



PROJECT REPORT

# English to Indonesian Translator Using Rule Based and Example Based Machine Translation Algorithm

Ferdinand Budi Kurniawan

09.02.0004

[2013]



FACULTY OF COMPUTER SCIENCE

SOEGIJAPRANATA CATHOLIC UNIVERSITY

Jl. Pawiyatan Luhur IV/1, Bendan Duwur, SEMARANG 50234

Telp. (024) 8441555 (hunting) Web: <http://www.unika.ac.id>

Email: [ikom@unika.ac.id](mailto:ikom@unika.ac.id)

|                     |                              |
|---------------------|------------------------------|
| <b>PERPUSTAKAAN</b> | NO. INV : 213 / 5 / IK / C.1 |
|                     | TGL : 7 Oktober 2013         |
|                     | PARAF :                      |



**APPROVAL AND RATIFICATION PAGE**

**PROJECT REPORT**

**English to Indonesian Translator Using Rule Based and Example Based**

**Machine Translation Algorithm**

This project report has been approved and ratified by the Dean of Faculty of  
Computer Science and Supervisor on July, 31<sup>th</sup> 2013

With approval,

Examiners,

Supervisor,

Suyanto E.A., Ir, M.Sc

NPP : 058.1.1992.116

R. Setiawan Aji Nugroho, ST., MCompIT

NPP : 058.1.2004.264

Examiners,

Examiners,

Shinta Estri Wahyuningrum, S.Si.

NPP: 058.1.2007.272

Rosita Herawati, ST., MIT

NPP : 058.1.2004.263

Examiners,

Dean of Faculty of Computer Science,

Hironimus Leong, S.Kom., M.Kom

NPP : 058.1.2007.273



Hironimus Leong, S.Kom., M.Kom

NPP : 058.1.2007.273

## STATEMENT OF ORIGINALITY

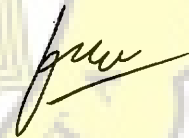
I, the undersigned:

Name : Ferdinand Budi Kurniawan

ID : 09.02.0004

Certify that this project was made by myself and not copy or plagiarize from other people, except that in writing expressed to the other article. If it is proven that this project was plagiarizes or copy the other, I am ready to accept a sanction.

Semarang, July 31<sup>th</sup> 2013



Ferdinand Budi Kurniawan

09.02.0004



## ABSTRACT

Example based machine translation (EBMT) is a method of machine translation that emphasizes the use of bilingual corpus as its knowledge base. Example based machine translation system stores bilingual corpus in a data structure that provides speed for the application to access the data. To this end, the system uses Hash Table combined with Binary Search Tree as the collision resolution.

Example based machine translation has several benefits over other methods of machine translation. The most prominent benefit of example based machine translation is the accuracy of translation it can produce to the user's input as long as the bilingual corpus contains the same sentence with the sentence inputted by users. However, this could also be the limitation of example based machine translation: limited only to the bilingual corpus the system has.

To increase the ability of the system, rule based machine translation method (RBMT) is also used as the preliminary step to analyze user's input. It analyzes the input sentence syntactically and then transfers the result to the corpus. The combination between Rule based machine translation and example based machine translation could give better result compared to machine translation that solely uses Example Based Machine Translation method.

*Keywords: example based machine translation, rule based machine translation, corpus*

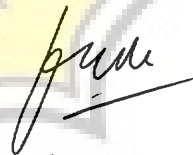
## FOREWORD

Praise to the Almighty God for all of his blessings that keep me spirited and inspired all days to finish my thesis entitled English to Indonesian Translator Using Rule Based and Example Based Machine Translation Algorithm. Also, without the role of these people, I would not be able to finish my thesis; I would like to thank to:

1. My parents and my family in Semarang and in Jakarta for all of their prayers and supports for me
2. My distinguished supervisor: Mr. Robertus Setiawan Aji Nugroho, ST., MCompIT for his endless supports, guides and suggestions for my thesis
3. All of my friends in Faculty of Computer Science, as well as any other friends of mine outside Faculty of Computer Science who keep motivate me.

