BONDED SEQUENTIAL AND PARALLEL INSERTION-DELETION SYSTEMS IN FORMAL LANGUAGE THEORY

AHMAD FIRDAUS YOSMAN

UNIVERSITI TEKNOLOGI MALAYSIA

BONDED SEQUENTIAL AND PARALLEL INSERTION-DELETION SYSTEMS IN FORMAL LANGUAGE THEORY

AHMAD FIRDAUS YOSMAN

A thesis submitted in fulfilment of the
requirements for the award of the degree of
Doctor of Philosophy

Faculty of Science
Universiti Teknologi Malaysia

NOVEMBER 2022

ACKNOWLEDGEMENT

I am indebted to so many people for the support and encouragement they have given throughout my research. I may not be able to name all of them individually, but they should know that I am thinking of them while writing this.

Personally, I would like to express my sincerest gratitude to my esteemed supervisor, Assoc. Prof. Dr. Fong Wan Heng for her selfless, tireless, and timeless guidance through all my years as her student. I truly believe that my life as a research student would not be so successful without her as my mentor. Furthermore, I would like to thank my co-supervisor, Dr. Hazzirah Izzati Mat Hassim for her openness and objectivity, and for always having meaningful insight to problems in research. Every discussion we had provided new points of view and approaches.

I would like to thank my friends and fellow members of Applied Algebra and Analysis Group (AAAG) for the support and companionship throughout my research. Though our in-person meetings were extremely limited due to the pandemic, I am proud that we had successfully organized several students' seminars, independent of our supervisors. I wish them all the best in their research endeavors.

A special mention goes out to the librarians of UTM, the unsung heroes during the COVID-19 pandemic. From the very beginning of Movement Control Order (MCO), they were ready to assist me with questions and requests for interlibrary loans (ILL), limited access articles, and extensions of my book loans. I owe much of my success to their efficiency and competence.

Finally, my deepest, most sincerest love and appreciation goes out to my parents and lovely brothers.

ABSTRACT

Insertion and deletion are set operations that act upon structures in a predetermined fashion to create language generating devices in formal language theory, a field that involves the mathematical aspects of logic and set theory. Earlier, insertion and deletion systems that simulate the process of chemical bonding between atoms had been introduced, called bonded insertion and deletion systems. Still, these bonded systems could not generate up to recursively enumerable languages, nor could they generate abstract families of languages. Thus, the aim of this research is to introduce bonded insertion-deletion systems by combining the two aforementioned systems. Additionally, the generative power and closure properties of the sequential and parallel variants of bonded insertion-deletion systems are determined by simulating computationally complete grammars and by direct proofs, respectively. From there, graphs are constructed to visualize the mechanism of the new bonded systems by using concepts in graph theory. Furthermore, a new cryptosystem that involves both encoding and encryption is constructed using bonded sequential and parallel insertion-deletion systems based on an American Standard Code for Information Interchange (ASCII) table framework involving a bonding alphabet. From this research, it has been shown that the generative power of the new systems is up to the family of recursively enumerable languages, hence capable of generating abstract families of languages. The graphs of bonded systems and language generating graphs have been constructed such that the language of the graphs is equivalent to the language of the systems. On the other hand, the newly constructed cryptosystem using bonded insertion-deletion systems is able to confidentially store or transfer information. The findings in this research are important to advance studies in the field of DNA computing by effectively increasing the generative power of previous bonded systems, visualizing the mechanism of derivations of bonded systems, as well as providing an alternative method for information security with a nexus of bio-inspired cryptosystems.

ABSTRAK

Penyisipan dan pengguguran merupakan operasi set yang bertindak ke atas struktur secara tertentu untuk mencipta peranti penjanaan bahasa dalam teori bahasa formal, suatu bidang yang melibatkan aspek-aspek matematik logik dan teori set. Sebelum ini, sistem-sistem penyisipan dan pengguguran yang menyelakukan proses ikatan kimia antara atom telah diperkenalkan, dipanggil sistem penyisipan dan pengguguran terikat. Namun, sistem-sistem terikat tersebut tidak berupaya menjana sehingga bahasa enumerasi rekursif atau menjana keluarga bahasa abstrak. Oleh itu, tujuan penyelidikan ini ialah untuk memperkenalkan sistem penyisipan-pengguguran terikat dengan menggabungkan kedua-dua sistem tersebut. Selain itu, kuasa penjanaan dan sifat tertutup varian berjujukan dan selari sistem penyisipan-pengguguran terikat, masing-masing dikenalpasti dengan mensimulasikan tatabahasa pengiraan lengkap dan secara pembuktian terus. Justeru, graf dibina untuk menggambarkan mekanisme sistem-sistem terikat baharu menggunakan konsep teori graf. Tambahan lagi, sebuah kriptosistem baharu yang melibatkan kedua-dua pengekodan dan penyulitan dibina menggunakan sistem penyisipan-pengguguran berjujukan dan selari terikat berdasarkan rangka kerja jadual Kod Piawai Amerika untuk Pertukaran Maklumat (ASCII) yang melibatkan abjad mengikat. Daripada penyelidikan ini, kuasa penjanaan sistem-sistem baharu telah ditunjukkan mencecah keluarga bahasa enumerasi rekursif, justeru berupaya menjana keluarga bahasa abstrak. Graf sistem terikat dan graf penjanaan bahasa telah dibina supaya bahasa graf setara dengan bahasa sistem-sistem tersebut. Di samping itu, kriptosistem yang baharu dibina menggunakan sistem penyisipan-pengguguran terikat mampu menyimpan atau memindahkan maklumat secara sulit. Hasil dapatan daripada penyelidikan ini penting untuk memajukan kajian dalam bidang pengkomputeran DNA dengan meningkatkan kuasa penjanaan sistem-sistem terikat terdahulu secara berkesan, menggambarkan mekanisme penerbitan sistem-sistem terikat, dan juga memberi kaedah alternatif untuk keselamatan maklumat dengan penghubungan kriptosistem berilhamkan biologi.

