**DECLARATION**

I hereby declare that I carried out the work reported in this thesis in the Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), under the supervision of Daw Su Sandar Phyo. I solemnly declare that to the best of my knowledge, no part of this thesis has been submitted here or elsewhere in a previous application for award of a degree. All sources of knowledge used have been duly acknowledged.

..…………………………

5th September, 2016

HTOO YANANT KHIN

5IS-61

**APPROVAL**

This is to certify that the mini thesis titled “**HASH BASED DUPLICATE WORDS CHECKING SYSTEM**” carried out by **HTOO YANANT KHIN (5IS-61)** has been read and approved for meeting part of the requirements and regulations governing the award of the degree of Bachelor of Engineering (Information Science and Technology), Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), Myanmar.

1. Dr. Soe Soe Khaing

Pro-Rector (Academic)

University of Technology ( Yatanarpon Cyber City ) …………………

(Chairman)

1. Dr. Hnin Aye Thant

Associate Professor and Head of Information Science

Faculty of Information and Communication Technology …………………

(Member)

1. Dr. Aye Nandar Hlaing

Lecturer and Course Coordinator of Information Science

Faculty of Information and Communication Technology …………………

(Member)

1. Dr. Nyo Nyo Yee

Lecturer

Faculty of Information and Communication Technology …………………

(Member)

1. Daw Su Sandar Phyo

Assistant Lecturer

Faculty of Information and Communication Technology …………………

(Supervisor)

**ACKNOWLEDGEMENTS**

First of all, I would like to express my special thanks to Dr. Aung Win, Rector, University of Technology (Yatanarpon Cyber City), for initiating the Bachelor programme at the University of Technology (Yatanarpon Cyber City).

I would like to express my respectful gratitude to Dr. Soe Soe Khaing, Pro-rector, University of Technology (Yatanarpon Cyber City), for her kind permission to complete this mini thesis. I am thankful to her for giving enough consideration to her ideas and views.

I would like to express grateful thank to Dr. Hnin Aye Thant, Associate Professor and Head of Information Science, Faculty of Information & Communication Technology, University of Technology (Yatanarpon Cyber City ), for her kind guidance, encouragement and supervision and patience in making my thesis to complete successfully.

I am very grateful to Dr. Renu and Dr. Aye Nandar Hlaing, Course Coordinators of Information Science and Technology, Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), for her kind guidance and encouragement and supervision and patience in making my thesis to complete successfully.

I am very grateful to my supervisor, Daw Su Sandar Phyo, Faculty of Information and Communication Technology, University of Technology (Yatanarpon Cyber City), for her kind guidance and encouragement. She has been very supportive in this mini-thesis, and also guides a lot, particularly, at the level of quality of presentation.

Very Special thanks to Daw Nan Tin Zar Zar Myint, Lecturer, Head of English Department, University of Technology (Yatanarpon Cyber City) for her valuable supports from the language point of view in my mini-thesis work.

I would like to thank a lot to all my teachers for their mentoring, encouragement, and recommending this dissertation.

Finally, I am grateful to my parents and friends who specially offered strong moral and physical support, care and kindness, during the year of my mini-thesis study.

**ABSTRACT**

Open Hashing (also known as separate chaining) is a technique in which the data is not directly stored at the hash key index (k) of the Hash table. Rather the data at the key index (k) in the hash table is a pointer to the head of the data structure where the data is actually stored. In the most simple and common implementations, the data structure adopted for storing the element is a linked-list. Chained hash tables with [linked lists](https://en.wikipedia.org/wiki/Linked_list) are popular because they require only basic data structures with simple algorithms, and can use simple hash functions that are unsuitable for other methods.

The system contains many English vocabularies and the user can only type words which are being existed in database. Initially, the system will generate a word in random order. The player must enter a word; its first character must be matched with the last character of random word. The system contains minus marking scheme for duplication, mismatch and words which are not existed in database. The player will gain -2 points for duplication, -1 point for mismatch and words which are not existed in database. One mark will be gained if the player enters a correct word which is consistent with the rules of the system. The system has five levels to test how much the player has known vocabularies in English. The system uses hash table with separate chaining mechanism to store input from user. This system is implemented by using java programming language (J2SE).