I would like to apologize if there are some drawbacks in my projects. Hopefully this project can be useful for other people.

Semarang, July 31, 2013



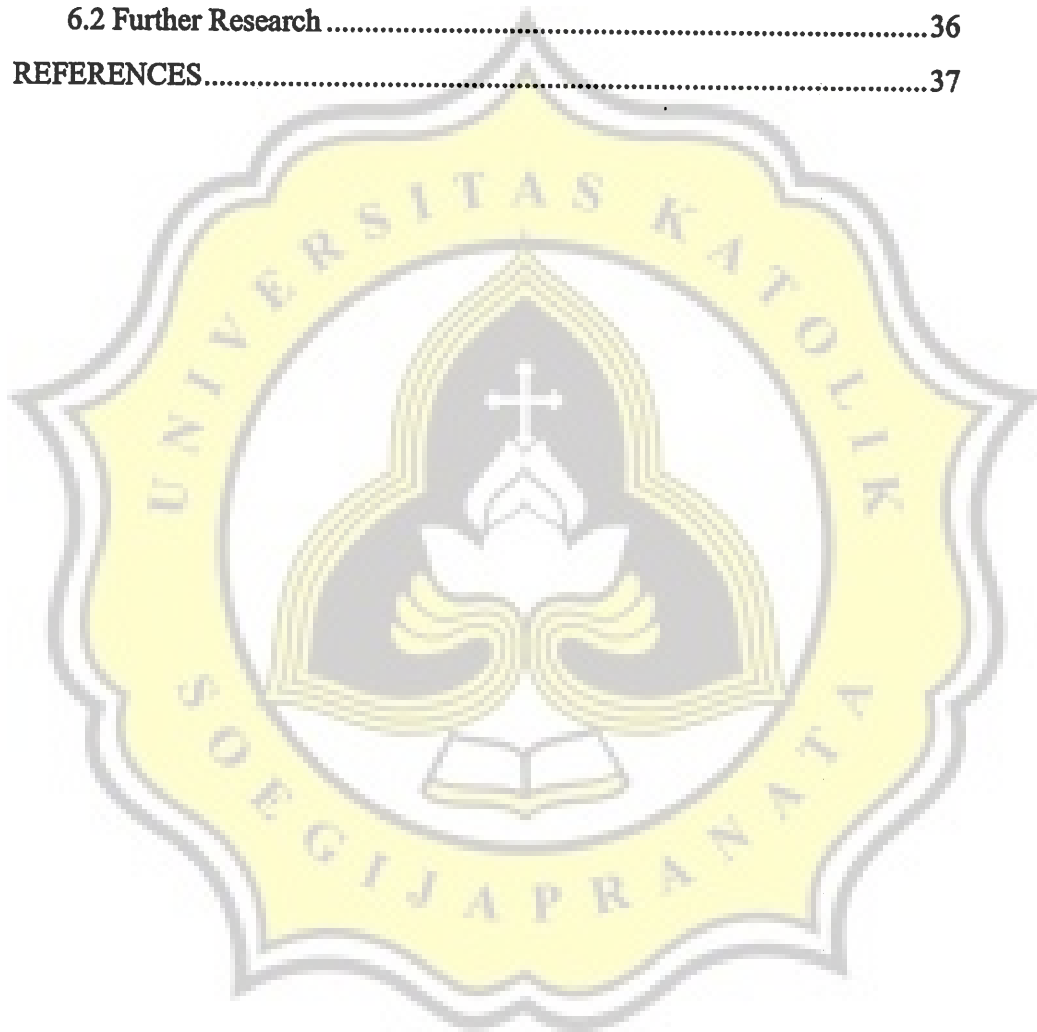
Ferdinand Budi Kurniawan

09.02.0004

## TABLE OF CONTENTS

|   |     |
|---|-----|
| APPROVAL AND RATIFICATION PAGE .....          | i   |
| STATEMENT OF ORIGINALITY .....                | ii  |
| ABSTRACT .....                                | iii |
| FOREWORD .....                                | iv  |
| TABLE OF CONTENTS .....                       | v   |
| TABLE OF FIGURES .....                        | vi  |
| TABLE OF TABLES.....                          | vii |
| CHAPTER 1: INTRODUCTION .....                 | 1   |
| 1.1 Background .....                          | 1   |
| 1.2 Scope.....                                | 1   |
| 1.3 Objective .....                           | 2   |
| CHAPTER 2: LITERATURE STUDY .....             | 3   |
| 2.1 Data Structure.....                       | 3   |
| 2.1.1 Hash Table .....                        | 3   |
| 2.1.2 Binary Search Tree.....                 | 4   |
| 2.2 Algorithm .....                           | 5   |
| 2.2.1 Rule Based Machine Translation .....    | 5   |
| 2.2.2 Example Based Machine Translation ..... | 6   |
| CHAPTER 3: PLANNING .....                     | 8   |
| 3.1 Research Methodology.....                 | 8   |
| 3.2 Project Management.....                   | 8   |
| CHAPTER 4: ANALYSIS & DESIGN .....            | 9   |
| 4.1 Analysis.....                             | 9   |
| 4.1.1 Use Case Diagram.....                   | 9   |
| 4.2 Design .....                              | 9   |
| 4.2.1 Flow Chart Diagram .....                | 9   |
| 4.2.2 Class Diagram .....                     | 10  |
| CHAPTER 5: IMPLEMENTATION & TESTING .....     | 15  |
| 5.1 Implementation .....                      | 15  |

|  |    |
|--|----|
| 5.2 Testing.....                               | 33 |