TABLE OF CONTENTS

	TITLE	PAGE
	DECLARATION	ii
	DEDICATION	iii
	ACKNOWLEDGEMENT	iv
	ABSTRACT	v
	ABSTRAK	vi
	TABLE OF CONTENTS	vii
	LIST OF TABLES	x
	LIST OF FIGURES	xi
	LIST OF ABBREVIATIONS	xii
	LIST OF SYMBOLS	xiv
	LIST OF APPENDICES	xvi
CHAPTER 1	**INTRODUCTION**	1
1.1	Introduction	1
1.2	Research Background	2
1.3	Problem Statement	3
1.4	Objectives of the Research	5
1.5	Scope of the Research	5
1.6	Significance of the Research	6
1.7	Research Methodology	6
1.8	Organization of the Thesis	9
1.9	Conclusion	13
CHAPTER 2	**LITERATURE REVIEW**	15
2.1	Introduction	15
2.2	Historical Background of Insertion Deletion Systems	15
2.3	Fundamental Concepts in Formal Language Theory	19
2.3.1	Preliminaries	19
2.3.2	Grammars	24
2.3.3	Lindenmayer Systems (L-systems)	27
2.3.4	Insertion and Deletion	32

2.4	Bonded Systems	37
2.5	Graph Theory	43
2.6	Abstract Machines	48
2.7	Cryptology	52
2.8	Conclusion	54
CHAPTER 3	**BONDED SEQUENTIAL INSERTION-DELETION SYSTEMS**	57
3.1	Introduction	57
3.2	Definition and Examples of Bonded Sequential Insertion-Deletion Systems	58
3.3	Generative Power of Bonded Sequential Insertion-Deletion Systems	63
3.4	Closure Properties of Bonded Sequential Insertion-Deletion Systems	71
3.5	Conclusion	78
CHAPTER 4	**BONDED PARALLEL INSERTION-DELETION SYSTEMS**	79
4.1	Introduction	79
4.2	Definition and Examples of Bonded Parallel Insertion-Deletion Systems	80
4.3	Generative Power of Bonded Parallel Insertion-Deletion Systems	86
4.4	Closure Properties of Bonded Parallel Insertion-Deletion Systems	90
4.5	Conclusion	99
CHAPTER 5	**GRAPHS OF BONDED INSERTION-DELETION SYSTEMS**	101
5.1	Introduction	101
5.2	Graphs of Bonded Sequential Insertion-Deletion Systems	101
5.3	Graphs of Bonded Parallel Insertion-Deletion Systems	125
5.4	Conclusion	147

CHAPTER 6 A CRYPTOSYSTEM USING BONDED INSERTION-DELETION SYSTEMS 149

6.1 Introduction 149

6.2 Bonded Cyptosystems 150

6.2.1 Encoding into DNA Bases 151

6.2.2 Encryption using Bonded Insertion-Deletion Systems 153

6.3 Increasing the Difficulty of Bonded Cyptosystems 161

6.3.1 Short Messages 162

6.3.2 More Insertion and Deletion Rules 162

6.3.3 Shuffling 165

6.4 Conclusion 167

CHAPTER 7 SUMMARY AND RECOMMENDATIONS 169

7.1 Summary 169

7.2 Recommendations 172

REFERENCES 175

LIST OF PUBLICATIONS 181

LIST OF TABLES

TABLE NO.	TITLE	PAGE
2.1	Past Research on Insertion-Deletion Systems	18
5.1	Number of k and m Insertions with Respect to n	123
6.1	Example of an Encoding Key for Bonded Cryptosystems	151
6.2	Shuffle 1 Encoding Key for Bonded Cryptosystems	166
7.1	Results on Closure Properties of Bonded Insertion-Deletion Systems	171

LIST OF FIGURES

FIGURE NO.	TITLE	PAGE
1.1	Flow chart of the research	9
1.2	Organization of the thesis	12
2.1	Hierarchy of L-systems	30
2.2	Hierarchy of Chomsky and L-systems	31
2.3	Hierarchy of language families	43
2.4	Graph of GCID Π_1	46
2.5	A finite automaton	48
2.6	Deterministic finite automaton A_1	49
2.7	Nondeterministic finite automaton A_2	50
5.1	Graph of bSINSDEL-system σ_0, Γ_{σ_0}	105
5.2	Language generating graph of bSINSDEL-system σ_0, $\Gamma(L(\sigma_0))$	108
5.3	Graph of bSINSDEL-system σ_1, Γ_{σ_1}	110
5.4	Language generating graph of bSINSDEL-system σ_1, $\Gamma(L(\sigma_1))$	112
5.5	Graph of bSINSDEL-system σ_2, Γ_{σ_2}	115
5.6	Language generating graph of bSINSDEL-system σ_2, $\Gamma(L(\sigma_2))$	117
5.7	Graph of bSINSDEL-system σ_3, Γ_{σ_3}	120
5.8	Language generating graph of bSINSDEL-system σ_3, $\Gamma(L(\sigma_3))$	121
5.9	Graph of bPINSDEL-system π_1, Γ_{π_1}	130
5.10	Language generating graph of bPINSDEL-system π_1, $\Gamma(L(\pi_1))$	132
5.11	Graph of bPINSDEL-system π_2, Γ_{π_2}	134
5.12	Language generating graph of bPINSDEL-system π_2, $\Gamma(L(\pi_2))$	136
5.13	Graph of bPINSDEL-system π_3, Γ_{π_3}	139
5.14	Language generating graph of bPINSDEL-system π_3, $\Gamma(L(\pi_3))$	144

LIST OF ABBREVIATIONS

ASCII	-	American Standard Code for Information Interchange
BAMNN	-	Bidirectional Associative Memory Neural Network
bPDEL-system	-	Bonded Parallel Deletion System
bPINS-system	-	Bonded Parallel Insertion System
bPINSDEL-system	-	Bonded Parallel Insertion-Deletion System
bSDEL-system	-	Bonded Sequential Deletion System
bSINS-system	-	Bonded Sequential Insertion System
bSINSDEL-system	-	Bonded Sequential Insertion-Deletion System
CDMB	-	Central Dogma of Molecular Biology
CRISPR	-	Clustered Regularly Interspaced Palindromic Repeat
CF	-	Context-free Grammar
CS	-	Context-sensitive Grammar
D0L-system	-	Deterministically Interactionless Lindenmayer System
DT0L-system	-	Deterministically Tabled Interactionless Lindenmayer System
DFA	-	Deterministic Finite Automaton
DNA	-	Deoxyribonucleic Acid
ED0L-system	-	Extended Deterministically Interactionless Lindenmayer System
EDT0L-system	-	Extended Deterministically Tabled Interactionless Lindenmayer System
E0L-system	-	Extended Interactionless Lindenmayer System
ET0L-system	-	Extended Tabled Interactionless Lindenmayer System
GCID	-	Graph-controlled Insertion-Deletion System
GUI	-	Graphical User Interface
HG Nord	-	*Heeresgruppe Nord* (German for Army Group North)

