Enigma and BOMBE simulation

Student Name: A.L. Gillies

Supervisor Name: M. Johnson

Submitted as part of the degree of MEng Computer Science to the

Board of Examiners in the School of Engineering and Computing Sciences, Durham University

Abstract — These instructions give you guidelines for preparing the design paper. DO NOT change any settings, such as margins and font sizes. Just use this as a template and modify the contents into your design paper. Do not cite references in the abstract.

The abstract must be a Structured Abstract with the headings **Context/Background**, **Aims**, **Method**, and **Proposed Solution**. This section should not be no longer than a page, and having no more than two or three sentences under each heading is advised.

This section briefly introduces the project, the research question you are addressing. Do not change the font sizes or line spacing in order to put in more text.

Note that the whole report, including the references, should not be longer than 12 pages in length (there is no penalty for short papers if the required content is included). There should be at least 5 referenced papers.

This section presents the proposed solutions of the problems in detail. The design details should all be placed in this section. You may create a number of subsections, each focusing on one issue.

This section should be up to 8 pages in length. The rest of this section shows the formats of subsections as well as some general formatting information. You should also consult the Word template.

The font used for the main text should be Times New Roman (Times) and the font size should be 12. The first line of all paragraphs should be indented by 0.25in, except for the first paragraph of each section, subsection, subsubsection etc. (the paragraph immediately after the header) where no indentation is needed.

In general, figures and tables should not appear before they are cited. Place figure captions below the figures; place table titles above the tables. If your figure has two parts, for example, include the labels "(a)" and "(b)" as part of the artwork. Please verify that figures and tables you mention in the text actually exist. make sure that all tables and figures are numbered as shown in Table 1 and Figure 1.

The list of cited references should appear at the end of the report, ordered alphabetically by the surnames of the first authors. The default style for references cited in the main text is the Harvard (author, date) format. When citing a section in a book, please give the relevant page numbers, as in (**?**, p293). When citing, where there are either one or two authors, use the names, but if there are more than two, give the first one and use "et al." as in , except where this would be ambiguous, in which case use all author names.

You need to give all authors' names in each reference. Do not use "et al." unless there are more than five authors. Papers that have not been published should be cited as "unpublished" (**?**). Papers that have been submitted or accepted for publication should be cited as "submitted for publication" as in (**?**) . You can also cite using just the year when the author's name appears in the text, as in "but according to Futher (**?**), we . . . ". Where an authors has more than one publication in a year, add 'a', 'b' etc. after the year.

I ABSTRACT

Context/Background - The world has changed significantly in the last fifty years. Ideas that were had before modern computing was fully realised need to be re-evaluated with a modern perspective. One

such example is the Engima machine and it's counterpart, the BOMBE. What is the computational speed up from 1940 to now?

Aims - The aim of this paper is to show the improvements of modern computing using the comparison of an antiquated design against several modern re-interpretations.

Method - After both the engima and BOMBE are implemented they will then be improved on using modern techniques. Each will then be tested. The time taken for a standard engima machine to encrypt, the computerised version to encrypt and the parallelised version to encrypt will be compared to one another to give a good indication of the improvements over time.Another comparison will be made between the standard enigma, the computerised enigma and the parallelised enigma, this should re-enforce the indication of computer improvements

Proposed Solution - The enigma machine will be implemented through c++ code, this will form the baseline for a modern interpretation of the machine as it will not run anything in parallel. This code will then be parallelised using different modern techniques, these will then be used as the final comparison to show the significant increases in performance that any antiquated idea can expect.

Keywords — Enigma, BOMBE, Parallel, Modern, Computation, Evaluation, Reinterpretation, Comparison, C++, Antiquated.

II   INTRODUCTION

A   *Description and Purpose*

This project is to be an investigation into the possible improvements that can be gained on an antiquated technique by parallelisation against unparallelised and a time accurate representation. This will be done by first creating a modern equivalent of the antiquated technique. The modern equivalent will then be improved upon using parallelisation techniques. This is to be done on the BOMBE using the enigma machine.

B   *What is the Enigma Machine?*

The enigma machine was an early to mid 20th century cryptography device, it was used in both commercial and military applications. Most famous for its use by Nazi Germany during the second world war, the electro-mechanical rotor device was made up of three rotors, a plugboard, and a reflector. The user interface for the device is a keyboard and a lightboard, once a letter is pressed on the keyboard, it is then encrypted and the corresponding encrypted letter will light up on the lightboard.

The device encrypts any given letter is the same way. Once a letter has been pressed on the keyboard a signal is passed through the plugboard this is a plug-based one to one mapping of all the letters on the keyboard to a unique letter of the alphabet. This output is then passed to the first of three rotors, a rotor is a set one to one mapping that will then step to the next letter once used. The order of either alphabet on the rotor cannot be changed but the mapping of the input letter to the output letter will change after each time it is used, this is called the stepping mechanism. The

first rotor will step each time a letter is pressed and the second will only step if the first rotor has done a complete rotation, the third will only step if the second has done a complete rotation. The final component is the reflector this is similar to the plugboard that it is a one to one mapping but has the issue that it cannot map a letter to itself. After the reflector the signal is sent back through all three rotors and then through the plugboard to light up one of the letters on the lightboard.