| 5.2.1 Testing Table .....                      | 33 |
| 5.2.2 Testing Screenshot.....                  | 34 |
| CHAPTER 6: CONCLUSION & FURTHER RESEARCH ..... | 36 |
| 6.1 Conclusion .....                           | 36 |
| 6.2 Further Research .....                     | 36 |
| REFERENCES.....                                | 37 |



## TABLE OF FIGURES

|   |    |
|---|----|
| Figure 2.1 Example of Hash Table .....              | 3  |
| Figure 2.2 Hash function.....                       | 3  |
| Figure 2.3 Illustration of Binary Search Tree ..... | 5  |
| Figure 4.1 Use case diagram .....                   | 9  |
| Figure 4.2 Flow chart diagram.....                  | 9  |
| Figure 4.3 Class Diagram.....                       | 10 |
| Figure 4.4 Dict.java.....                           | 10 |
| Figure 4.5 Corpus.java .....                        | 11 |
| Figure 4.6 TheBinaryTree.java .....                 | 11 |
| Figure 4.7 Node.java .....                          | 11 |
| Figure 4.8 UserInput.java.....                      | 12 |
| Figure 5.1 AppServletContextListener.java.....      | 13 |
| Figure 5.2 Index.jsp .....                          | 15 |
| Figure 5.3 Node.java .....                          | 16 |
| Figure 5.4 TheBinaryTree.java .....                 | 22 |
| Figure 5.5 Dict.java .....                          | 24 |
| Figure 5.6 Corpus.java.....                         | 32 |
| Figure 5.7 One word translations .....              | 34 |
| Figure 5.8 two words translations.....              | 34 |
| Figure 5.9 three words translations .....           | 34 |
| Figure 5.10 four words translation.....             | 35 |
| Figure 5.11 more than four words translation.....   | 35 |
| Figure 5.12 suggest an input .....                  | 35 |



## TABLE OF TABLES

|  |    |
|--|----|
| Table 2.1 Patterns.....                    | 6  |
| Table 3.1 project management timeline..... | 8  |
| Table 5.1 Testing table.....               | 33 |