0L-system	-	Interactionless Lindenmayer System
L-system	-	Lindenmayer System
IL-system	-	Lindenmayer System with Interactions
MATINSDEL	-	Matrix Insertion-Deletion System
NFA	-	Nondeterministic Finite Automaton
PG0L	-	Propagating Graph 0L-system
REG	-	Regular Grammar
T0L-system	-	Tabled Interactionless Lindenmayer System
RE	-	Unrestricted Grammar
WWI	-	World War I
WWII	-	World War II

LIST OF SYMBOLS

α	-	Axiom word
η	-	Bonded deletion system
γ	-	Bonded insertion system
π	-	Bonded parallel insertion-deletion system
σ	-	Bonded sequential insertion-deletion system
$\mathcal{B}_\Sigma$	-	Bonding alphabet
$\|S\|$	-	Cardinality of the set S
$\mathbb{Z}_n$	-	Cyclic group of order n
$\Rightarrow$	-	Derivation relation
β	-	Derived word
$\in$	-	Element of
$\emptyset$	-	Empty set
λ	-	Empty word
$\mathcal{L}(X)$	-	Family of languages generated by X
Γ	-	Graph
Γ_{π_i}	-	Graph of bonded parallel insertion-deletion system π_i
Γ_{σ_i}	-	Graph of bonded sequential insertion-deletion system σ_i
$>$	-	Greater than
$\geq$	-	Greater than or equal to
$\subseteq$	-	Inclusion
δ_i	-	Insertion or deletion rule
$\cap$	-	Intersection
a^*	-	Kleene closure on the symbol a

$\|w\|$	-	Length of the word w
$<$	-	Less than
$\leq$	-	Less than or equal to
$\rightarrow$	-	Maps to
$\not\subseteq$	-	Not a subset
$\neq$	-	Not equal to
$\|\|\Gamma\|\|$	-	Number of edges of a graph Γ
$\subset$	-	Proper inclusion
$\Rightarrow^*$	-	Reflexive and transitive closure of $\Rightarrow$
$\mathcal{B}_\Sigma^{\circledast}$	-	Set of all balanced words built from letters of $\mathcal{B}_\Sigma$
$\mathbb{Z}$	-	Set of all integers
$\mathbb{N}$	-	Set of all natural numbers i.e. integers ≥ 1
$\mathcal{B}_\Sigma^*$	-	Set of all well-formed words built from letters of $\mathcal{B}_\Sigma$
Σ^*	-	Set of all words over the alphabet Σ
Σ	-	Set of letters called the alphabet
$\sigma(a)$	-	Substitution of the letter a
$\Rrightarrow$	-	Transitions to
$\cup$	-	Union

LIST OF APPENDICES

APPENDIX	TITLE	PAGE
A	List of Publications	181

CHAPTER 1

INTRODUCTION

1.1 Introduction

The field of formal language theory has been around ever since the early 20th century when Thue studied the properties of patterns in extremely long words [1]. Though Thue and Post's work was fairly unrecognized for decades due to the limited and near-reaching resources back in the 1900s, they are regarded as the pioneers of the construction of formal grammars [1]. The mathematical aspect of formal language theory bears its roots from mathematical logic, algebra, and combinatorics, where set operations, properties, and structures are used as the building blocks of language generating devices. Throughout the years, formal language theory has evolved from a synthesis of mathematics, computing and linguistics into the very foundation of theoretical computer science and deoxyribonucleic acid (DNA) computing.

The set operations initially applied in the formation of formal grammars were concatenation and also quotient, such that the addition or removal of letters from words are done only at the extremities of words. Kari [2] in her groundbreaking work provided generalizations for the aforementioned operations, namely insertion and deletion, respectively. Now, letters can be added into or removed from anywhere in a word. By applying the operation of Kleene closure, the continual iteration of the operations can be achieved, thus creating systems of language generation. These systems have since been explored by researchers from a plethora of backgrounds, mainly attracting those working in the area of DNA computing, where the formalization of the biological phenomena of recombinant processes of DNA is of utmost priority. In [3], Fong et al. introduced insertion systems that simulate the biochemical process of chemical bonding, called bonded insertion systems. By appending integers on to the left and right side of each letter, so-called bonds now affect the permissible positions for insertion. These systems showed immense potential in generating families of languages of a higher power in the hierarchy of languages. Later, bonded deletion systems were introduced in [4] along with their generative power.

In this research, the bonded systems of insertion and deletion are advanced by combining them to create new insertion-deletion systems, where their properties and computing capabilities are investigated. From there, the graphs of these systems are constructed in addition to being used to introduce a new cryptographic system (shortly, cryptosystem).

1.2 Research Background

Bonded systems were first introduced by Fong et al. [3], where bonded sequential insertion and parallel insertion systems were defined. These systems were constructed to not only utilize the operations of insertion to model recombinant DNA but to also take into account the sharing and transferring of valence electrons in the formation or breaking of chemical bonds during the recombination process. In their work, the generative power of the bonded insertion systems were shown such that the relation $\mathcal{L}(\mathrm{REG}) \subset \mathcal{L}(\mathrm{bSINS}) \subset \mathcal{L}(\mathrm{bPINS}) \subset \mathcal{L}(\mathrm{E0L})$ holds; where $\mathcal{L}(\mathrm{REG})$ represents the family of languages generated by regular grammars, $\mathcal{L}(\mathrm{bSINS})$ represents bonded sequential insertion systems, $\mathcal{L}(\mathrm{bPINS})$ represents bonded parallel insertion systems, and $\mathcal{L}(\mathrm{E0L})$ represents extended interactionless Lindenmayer systems. Two variants of parallel insertion systems, called bonded Indian parallel insertion systems and bonded uniformly parallel insertion systems, had also been introduced along with their generative power in [5]. The closure properties of bonded insertion systems were investigated in [6]. In this case, the generative power of a system refers to the capability of a system in generating the family of languages contained within or equivalent to a family of languages in the Chomsky hierarchy.

However, the Chomsky hierarchy deals with sequential productions or derivations. Thus, the generative power of systems with parallel derivations are compared to the families of languages in the hierarchy of Lindenmayer systems instead due to their parallel nature.

Naturally, the concept of bonded systems was extended to the deletion operation, which included bonded sequential and parallel deletion systems [4]. However, the generative power of these systems were significantly higher than bonded insertion systems, where both the bonded deletion systems could generate up to recursively enumerable languages, albeit with some caveats.

