AGENT-BASED SOLUTION FOR IMPROVING ABSTRACTS

Adhikari Mudiyanselage Tharindu Buddhika Adhikari

Degree of Master of Science in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa
Sri Lanka

December 2015
AGENT-BASED SOLUTION FOR IMPROVING ABSTRACTS

Adhikari Mudiyanselage Tharindu Buddhika Adhikari

Thesis submitted in partial fulfillment of the requirements for the degree of MSc in Artificial Intelligence

Department of Computational Mathematics

University of Moratuwa

Sri Lanka

December 2015
Declaration

I declare that this dissertation does not incorporate, without acknowledgment, any material previously submitted for a Degree or a Diploma in any University and to the best of my knowledge and belief, it does not contain any material previously published or written by another person or myself except where due reference is made in the text. I also hereby give consent for my dissertation, if accepted, to be made available for photocopying and for interlibrary loans, and for the title and summary to be made available to outside organization.

A.M.T.B. Adhikari

Date: 31st December 2015

Supervised by

Prof. Asoka Karunananda

Date: 31st December 2015
Acknowledgements

I would like to express my sincere gratitude to my supervisor Prof. Asoka Karunananda for guiding me throughout this research project. He has given his fullest corporation to me whenever I sought for advice. Also, Prof. Karunnanda’s teaching techniques helped me a lot to consider problems in different avenues. I must also thank all the members of the lecturer panel. During the period of one year and three months, these valuable lectures helped me a lot to think differently.

I would like to place my gratitude to my loving parents for always encouraging for me on higher studies. Also, I would like to appreciate Mr. Dimuthu and Ms. Nadeeka from the non-academic staff for their support that they provided throughout the period of the course.

For this research, I had to refer to many books and research papers as reference. I would like to thank all the authors of those publications. Also, my batch mates helped me a lot in many different ways. I would like to pay my gratitude to them. They made my life enjoyable during the course period.

Last but not least, I would like to thank all my colleagues and others who are not mentioned, for all the support extended to me. My co-workers helped me a lot. So, that I can focus much more on the research and the MSc. Without their dedication the project would not have been successful.
Abstract

Writing abstracts in a comprehensive and meaningful manner is a challenge for any researcher. However, an abstract includes limited set of verbs and standard phrases and other good practices of structuring the contents. A research has been conducted to develop an Agent-based Solution for Improving Abstracts. This solution is based on multi agent systems technology and natural language processing together with commonly used verb phrases and other good practices. The system has been developed with nine agents, namely, coordination agent, parser agent, problem agent, solution agent, conclusion agent, content agent, synonym agent, improvement agent and restructure agent. The coordination agent coordinates entire process. The parser agent identifies syntactic information of each sentence and prepares the contents of the abstract for further analysis. The problem agent ensures whether the research problem has been stated in the early part of the abstract and its proportion within the abstract. The solution agent checks for the contents in terms of concepts such as hypothesis, methodology, approach, design, implementation, methods, theoretical framework, technology, hardware, software, and sampling based on the key words. The conclusion agent searches for concepts such as testing, evaluation, data analysis and statistical significance based on the key words. The content agent, improvement agent, synonym agent, and restructure agent are responsible to offer guidelines to modify and improving of the abstract. More importantly, these agents interact with each other and deliberate to reach consensus regarding a solution. For instance, problem agent and solution agent may agree on the proportion of respective contents within the abstract. Each agent has its own Ontology for deliberating with other agents. The Stanford CoreNLP Natural Language Processing Toolkit has been used to develop parser and JADE has been used for development of the entire multi agent system. The system has been developed with JAVA to run on Windows. It has been incrementally tested, and shown interesting results related to checking for completeness of the abstract in terms required materials and suggestion for improvements.
# Contents

Chapter 1 Introduction ............................................................................................ 14  
1.1 Prolegomena ................................................................................................... 14  
1.2 State of the art Artificial Intelligence field ................................................... 15  
1.3 Aim and Objectives ........................................................................................ 19  
1.4 Background and Motivation ......................................................................... 19  
1.5 Problem in brief ............................................................................................ 19  
1.6 Approach ........................................................................................................ 19  
1.7 Structure of the thesis .................................................................................. 20  
1.8 Summary ....................................................................................................... 20