The device has several settings that can be changed. Firstly the plugboard is reprogrammable, the plugboard is a one to one mapping between two alphabets. Secondly, the device typically has three rotors that can be used at any one time, there are eight rotors that can be put in any one of the three position, left, right and centre. Having chosen three of the rotors and their positions in the device, the initial rotation of each also has to be chosen. Finally, the reflector is also programmable, this is another one to one mapping between two alphabets.

C   What is the BOMBE?

The BOMBE is the antithesis of the Engima machine, it was specifically designed by British cryptologists in 1939 to counter the enigma encoded messages that were used by the axis at the time. The BOMBE was designed to to discover the settings used by the enigma operators that were changed daily.

The BOMBE was an electro-mechanical device that replicated the action of several enigma machines wired together; the British BOMBE contained 36 enigma representations.

A lot more to write...

D   What are the parallelisation techniques?

They are...

E   Deliverables

E.1   Minimum Objectives

- Simulate Enigma and show its correctness, by encrypting plaintext and comparing this to its known ciphertext to show this correctness, this should be done over a significant database of plaintext and ciphertext.
- Simulate BOMBE and show its correctness, taking some of the aforementioned stock ciphertext and running the BOMBE on this input such that the plaintext is output as the result.

E.2   Intermediate Objectives

- Prove the correctness of the Enigma simulation by encrypting plaintext and comparing this to its known ciphertext to show this correctness. This should be done over a significant database of plaintext and corresponding ciphertext.
- Prove that the BOMBE simulation is correct by taking all of the aforementioned stock ciphertext and running the BOMBE on this input such that the plaintext is output as the result.
- Both simulations should also be proved to be effective on random inputs and also in relation to one another; the Enigma machine should be run on a random input and the output should then be used as the input on the BOMBE which should in turn return the random plaintext.
- Evaluate average computational time on inputs of a set size.

E.3 Advanced Objectives

- Increase the scope of the BOMBE simuation by running the simulation in parallel. This means that different parts of the ciphertext could be run in parallel or the whole cipher text being run with each parallel part testing different configurations of the Engima machine.

III DESIGN

A Tools

The only tools I will need are the Intel compiler, this will allow for the unimproved version of both the Enigma and BOMBE to be written in C++ meaning that it can be compiled by a standard C++ compiler. The Intel tools will allow for the parallelisation techniques to be used on the BOMBE simulation, at this point only the intel compiler will be able to compile the code to fully utilise the additions to the code that parallelization require. I have chosen C++ as my base language mainly because it has an expansive library of parlallel tools, this is the intel comiler, and it is also a language I am very familiar with, having used it extensively throughout my academic career.

B Development and Maintenance

The project will be designed and produced in an agile manner, it will also incorporate a modular progression. Each system, enigma and BOMBE, will be produced independent of one another. This will allow verification of correctness to be acquired at several different stages throughout the implementation and not as a final progression. The Engima machine will be designed and implemented first, it will then be verified. Following that the standard BOMBE will be implemented and using the already verified engima as a means to show to correctness of the BOMBE, it will also be tested on known (ciphertext, plaintext) pairs as well. This will lead onto the final stage, the parallelisation and testing of the BOMBE, it will continuously be verified during this process to produce a fast and correct BOMBE implementation that will then be tested against it unparallelised counterpart, leading to a conclusion.

The maintenance of this system will most likely not extend far past the deadline date, as by this point both systems will be fully working in the latest version of the intel compilers and will not be changed for any updated version of the compilers. The systems will work on any operating system that can install the latest version of the intel compilers at the time of the deadline.

C System Architecture and Design

The system will consist of two stand alone C++ programs, each will have an input file that will tune parameters and inputs. Each system will have a shell script that will verify the inputs, recompile on the current system, tune the parameters and then run the system with these settings. This will be the same for both the enigma and BOMBE. The settings file will be in the form of a basic .txt file, this will have a standard format. The settings file will contain; which rotors to use in which positions, what their offset is, the plugboard and reflector mappings, as well as the plaintext.

Neither the enigma or the BOMBE will allow for continuous input of characters as the plaintext or ciphertext respectively as this would be the dictate of the speed of the system, the char-

acters cannot be input fast enough for there to be no machine idle time between two adjacent characters. Thus rendering a timed run completely dependant on how fast the user can type and not on the system itself.

D   Functional Requirements

Table 1: Functional Requirements

Unique ID	Deliverable	Description	Priority
DL1	Simulate the Enigma Machine	This is a description of a thing	High
DL2	Simulate the Bombe	This is a description of a thing	High
DL3	Show the correctness of the Enigma Machine	This is a description of a thing	Medium
DL4	Show the correctness of the BOMBE	This is a description of a thing	Medium
DL5	Evaluate average computation time over set inputs for both Enigma and BOMBE	This is a description of a thing	High
DL6	Parallelise The BOMBE and evaluate average computation time	This is a description of a thing	High

E   Evaluation

Each system will be evaluated individually first then in conjunction to one another. To evaluate the enigma, it will be tested over known (plaintext,ciphertext) pairs to show output is valid. The BOMBE is tested over known and unknown ciphertext, the output of the known ciphertext will then be reencrypted using the settings that were used to decrypt it and they will be compared to show correctness. Once correctness has been established then decryption time will compared between the parallelised and standard BOMBE over the same ciphertext.

F   References