Therefore, in this research, which is a continuation of the work in [4], bonded sequential and parallel insertion-deletion systems are introduced, such that both insertion and deletion operations acting on the bonding alphabet are included in one construct. Subsequently, the generative power and closure properties of the newly introduced systems are determined.

In the past, graphs of grammars had been introduced to model how the production rules of grammars work [7]. This work inspired the next results in this research, where the graphs of bonded sequential and parallel insertion-deletion systems are introduced to illustrate the mechanism of the derivation of these systems.

Lastly, inspired by Turing's contributions to cryptanalysis during World War II (WWII), the newly introduced bonded systems are then utilized in the construction of a new cryptosystem involving encoding and encryption, called bonded cryptosystems.

1.3 Problem Statement

Previously, bonded sequential and parallel insertion systems alone do not generate up to recursively enumerable languages, where bonded sequential insertion systems could generate only up to context-free languages while bonded parallel insertion systems only up to extended interactionless Lindenmayer systems; nor do they generate abstract families of languages. Therefore, bonded insertion-deletion systems are constructed to increase the generative power and to determine the closure under language operations.

Furthermore, the graphs of bonded systems have yet to be introduced in past research, thus allowing inconsistencies in the mechanism of derivations of the bonded systems that lead to the generation of languages other than the one desired. Hence, in this research, graphs of bonded insertion-deletion systems are constructed to provide a visual illustration of the mechanism of derivations of these systems.

Lastly, the supposed myriad of applications of bonded systems have not been specified before. This leads to the introduction of a new crpytosystem in this research that involves encoding and encryption which utilize the newly introduced bonded systems.

Thus, the findings of this research can answer the following questions:

1. How to increase the generative power of bonded insertion systems and bonded deletion systems?
2. What is the generative power of bonded sequential and parallel insertion-deletion systems with regards to the hierarchy of languages?
3. What are the closure properties of bonded sequential and parallel insertion-deletion systems?
4. How to visualize the mechanism of derivations of bonded sequential and parallel insertion-deletion systems?
5. How to encode and encrypt messages using bonded sequential and parallel

insertion-deletion systems?

1.4 Objectives of the Research

The objectives of this research are:

1. To introduce bonded sequential and parallel insertion-deletion systems with integer contexts.
2. To determine the generative power of bonded sequential and parallel insertion-deletion systems by simulation of language generating devices.
3. To determine the closure properties of bonded sequential and parallel insertion-deletion systems with regards to language operations.
4. To construct graphs to visualize the mechanism of derivations of bonded sequential and parallel insertion-deletion systems.
5. To construct bonded cryptosystems that involve encoding and encryption using bonded sequential and parallel insertion-deletion systems.

1.5 Scope of the Research

This research involves bonded insertion systems and bonded deletion systems of the sequential and parallel variant. Additionally, fundamental concepts in formal language theory, graph theory, and automata theory are used to simulate language generating devices and to construct graphs of bonded sequential and parallel insertion-deletion systems. The language operations considered are those required to decide whether the systems generate abstract families of languages, which are union, concatenation, concatenation closure, λ-free concatenation closure, intersection, and substitution. The families of languages in the Chomsky hierarchy and Lindenmayer systems are used to determine the generative power of the languages generated by bonded sequential and parallel insertion-deletion systems. Once determined, the generative power of the bonded sequential and parallel insertion-deletion systems is used to verify the results on the closure properties. Basic concepts in encoding and encryption are used in the final part of this research to introduce bonded cryptosystems, which are systems that utilize bond words to encode text into the DNA alphabet consisting of the letters A, T, G, and C before encrypting the encoded text using bonded insertion-deletion systems.

1.6 Significance of the Research

Biomathematics has seen exponential growth in recent years as more researchers realize the many benefits of DNA computing, which has now expanded beyond DNA recombination into genetic engineering and advanced theoretical computer science. The direct implications of this research is in the improvement of the accuracy of abstract DNA computing by bonded sequential and parallel insertion-deletion systems due to the combination of both operations. Besides that, newly introduced graphs of the systems aid in visualizing the mechanism of derivations of the systems. Furthermore, the newly introduced encoding and encryption systems in this research provide an alternative method to securing confidential information using molecular biology by presenting or storing digital information as a DNA string. Organically, this research will further propel studies in computing, thus bringing the world closer to the realization of true DNA computers through extensive interdisciplinary research.

1.7 Research Methodology

This research begins with the literature review on the fundamental concepts in formal language theory, bonded systems, graph theory, abstract machines, and cryptology. Numerous materials from a variety of online and physical sources have been studied to ensure an all-encompassing view of the concepts. The literature review clarified the relationship between formal language theory and graph theory, explained the mechanism of abstract machines in language generating devices, specifically Turing machines, and linked bonded systems to cryptology.

From there, bonded sequential and parallel insertion-deletion systems are introduced by combining the concepts of bonded insertion systems and bonded deletion systems, where the operations of insertion and deletion act upon the bonding alphabet. The formal definitions of these systems have been constructed according to conventional formal language theory and examples have been presented to demonstrate the new systems. In both bonded insertion-deletion systems, the insertion and deletion rules have been loosened to include unbalanced rules, which are extremely useful in increasing the generative power of the systems.

Next, the generative power of bonded sequential and parallel insertion-deletion systems has been determined. Firstly, the generative power of bonded sequential insertion-deletion systems (bSINSDEL-systems) has been determined by simulations of the production rules of a grammar and an insertion-deletion system with matrix control, called a matrix insertion-deletion system, where both generate the family of recursively enumerable languages. On the other hand, the generative power of bonded parallel insertion-deletion systems (bPINSDEL-systems) has also been determined by simulation of a Turing machine.

Once the generative power of the systems had been determined, the next step is to determine the closure properties of bSINSDEL-systems and bPINSDEL-systems with regards to language operations. Here, the family of languages generated by bSINSDEL-systems and bPINSDEL-systems, denoted by $\mathcal{L}(\mathrm{bSINSDEL})$ and $\mathcal{L}(\mathrm{bPINSDEL})$ respectively, are subjected to language operations of union, concatenation, concatenation closure, λ-free concatenation closure, and intersection with regular languages. Additionally, $\mathcal{L}(\mathrm{bSINSDEL})$ is subjected to substitution while $\mathcal{L}(\mathrm{bPINSDEL})$ is subjected to homomorphism and inverse homomorphism.