Chapter 2 Review of the State of the Art ............................................................. 21  
2.1 Introduction .................................................................................................... 21  
2.2 An Overview of Natural Language Processing by Computers .................. 21  
2.3 Natural Language Processing using Machine Learning Algorithms ........ 23  
2.4 Most Famous Research Areas in Natural Language Processing ............... 24  
  2.4.1 Automatic summarization ..................................................................... 24  
  2.4.2 Coreference resolution ......................................................................... 24  
  2.4.3 Discourse analysis ............................................................................... 24  
  2.4.4 Machine translation ............................................................................. 25  
  2.4.5 Morphological segmentation ............................................................... 25  
  2.4.6 Named Entity Recognition (NER) ...................................................... 25  
  2.4.7 Natural language generation ................................................................. 25  
  2.4.8 Natural language understanding ............................................................ 25  
  2.4.9 Optical Character Recognition (OCR) ................................................. 26  
  2.4.10 Part-of-speech tagging ....................................................................... 26  
  2.4.11 Parsing .................................................................................................. 26  
  2.4.12 Question Answering ........................................................................... 26  
  2.4.13 Relationship Extraction ...................................................................... 26  
  2.4.14 Sentence Breaking ............................................................................... 26  
  2.4.15 Sentiment Analysis ............................................................................ 26  
  2.4.16 Speech Recognition ............................................................................ 27
Chapter 3 Cutting Edge Technologies Integrated

3.1 Introduction

3.2 Natural Language Processing

3.3 Bayesian approach
6.3 Coordination Agent and Parser Agent ....................................................... 68
6.4 Problem Agent .......................................................................................... 68
6.5 Solution Agent .......................................................................................... 68
6.6 Conclusion Agent....................................................................................... 68
6.7 Content Agent, Synonym Agent, Improvement Agent and Restructure 
Agent ................................................................................................................ 69
6.8 Multi Agent System ................................................................................... 70
6.9 Testing........................................................................................................ 70
6.10 The Stanford CoreNLP Natural Language Processing Toolkit ............... 71
6.11 JADE Framework .................................................................................... 72
6.12 Summary .................................................................................................. 73

Chapter 7 Evaluation of the Proposed Solution ..................................................... 74
7.1 Introduction ............................................................................................ 74
7.2 Evaluating the Overall System................................................................ 74
7.3 Evaluating the Title and Abstract Overall Match ........................................ 75
7.4 Evaluating the Checking whether the Problem is stated .......................... 75
7.5 Evaluating the Checking whether the Solution is stated ......................... 76
7.6 Evaluating the Checking whether the Conclusion is stated .................... 77
7.7 Evaluating the Checking the Balance between Problem, Solution & 
Conclusion ....................................................................................................... 78
7.8 Evaluating the Checking the Content .................................................... 79
7.9 Evaluating the Suggesting the Synonyms .............................................. 80
7.10 Evaluating the Suggesting the Improvements ....................................... 81
7.11 Evaluating the Suggesting the Restructurings ..................................... 82
7.12 Summary ............................................................................................ 93

Chapter 8 Conclusion and Further Work of the Research ................................... 94
8.1 Introduction ............................................................................................ 94
8.2 Conclusion ............................................................................................. 94
8.3 Further Work.......................................................................................... 94
8.4 Improving the Multi Agent System ....................................................... 95
8.5 Improving the Ontologies ................................................................. 95
8.6 Improving the Natural Language Processing Capabilities ..................... 96
8.7 Improving the User Friendliness......................................................... 96
8.8 Summary...........................................................................................................96

Chapter 9 References..............................................................................................97

