Synonym Searching using Stemming and Levensthein Distance Algorithms

Miltiades Aditya Siswa Hendratama
09.02.0052
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
Synonym Searching Using Stemming and Levensthein Distance Algorithms

This project report has been approved and ratified by the Dean of faculty of Computer Science and Supervisor on January 10, 2013.

With Approval,

Examiner,

Shinta Estri Wahyuningrum, S.Si., M.Cs
NPP: 058.1.2007.272

R. Setiawan Aji Nugroho, ST., M.ComplT
NPP: 058.1.2004.264

Examiner,

Rosita Herawati, ST., MIT
NPP: 058.1.2004.263

Hironimus Leong, S.Kom., M.Kom
NPP: 058.1.2007.273

Examiner,

Suyanto EA., Ir., M.Sc

Hironimus Leong, S.Kom., M.Kom
NPP: 058.1.2007.273
STATEMENT OF ORIGINALITY

Hereby signed:

Name : Miltiades Aditya Siswa Hendratama

ID : 09.02.0052

Declare that the final project entitled:

_Synonym Searching Using Stemming and Levensthein Distance Algorithms_

Was created by myself and wasn’t plagiarized from any other source, except that in writing expressed to the other article. If it proven that this final project was plagiarized from other source, then I’m ready to face the consequences.

Semarang, January 10, 2013

[Signature]

Miltiades Aditya Siswa Hendratama

09.02.0052
FOREWORD

Thanks to everyone who helped me to finish this project, which titled: Synonym Searching Using Stemming and Levenstein Distance Algorithms.

In this opportunity, I would especially thank to:

1. Yesus Kristus and Bunda Maria who blessed me with passion to finish this project.

2. My Mother, My Father and My Big Family for their support and prayer.

3. R.Setiawan Aji Nugroho, ST, M.Comput for always guides me from the start to the end.

4. All lecturers in Faculty of Computer Science, Hironimus Leong, S.Kom., M.Kom, Suyanto EA., Ir, M.Sc, Shinta Estri Wahyuningrum, S.Si., M.Sc, Rosita Herawati ST.,M.I.T. for their help and advice in programming.

5. Veronica Anjas and All my friend who supported me in the making of this project.

Finally, I would like to apologize for any imperfection in this project. I hope this project can encourage other people to create a better one. I will accept any critics and suggestion to improve this project.

Semarang, January 10, 2013

Miltiades Aditya Siswa Hendratama
09.02.0052
ABSTRACT

Synonym Searching Using Stemming and Levenshtein Distance Algorithms is a program that can find common words, program is very practical to use. The program is built with the collusion detection data structure hash tree, and using Stemming algorithm to obtain basic words and Levenshtein distance as algorithms for suggestion words, if the search terms are not met.

Keywords: Hash function, Binary Tree, Stemming and Levenshtein Distance
# TABLE OF CONTENTS

- Approval and Ratification Page ................................................. ii
- Statement of Originality .......................................................... iii
- Foreword ....................................................................................... iv
- Abstract ......................................................................................... v
- Table of Contents ......................................................................... vi
- Table of Figures ........................................................................... viii
- Table of Tables ............................................................................. x

**Chapter I** Introduction ..................................................................... 1

1.1 Background .................................................................................. 1

1.2 Scope .......................................................................................... 1

1.3 Objective ..................................................................................... 2

**Chapter II** Literature Study ............................................................ 3

2.1 Data Structures ........................................................................... 3

2.1.1 Binary Tree ............................................................................. 4

2.1.2 Linked List ............................................................................. 5

2.2 Algorithms .................................................................................. 6

2.2.1 Algorithms stemming word ................................................. 6

2.2.2 Algorithms levenstein distance ........................................... 7

**Chapter III** Planning ....................................................................... 9

3.1 Research Methodologies .............................................................. 9

3.2 Project Management .................................................................... 9
<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Figure 2.1</td>
<td>Hash Function</td>
<td>3</td>
</tr>
<tr>
<td>Figure 4.1</td>
<td>Use Case Diagram</td>
<td>10</td>
</tr>
<tr>
<td>Figure 4.2</td>
<td>Class Diagram Kamus</td>
<td>11</td>
</tr>
<tr>
<td>Figure 4.3</td>
<td>BacaFile Class</td>
<td>12</td>
</tr>
<tr>
<td>Figure 4.4</td>
<td>BinaryTree Class</td>
<td>12</td>
</tr>
<tr>
<td>Figure 4.5</td>
<td>HashTable Class</td>
<td>13</td>
</tr>
<tr>
<td>Figure 4.6</td>
<td>InggrisStemmer Class</td>
<td>13</td>
</tr>
<tr>
<td>Figure 4.7</td>
<td>IndonesiaStemmer Class</td>
<td>14</td>
</tr>
<tr>
<td>Figure 4.8</td>
<td>Kamus Class</td>
<td>14</td>
</tr>
<tr>
<td>Figure 4.9</td>
<td>LevenstheinDistance Class</td>
<td>15</td>
</tr>
<tr>
<td>Figure 4.10</td>
<td>LinkedList Class</td>
<td>15</td>
</tr>
<tr>
<td>Figure 4.11</td>
<td>MainKamus Class</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4.12</td>
<td>Node Class</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4.13</td>
<td>Stemmer Class</td>
<td>16</td>
</tr>
<tr>
<td>Figure 4.14</td>
<td>Tampilan Class</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4.15</td>
<td>Translator Class</td>
<td>17</td>
</tr>
<tr>
<td>Figure 4.16</td>
<td>TreeNode Class</td>
<td>18</td>
</tr>
<tr>
<td>Figure 4.17</td>
<td>Word Class</td>
<td>18</td>
</tr>
<tr>
<td>Figure 5.1</td>
<td>Running Program</td>
<td>36</td>
</tr>
<tr>
<td>Figure 5.2</td>
<td>Input word Prgram</td>
<td>........................................... 36</td>
</tr>
<tr>
<td>Figure 5.3</td>
<td>Show Result</td>
<td>........................................... 37</td>
</tr>
<tr>
<td>Figure 5.4</td>
<td>Show Suggestion Word</td>
<td>........................................... 37</td>
</tr>
<tr>
<td>Figure 5.5</td>
<td>GUI Program</td>
<td>........................................... 38</td>
</tr>
<tr>
<td>Figure 5.6</td>
<td>Input Word with Gui</td>
<td>........................................... 38</td>
</tr>
<tr>
<td>Figure 5.7</td>
<td>Show Result with Gui</td>
<td>........................................... 39</td>
</tr>
<tr>
<td>Figure 5.8</td>
<td>Show Suggestion with Gui</td>
<td>........................................... 39</td>
</tr>
</tbody>
</table>
TABLE OF TABLES

Table 3.1  Project Schedules .................................................. 8