Furthermore, graphs of bonded insertion-deletion systems are introduced to visualize the mechanism of bonded sequential and parallel insertion-deletion systems. By representing all the letters in a word as vertices, the derivation relation as edges, and the insertion and deletion rules as edge labels, graphs of bonded systems have been introduced. To depict the derivations more accurately, the language generating graphs of bonded systems have also been introduced by including restrictions on the edges.

Lastly, an encoding and encryption system that utilizes the definition of bSINSDEL-systems and bPINSDEL-systems is introduced. First, alphanumeric messages are encoded into bond words over the DNA alphabet: A, T, G, and C, using an encoding key, which is a table containing all 26 letters of the English alphabet, arranged into four columns and ten rows. The columns are labeled with the letters A, T, G, and C and represent the cipher letter of the original character in the plain text. The rows are labeled with the numbers 1-10 and represent the right bonds of the letter. The left bonds of each letter are equal to the right bonds of the previous letter, whereas the leftmost and rightmost bonds are 0. Thus, by encoding a message using the table, a bond word is formed. Next, the letters are either added and removed using a bSINSDEL-system or a bPINSDEL-system as a form of encryption. This way, messages can be safely stored or transfered in the form of DNA strings that can only be decrypted and read with the appropriate decryption key. Here, the decryption key is constructed according to the definition presented, which includes the number of derivations k, the rules used, and the right bonds of the encrypted word.

The flow of the research is depicted in Figure 1.1

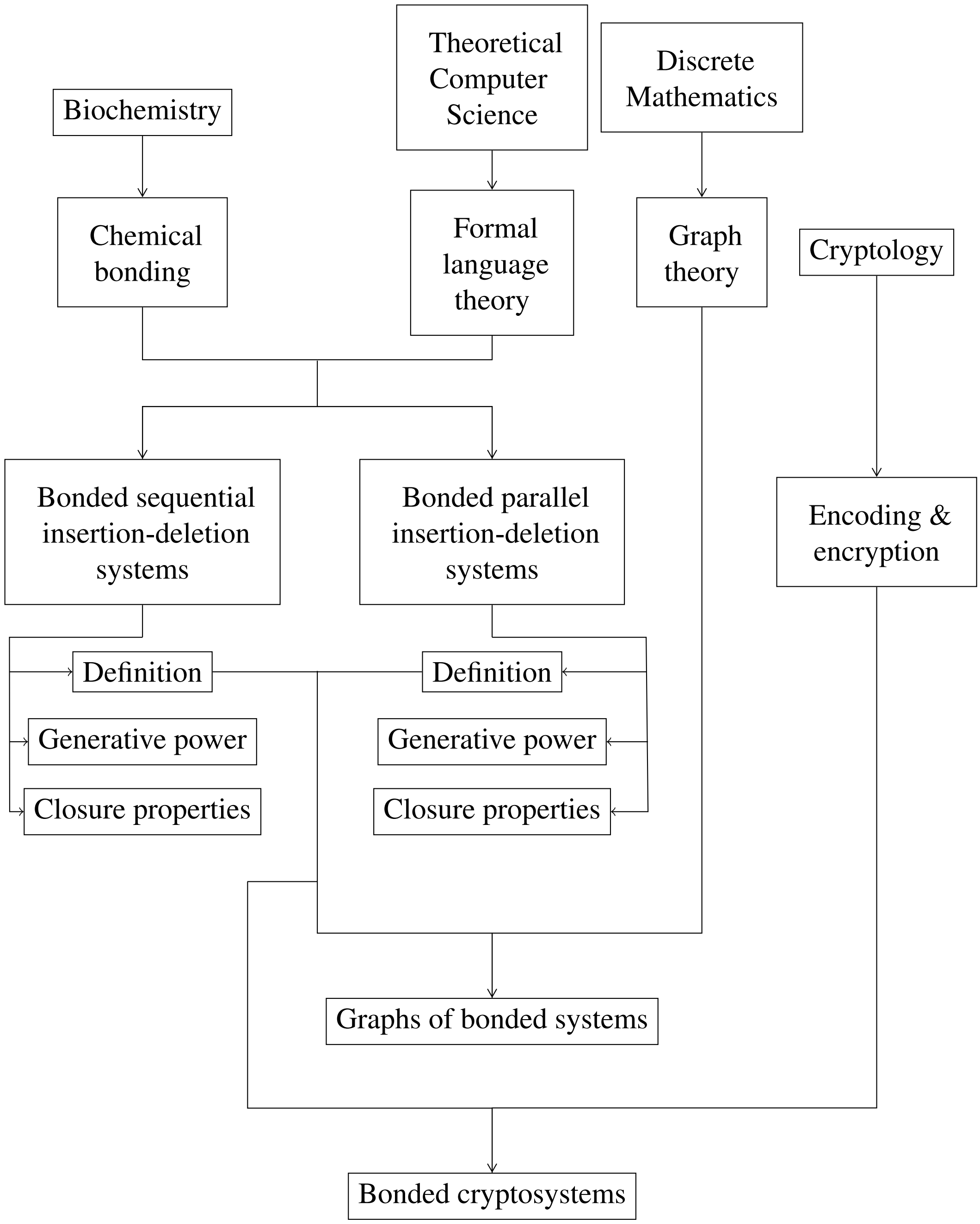

Figure 1.1 Flow chart of the research

1.8 Organization of the Thesis

The thesis is organized as follows: In Chapter 1, the introduction to the research is provided, which includes the research background, problem statement, objectives of the research, scope of the research, significance of the research, and research methodology. Lastly, the organization of the thesis is presented.

In Chapter 2, the literature review of the research is presented, beginning with the historical background of insertion-deletion systems explained chronologically through the progress and findings in the research area. Next, some fundamental concepts of formal language theory are presented, such as the definitions and notations of important terms and operations used in this research. Besides that, the concept of bonded systems is presented along with examples to further clarify the mechanism of these systems. In addition, basic concepts in graph theory and abstract machines are provided. An introduction and brief history in cryptology is provided in the final part of this chapter.

In Chapter 3, results on bonded sequential insertion-deletion systems are presented. Firstly, the formal definition of the systems is provided, followed by some examples. Once the definition and mechanism have been established, the generative power of the systems is shown in the next part of the chapter by two simulations: the production rules of a grammar in special Geffert normal form and the production rules of a matrix insertion-deletion system. In the end, results on the closure properties of the family of languages generated by bonded sequential insertion-deletion systems are presented.