Appendix A: Software Code ................................................................................100
A.1 Introduction .........................................................................................................100
A.2 Coordination Agent .............................................................................................100
A.3 Parser Agent ......................................................................................................100
A.4 Problem Agent ...................................................................................................100
A.5 Solution Agent ...................................................................................................100
A.6 Conclusion Agent ...............................................................................................100
A.7 Content Agent ....................................................................................................100
A.8 Synonym Agent ..................................................................................................101
A.9 Improvement Agent ...........................................................................................101
A.10 Restructure Agent ............................................................................................101
A.11 Abstract Concept ..............................................................................................101
A.12 Message Concept ...............................................................................................101
A.13 Parser Ontology ................................................................................................101
A.14 Problem Ontology ..............................................................................................101
A.15 Solution Ontology ..............................................................................................102
A.16 Conclusion Ontology ..........................................................................................102
A.17 Content Ontology ..............................................................................................102
A.18 Synonym Ontology ............................................................................................102
A.19 Improvement Ontology .....................................................................................102
A.20 Restructure Ontology .........................................................................................102
A.21 Abstract GUI .....................................................................................................102
A.22 Text Area Output Stream ..................................................................................102

Appendix B: User Interfaces of the Proposed System ..............................................103
B.1 Introduction ..........................................................................................................103
B.2 Main User Interface of ASIA ...............................................................................103
B.3 JADE Remote Agent Management GUI ............................................................104
B.4 Open Abstract Dialog GUI ..................................................................................104
B.5 Abstract Open in the Text Area ............................................................................105
B.6 Enter the Title of the Abstract in the Text Area .....................................................106
B.7  Submit the Abstract to the System.......................................................107
B.8  Save the Suggested Improvements .....................................................108
B.9  Save the Agent Message Space..........................................................109
B.10 Clear the Agent Message Space..........................................................110
B.11 Clear the Suggested Improvements.....................................................111
B.12 Clear the Abstract and Title ...............................................................112
B.13 Project Workspace in NetBeans IDE .................................................113
B.14 Project Run Output in NetBeans IDE ..................................................114
List of Figures

Figure 1.1: Acting humanly: The Turing Test approach 15
Figure 1.2: Thinking humanly: The cognitive modeling approach 16
Figure 1.3: Acting rationally: The rational agent approach 17
Figure 1.4: Thinking rationally: The "laws of thought" approach 18
Figure 2.1: Major research areas in Natural Language Processing 35
Figure 3.1: System Architecture of Stanford CoreNLP Toolkit 54
Figure 4.1: High Level Process Diagram 59
Figure 5.1: Design Architecture Diagram 64
Figure 6.1: System Architecture of Stanford CoreNLP Toolkit 72
Figure 7.1: Evaluation process of the proposed solution 75
Figure 7.2: Evaluation process of the Title and Abstract Overall Match 76
Figure 7.3: Evaluation process of the Checking whether the Problem is stated 77
Figure 7.4: Evaluation process of the Checking whether the Solution is stated 78
Figure 7.5: Evaluation process of the Checking whether the Conclusion is stated 79
Figure 7.6: Evaluation process of the Checking the Balance between Problem, Solution & Conclusion 80
Figure 7.7: Evaluation process of the Checking the Content 81
Figure 7.8: Evaluation process of the Suggesting the Synonyms 82
Figure 7.9: Evaluation process of the Suggesting the Improvements 83
Figure 7.10: Evaluation process of the Suggesting the Restructurings 84
Figure 8.1: Further Work of the Research 86
Figure B.1: Main User Interface of ASIA 95
Figure B.2: JADE Remote Agent Management GUI 95
Figure B.3: Open Abstract Dialog GUI 96
Figure B.4: Abstract Open in the Text Area 97
Figure B.5: Enter the Title of the Abstract in the Text Area 98
Figure B.6: Submit the Abstract to the System 99
Figure B.7: Save the Suggested Improvements 100
Figure B.8: Save the Agent Message Space 101
Figure B.9: Clear the Agent Message Space 102
Figure B.10: Clear the Suggested Improvements 103
Figure B.11: Clear the Abstract and Title  104
Figure B.12: Project Workspace in NetBeans IDE  105
Figure B.13: Project Run Output in NetBeans IDE  105
List of Tables

Table 1.1: Applications of Artificial Intelligence Technologies 18
Table 2.1: Applications of Multi Agent Systems 40
Table 3.1: Adapted Cutting Edge Technologies 56
Table 4.1: Inputs and Outputs of the System 60
Table 5.1: Process Description in the Proposed Design 67