. |
| [18] | L. P. a. J. C. C. Sahin, “How Do Code Refactoring Affect Energy Usage?,” 2019. |
| [19] | R. S. F. K. F. C. a. G. A. R. Morales, “Anti-patterns and the energy efficiency of Android applications,” 2019. |
| [20] | A. Hindle, “Green mining: a methodology of relating software change and configuration to power consumption,” *Empirical Software Engineering,* vol. 20, no. 2, pp. 374-409, 2013. |
| [21] | J. Watts, “We have 12 years to limit climate change catastrophe, warns UN,” The Guardian, 08 10 2018. [Online]. Available: https://www.theguardian.com/environment/2018/oct/08/global-warming-must-not-exceed-15c-warns-landmark-un-report. [Accessed 24 01 2019]. |
| [22] | T. A. S. Foundation, “Apache Maven Project,” Apache, 29 01 2019. [Online]. Available: https://maven.apache.org/what-is-maven.html. [Accessed 30 01 2019]. |
| [23] | Android, “Android 9 Pie,” Android, [Online]. Available: https://developer.android.com/about/versions/pie/. [Accessed 24 01 2019]. |
| [24] | Android, “App Manifest Overview,” Android Developers, [Online]. Available: https://developer.android.com/guide/topics/manifest/manifest-intro. [Accessed 24 01 2019]. |
| [25] | J. G. a. C. B Evans, Optimizing Java, O'Reilly Media, 2018. |
| [26] | M. Negnevitsky, Artificial Intelligence, Harlow: Addison-Wesley/Pearson, 2011. |
| [27] | M. Mitchell, An Introduction to Genetic Algorithms, Cambridge, Mass: MIT Press, 1996. |
| [28] | S. Luke, Essentials of Metaheuristics, Lulu, 2015. |
| [29] | E. B. A. B. Z. K. J. S. a. N. V. N. Burles, “Object-Oriented Genetic Improvement for Improved Energy Consumption in Google Guava,” Search-Based Software Engineering , 2015. |
| [30] | E. B. B. B. a. K. S. N. Burles, “Specialising Guava's Cache to Reduce Energy Consumption,” 2018. |
| [31] | W. B. L. a. J. P. Petke, “Software is not fragile,” *First Complex Systems Digital Campus World E-Conference 2015,* pp. 203-211, 2017. |
| [32] | Oracle, “The Java Tutorials,” Java Documentation, [Online]. Available: https://docs.oracle.com/javase/tutorial/java/javaOO/lambdaexpressions.html. [Accessed 24 01 2019]. |

Appendix 1

You may have one or more appendices containing detail, bulky or reference material that is relevant though supplementary to the main text: perhaps additional specifications, tables or diagrams that would distract the reader is placed in the main part of the dissertation. Make sure that you place appropriate cross-references in the main text to direct the reader to the relevant appendices.

*Note that you must* ***not*** *include your program listings as an appendix or appendices*. You should submit such material to the project *digital repository*.

Appendix 2 – User guide

If you produced software that is intended for others to use, or that others may wish to extend/improve, then it is advisable to include user guide and installation guide appendices.

Appendix 3 – Installation guide

If you produced software that is intended for others to use, or that others may wish to extend/improve, then it is advisable to include user guide and installation guide appendices.