Meanwhile, in Chapter 4, results on bonded parallel insertion-deletion systems are presented. The chapter begins with the formal definition of the systems accompanied by some examples. Following that, the results on the generative power by simulation of a Turing machine and closure properties are presented.

In Chapter 5, graphs of bonded systems are introduced, where the components of bonded sequential and parallel insertion-deletion systems are represented by components of graphs. Furthermore, language generating graphs of bonded systems are introduced to ensure that only the desired language is generated by the graph of bonded systems. Examples and figures are provided to illustrate the graphical representations.

In Chapter 6, a new method to encoding and encryption called bonded cryptosystems is introduced. The chapter begins with an overview of the idea to introduce such a system. The next part of the chapter dives into the construction of the bonded cryptosystems, beginning with the encoding into a DNA alphabet, followed up by encryption using bonded insertion-deletion systems. The description of the decryption key is also provided. The chapter ends with some methods to increase the difficulty of the bonded cryptosystems.

In the end, the summary of this research and recommendations for future work are provided in Chapter 7.

The organization of the thesis is shown in Figure 1.2.

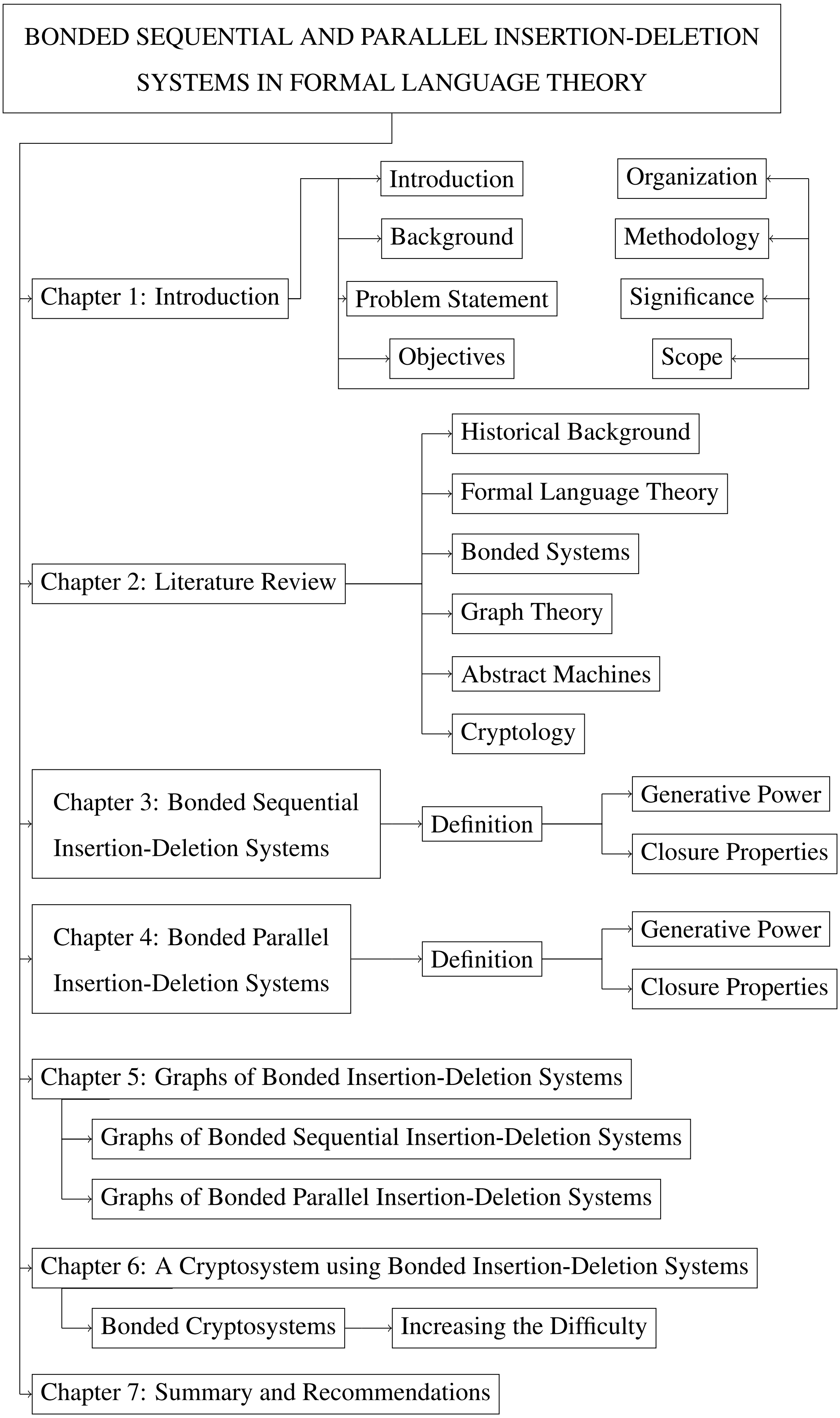

Figure 1.2 Organization of the thesis

1.9 Conclusion

In this chapter, the framework of the research has been explained in intricate detail. The introduction and research background have provided insight into the conception of bonded insertion-deletion systems. The challenges faced by previous researchers and questions solved through the work in this research have been outlined in the problem statement. Furthermore, the objectives and scope of the research have been presented, as well as the significance of the research being highlighted. In addition, the research methodology has been elaborated with details of the steps taken to arrive at the results in this research. Lastly, the organization of thesis has been provided to display the overall flow of the presentation of the findings of this research.

Much of what is known today in DNA computing has been the result of painstaking studies into the depths of human genetics and its mathematical formalization. With every new finding, more light is shed on the abstract occurrences in nature. By bringing research in DNA computing to its fundamental concepts, this research will enable a better understanding of how nature's natural order is kept intact after eons of evolution and adaptation. From there, applications of DNA computing can extend beyond biomathematics into fields such as cryptology as shown in this research.

REFERENCES

1. Rozenberg, G. and Salomaa, A. *Handbook of Formal Languages*. Berlin: Springer-Verlag. 1997.

2. Kari, L. *On Insertion and Deletion in Formal Languages*. University of Turku: Ph.D thesis. 1991.

3. Fong, W. H., Holzer, M., Truthe, B., Turaev, S. and Yosman, A. F. On bonded sequential and parallel insertion systems. In *8th Workshop on Non-Classical Models of Automata and Applications (NCMA), Debrecen, Hungary, August 29 – 30, 2016, Proceedings*. Österreichische Computer Gesellschaft, Austria. 2016. books@ocg.at. 163–178.