**CONTENTS**

**Page**

DECLARATION i

APPROVAL ii

ACKNOWLEDGEMENTS iii

ABSTRACT iv

CONTENTS v

LIST OF FIGURES vii

**CHAPTER 1 INTRODUCTION**

1.1 Objectives 2

1.2 Analysis of Algorithm 2

1.3 Overview of the System 3

1.4 Organization of the Thesis 4

**CHAPTER 2 THEORETICAL BACKGROUND**

2.1 History of Hashing 5

2.1.1 Library Hash Functions 7

2.1.1.1 CBU Hash 7

2.1.1.2 PKP Hash 8

2.1.1.3 BUZ Hash 8

2.1.1.4 Other Hash Functions 9

2.2 Hash Table 10

2.2.1 Conflicting Requirements of Hash Table 11 2.3 Collision in Hashing 11

* 1. Collision Handling Techniques 12

2.4.1 Separate Chaining 12

#### 2.4.1.1 Separate Chaining with Linked Lists 13

2.4.1.2 Separate Chaining with List Head Cells 14

2.4.1.3 Separate Chaining with Other Structures 15

2.4.1.4 Two Most Important Chaining Techniques 15

2.4.1.4.1 Move to Front 16

* + - * 1. Exact Fit 16

2.4.2 Open Addressing 17

2.4.2.1 Linear Probing 18

2.4.2.2 Quadratic Probing 19

2.4.2.3 Double Hashing 20

2.5 Choosing a Good Hash Function 21

2.6 Perfect Hash Function 22

2.7 Some Uses of Hash Function 22

2.7.1 Using in Hash Table 23

### 2.7.2 Using in Building Caches 23

### 2.7.3 Usage in Finding Duplicate Records 23

### 2.7.4 Useful in Protecting Data 24

### 2.7.5 Usage in Finding Similar Records 24

### 2.7.6 Finding Similar Substrings 24

### 2.7.7 Geometric Hashing 25

### 2.7.8 Standard Uses of Hashing in Cryptography 25

2.8 Advantages of Hash Tables 26

2.9 Disadvantages of Hash Tables 26

2.10 Summary 28

**CHAPTER 3 SYSTEM DESIGN AND IMPLEMENTAION**

3.1 System Design 29

3.2 System Flow 30

3.3 System Implementation 31

3.4 Output of the System 47

**CHAPTER 4 CONCLUSION**

4.1 Limitations of the System 48

4.2 Further Extension 49

**REFERENCES** 50

**LIST OF FIGURES**

**Figure Page**

2.1 Sample Code of CBU Hash 7

2.2 Sample Code of PKP Hash 8

2.3 Simple Code of PKP Hash with Inner Loop 8

2.4 Simple Code of BUZ Hash 9

2.5 Simple Code of Other Hash Functions 10

2.6 Hash Table 11

2.7 Collision of Two Keys in Hashing : h(Ki) = h(Kj) 12

2.8 Hash Collision Resolved by Separate Chaining 13

2.9 Hash Collision by Separate Chaining with Head Records in the Bucket Array 14

2.10 Hash Table Applying Exact Fit Technique 17

2.11 Hash Collision Resolved by Linear Probing 19

2.12 Hash Collision Resolved by Quadratic Probing 20

2.13 Hash Collision Resolved by Double Hashing 21

2.14 A Cryptographic Hash Function at Work 26

3.1 System Design 29

3.2 System Flow 30

3.3 Intro Page 31

3.4 Home Page 32

3.5 Get Info Page 33

3.6 Rules Page 34

3.7 Recent Player Page 35

3.8 Exit Dialog 35

3.9 Old Player Dialog 36

3.10 Invalid Player Dialog 36

3.11 New Player Dialog 37

3.12 Invalid Player Dialog 37

3.13 Invalid Player Dialog of Existing Name 37

3.14 Invalid Player Dialog of Mismatched Password 38

3.15 Levels Page of New Player 38

3.16 Levels Page for Old Player 39

3.17 Initial State for Level One 40

3.18 State for Invalid Word in Level One 41

3.19 State for Valid Word in Level One 42

3.20 Level Two Page 43

3.21 Level Three Page 43

3.22 Level Four Page 44

3.23 Level Five Page 45

3.24 Level One Completed Dialog 45

3.25 Level One Fail Dialog 46

3.26 Congratulation Dialog for Level Five 46