4. Yosman, A. F. *New Variants of Insertion and Deletion Systems in Formal Languages*. Universiti Teknologi Malaysia: M.Phil thesis. 2017.

5. Yosman, A. F., Holzer, M., Truthe, B., Fong, W. H. and Turaev, S. On bonded Indian and uniformly parallel insertion systems and their generative power. *Malaysian Journal of Fundamental and Applied Sciences*. 2017. 13: 769–773.

6. Holzer, M., Truthe, B. and Yosman, A. F. On bonded sequential and parallel insertion systems. *RAIRO - Theoretical Informatics and Applications*. 2018. 52(2-3-4): 127–151.

7. Nagy, B. Graphs of grammars–derivations as parallel processes. In *Computational Intelligence in Engineering*. Springer. 1–13. 2010.

8. Head, T. Formal language theory and DNA: An analysis of the generative capacity of specific recombinant behaviors. *Bulletin of Mathematical Biology*. 1987. 49(6): 737–759.

9. Adleman, L. M. Molecular computation of solutions to combinatorial problems. *Science*. 1994. 266(5187): 1021–1024.

10. Linz, P. *An Introduction to Formal Languages and Automata, 5th Ed.* Jones and Bartlett Publishers. 2011.

11. Paun, G., Rozenberg, G. and Salomaa, A. *DNA Computing: New Computing Paradigms*. Springer-Verlag Berlin Heidelberg. 1998.

12. Kari, L. Insertion and deletion of words: Determinism and reversibility. In *Mathematical Foundations of Computer Science*. Manhattan, NYC: Springer. 1992. no. 629 in LNCS. 315–327.

13. Kari, L. Generalized derivatives. *Fundamenta Informaticae*. 1993. 18(1): 61–77.

14. Kari, L., Mateescu, A., Păun, Gh. and Salomaa, A. Deletion sets. *Fundamenta Informaticae*. 1993. 19: 355–370.

15. Kari, L. Insertion operations: closure properties. *Bulletin of the EATCS*. 1993. 51: 181–191.

16. Kari, L. Deletion operations: closure properties. *International Journal of Computer Mathematics*. 1994. 52: 23–42.

17. Kari, L., Mateescu, A., Păun, Gh. and Salomaa, A. On parallel deletions applied to a word. *RAIRO – Theoretical Informatics and Applications*. 1995. 29(2): 129–144.

18. Ito, M., Kari, L. and Thierrin, G. Insertion and deletion closure of languages. *Theoretical Computer Science*. 1997. 183(2): 3–19.

19. Kari, L. and Thierrin, G. Contextual insertions/deletions and computability. *Information and Computation*. 1996. 131(1): 47–61.

20. Kari, L., Păun, Gh., Thierrin, G. and Yu, Sh. At the crossroads of dna computing and formal languages: Characterizing RE using insertion-deletion systems. In *Proceedings of 3rd DIMACS Workshop on DNA Based Computing*. 1997. 318–333.

21. Daley, M., Kari, L., Gloor, G. and Siromoney, R. Circular contextual insertions/ deletions with applications to biomolecular computation. In *SPIRE/CRIWG*. 1999. 47–54.

22. Margenstern, M., Pun, G., Rogozhin, Y. and Verlan, S. Context-free insertion-deletion systems. *Theoretical Computer Science*. 2005. 320(2): 339–348.

23. Freund, R., Kogler, M., Rogozhin, Y. and Verlan, S. Graph-controlled insertion-deletion systems. In *DCFS*. 2009.

24. Petre, I. and Verlan, S. Matrix insertion-deletion systems. *Theoretical Computer Science*. 2012. 456: 80–88. ISSN 0304-3975.

25. Kuppusamy, L., Mahendran, A. and Krishna, S. N. Matrix insertion-deletion systems for bio-molecular structures. In Natarajan, R. and Ojo, A. (Eds.). *Distributed Computing and Internet Technology*. Berlin, Heidelberg: Springer Berlin Heidelberg. 2011. 301–312.

26. Fernau, H., Kuppusamy, L. and Raman, I. On the generative capacity of matrix insertion-deletion systems of small sum-norm. *Natural Computing*. 2021. 20(4): 671–689.

27. Vu, M. and Fernau, H. Insertion-Deletion Systems with Substitutions I. In Anselmo, M., Della Vedova, G., Manea, F. and Pauly, A. (Eds.). *Beyond the Horizon of Computability*. Cham: Springer International Publishing. 2020. 366–378.

28. Vu, M. and Fernau, H. *Insertion-Deletion with Substitutions II. Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 124422. Springer, Cham. 2020.

29. Vu, M. and Fernau, H. Adding matrix control: Insertion-deletion systems with substitutions III. *Algorithms*. 2021. 14(5).

30. Fernau, H., Kuppusamy, L. and Raman, I. On path-controlled insertion-deletion systems. *Acta Informatica*. 2019. 56(1): 35–59.

31. Fernau, H., Kuppusamy, L. and Raman, I. Generative power of matrix insertion-deletion systems with context-free insertion or deletion. In *Proceedings of the 15th International Conference on Unconventional Computation and Natural Computation - Volume 9726*. Berlin, Heidelberg: Springer-Verlag. 2016. UCNC 2016. 3548.

32. Salomaa, A. *Formal Languages*. ACM Monograph Series. Academic Press, Inc. 1973.

33. Lindenmayer, A. Mathematical models for cellular interactions in development I. filaments with one-sided inputs. *Journal of Theoretical Biology*. 1968. 18(3): 280–299.

34. Lindenmayer, A. Mathematical models for cellular interactions in development II. simple and branching filaments with two-sided inputs. *Journal of Theoretical Biology*. 1968. 18(3): 300–315.

35. Vitanyi, P. M. B. Deterministic Lindenmayer languages, nonterminals and homomorphisms. *Theoretical Computer Science*. 1976. 2(1): 49 – 71.

36. Perttunen, J. and Sievnen, R. Incorporating Lindenmayer systems for architectural development in a functional-structural tree model. *Ecological Modelling*. 2005. 181(4): 479 – 491.

37. Rozenberg, G. and Salomaa, A. *Mathematical Theory of L Systems*. USA: Academic Press, Inc. 1980.

38. Vitanyi, P. M. B. Context sensitive table Lindenmayer languages and a relation to the LBA problem. *Information and Control*. 1977. 33(3): 217 – 226.

39. Dassow, J. and Paun, G. *Regulated Rewriting in Formal Language Theory*. Monographs in Theoretical Computer Science. An EATCS Series. Springer-Verlag Berlin Heidelberg. 1989.

40. Wang, Y. *The Goldbach Conjecture, 2nd Edition*. Singapore: World Scientific. 2002.

41. Nagl, M. Formal languages of labelled graphs. *Computing*. 1976. 16(1-2): 113–137.

42. Ehrig, H., Pfender, M. and Schneider, H. J. Graph-grammars: An algebraic approach. In *14th Annual Symposium on Switching and Automata Theory (swat 1973)*. 1973. 167–180.

43. Janssens, D., Rozenberg, G. and Verraedt, R. On sequential and parallel node-rewriting graph grammars. *Computer Graphics and Image Processing*. 1982. 18(3): 279–304.

44. Janssens, D., Rozenberg, G. and Verraedt, R. On sequential and parallel node-rewriting graph grammars, II. *Computer Vision, Graphics, and Image Processing*. 1983. 23(3): 295–312.

45. Fernau, H., Kuppusamy, L. and Raman, I. Descriptional Complexity of Graph-Controlled Insertion-Deletion Systems. In Câmpeanu, C., Manea, F. and Shallit, J. (Eds.). *Descriptional Complexity of Formal Systems*. Cham: Springer International Publishing. 2016. 111–125.

46. Fernau, H., Kuppusamy, L. and Raman, I. On the computational completeness of graph-controlled insertion-deletion systems with binary sizes. *Theoretical Computer Science*. 2017. 682: 100–121.

47. Diestel, R. *Graph Theory, 5th Edition*. Springer Publishing Company, Incorporated. 2017.

48. Davis, M. D. and Weyuker, E. J. *Computability, Complexity and Languages: Fundamentals of Theoretical Computer Science*. Academic Press. 1983.

49. Kahn, D. *The Codebreakers: The Story of Secret Writing*. New York, New York: The New American Library, Inc. 1973.

50. Van Dalen, D. A note on some systems of Lindenmayer. *Mathematical Systems Theory*. 1971. 5(2): 128–140.

51. Culik, II, K. and Lindenmayer, A. Parallel graph generating and graph recurrence systems for multicellular development. *International Journal of General Systems*. 1976. 3(1): 53–66.

52. Mayoh, B. H. Multidimensional Lindenmayer Organisms. In Rozenberg, G. and Salomaa, A. (Eds.). *L Systems*. Springer. 1974. *Lecture Notes in Computer Science*, vol. 15. 302–326.

53. Mandelbrot, B. B. *The fractal geometry of nature*. vol. 1. WH Freeman, New York. 1982.

54. Banal, J. L., Shepherd, T. R., Berleant, J., Huang, H., Reyes, M., Ackerman, C. M., Blainey, P. C. and Bathe, M. Random access DNA memory using Boolean search in an archival file storage system. *Nature Materials*. 2021. 20(9): 1272–1280.

55. Ugandran, I., Mahendran, A., S., A., Hamada, M., Alnumay, W. S., Ghosh, U. and Sharma, P. K. A novel cryptosystem using DNA sequencing and contextual array splicing system for medical internet of things. *Computers Electrical Engineering*. 2021. 96: 107429.

56. Hamer, D., Sullivan, G. and Weierud, F. Enigma variations: An extended family of machines. *Cryptologia*. 1998. 22(3): 211–229.

57. Ostwald, O. and Weierud, F. Modern breaking of Enigma ciphertexts. *Cryptologia*. 2017. 41: 395–421.

58. Sullivan, G. and Weierud, F. Breaking German army ciphers. *Cryptologia*. 2005. 29(3): 193–232.

59. Ryan, P. Y. A., Naccache, D. and Quisquater, J. J. (Eds.). *The New Codebreakers: Essays Dedicated to David Kahn on the Occasion of His 85th Birthday*. Lecture Notes in Computer Science. Springer Berlin Heidelberg. 2016.

APPENDIX A

LIST OF PUBLICATIONS

Some results from this research have been published in journals or presented at conferences as listed in the following.

A.1 Journals

A.1.1 Indexed Journals: 1

1. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. On Bonded Sequential and Parallel Deletion Systems. *Journal of Critical Reviews*. 2020. 17(6): 902–909. - indexed by Scopus.

A.2 Proceedings

A.2.1 Indexed Proceedings: 2

1. Fong W. H., Yosman, A. F., and Mat Hassim, H. I. Closure Properties of Bonded Sequential Insertion-Deletion Systems. *Journal of Physics: Conference Series*. 2021. 1988: 12075 1–10. - indexed by Scopus.
2. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Bonded Parallel Insertion-Deletion Systems and Their Generative Power. *AIP Conference Proceedings*.

2022. 2465: 020002 1–10. - indexed by Web of Science & Scopus.

A.2.2 Non-indexed Proceedings: 2

1. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Bonded Sequential Insertion-Deletion Systems and Their Generative Power. *Proceedings of the 8th International Graduate Conference on Engineering, Science and Humanities (IGCESH 2020)*. UTM Johor Bahru, Malaysia, 18–19 August 2020. 228-231. eISSN: 2735-055X.
2. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Characteristics of Formal Languages Generated by Bonded Systems. *Proceedings of Science and Mathematics*. UTM Johor Bahru, Malaysia, 24 March 2022. 6: 81–85. eISSN: 2756-8857.

A.3 Conferences

A.3.1 International Conferences: 4

1. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. On Bonded Sequential and Parallel Deletion Systems. ASIA International Multidisciplinary Conference 2020 (AIMC2020). UTM Johor Bahru, Malaysia. 18–19 April 2020.
2. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Bonded Sequential Insertion-Deletion Systems and Their Generative Power. 8th International Graduate Conference on Engineering, Science and Humanities (IGCESH 2020). UTM Johor Bahru, Malaysia. 18–19 August 2020.
3. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Bonded parallel insertion-deletion systems and their generative power. The 1st International Conference

on Mathematics, Statistics and Operational Research (IMSO) 2021. Universiti Malaysia Perlis, Malaysia. 12–13 July 2021.

4. Yosman, A. F., Fong W. H., and Mat Hassim, H. I. Characteristics of Formal Languages Generated by Bonded Systems. International Postgraduate Faculty of Science Symposium (IPGFS 2022). UTM Johor Bahru, Malaysia. 24 March 2022.