Ontological solution for enabling Continuous Learning

Tharanga K.G. D.

129110D

Dissertation submitted to the Faculty of Information Technology, University of Moratuwa, Sri Lanka for the partial fulfilment of the requirements of the Degree of MSc in Artificial Intelligence

November 2014

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.

Name of Student: Tharanga K.G.D.

Signature of Student

Date: 22.11.2014.

Supervised by

Name of Supervisor(s): Prof. Asoka S. Karunananda

Signature of Supervisor(s)

Date: 22/11/14

Acknowledgement

First and foremost I would like to extend the sincere gratitude to my supervisor Professor Asoka S. Karunananda, who always had time to spare with us to give assistance and guidance, to make the project a success. I would like to thank University of Moratuwa for giving an opportunity to carry out this research project and its continuing support during the project.

I would like to pay gratitude to all the academic and non-academic staff members of University of Moratuwa and the batch mates for their generous support, comments and encouragements throughout the project.

At last but not least I would like to humbly thank my parents, husband and friends for standing beside me at every time. If not for your support none of this would have been possible.

Abstract

Human being starts the learning process with a minimal knowledge and then constructs the knowledge by adding new contents or updating the existing knowledge. The research 'Ontological solution for enabling Continuous Learning' provides an approach, which can learn and expand the knowledge somewhat similar to a human being. Human knowledge is constructed by gathering totally new concepts or expanding the already existing content. People have identified the importance of intelligent machines which can learn and behave as human beings. Intelligent and advanced systems have become popular with the development and rapid changes of the technology. Knowledge representation and updating can be defined as a most crucial requirement in intelligent systems. Knowledge management researches focus on the development of concepts, methods, and tools supporting the management of human knowledge. In most of the intelligent systems, knowledge representation and knowledge updating processes are done by a human expert and there is no proper way to represent and update knowledge in an automated way without human involvement. Knowledge representation via ontology can be introduced as a most popular approach to represent and manipulate knowledge. 'Ontological solution for enabling Continuous Learning' can be defined as an automated knowledge acquisition mechanism by constructing ontology and integrating two or more ontology. The research reveals a mechanism to update the existent knowledge by comparing and integrating new knowledge. This thesis is meant to delineate the background, research review, functions and features of the research. It further explained the purpose, literature, approach, methodology, evaluation of the research, the constraints under which it must operate, conclusion and benefits of the research. The software is endowed with inventions involving in knowledge representation via ontology, which are still under research. Knowledge extraction, knowledge comparison, knowledge representation and knowledge integration can be defined as the main functions of the system. This solution can be used as the main learning process of any application and the knowledge will be easily manipulated in any intelligent system by embedding this ontology based solution.

Table of Contents

Declaration II

Acknowledgement III

Abstract IV

List of Figures VIII

List of Tables IX

Chapter 1 Introduction 1

1.1 Prolegomena 1

1.2 Learning in Intelligence Systems 1

1.2 Background & Motivation 3

1.3 Aim and Objectives 4

1.4 Structure of the thesis 5

1.5 Summary 6

Chapter 2 Current strengths & weaknesses in learning mechanism of intelligent systems 7

2.1 Introduction 7

2.2 Major researches in learning systems 7

2.2.1 OntoCase - A Pattern-Based Ontology Construction Approach 7

2.2.2 Aspects of Automatic Ontology Extension: Adapting and Re-Generalizing Dynamic Updates 8

2.2.3 SMART: Algorithm and Tool for Automated Ontology Merging and Alignment 9

2.2.4 A Protégé Plug-In for Ontology Extraction from Text Based on Linguistic Analysis 10

2.2.6 OntoPop or how to annotate documents and populate ontologies from texts 12

2.2.7 Apolda: A Practical Tool for Semantic Annotation 13

2.2.8 Semantic Similarity Methods in WordNet and their Application to Information Retrieval on the Web 14

2.2.9 OntoLearn 15

2.2.10 Flexible Ontology Population from Text: The OwlExporter 16

2.2.11 Incorporating ontological background knowledge into Information Extraction 17

2.2.12 TextOntoEx: Automatic Ontology Construction from Natural English Text 18

2.3 Summary 19

Chapter 3 Power of Ontological Modelling 20

3.1 Introduction 20

3.2 Ontology in Learning 20

3.3 Summary 22

Chapter 4 Ontology modelling for continuous learning 23

4.1 Introduction 23

4.2 Hypothesis 23

4.3 Input 23

4.4 Process 23

4.4 Output 25

4.5 Features 25

4.7 Summary 25

Chapter 5 Design 26

5.1 Introduction 26

5.2 Architecture of the System 26

5.3 Summary 33

Chapter 6 Implementation 34

6.1 Introduction 34

6.2. Tools and Technologies 34

6.3 Implementation of 'Ontological solution for enabling Continuous Learning' 35
6.3.1 Development of Semantic information extraction module 35
6.3.2 Ontology comparison module 36
6.3.3 Knowledge Integration module 37

6.4 Summary 41

Chapter 7 Evaluation 42

7.1 Introduction 42

7.2 Statistical Analysis 42
7.2.1 Ontology construction from the concept 42
7.2.2 Update the knowledge by integrating main ontology and new ontology. 44

7.3 Results evaluation 46

7.4 Summary 47

Chapter 8 Conclusion 48

8.1 Introduction 48
8.2 Closing Remarks of the research 48
8.3 Future Work 50
8.4 Summary 51
References **52**
Appendix A – Resources of the system **55**
Appendix B – Related Work **75**

List of Figures

Figure 5.1: Module flow diagram 27

Figure 5.2 : High-level architecture of the system 28

Figure 5.3 : New concept 29

Figure 5.4: Identification of tokens 29

Figure 5.5 : Created Ontology – owl file 30

Figure 5.6 : Main ontology 32

Figure 5.7 : Updated ontology 33

Figure 6.1: Knowledge Integration algorithm 38

Figure 6.2 : WordAnalyzer algorithm 38

Figure 6.3 : Semantic matching algorithm 39

Figure 6.4: Similarity checking algorithm 40

Figure 6.5: Class diagram of the system 41

List of Tables

Table 7.1 : Statistical analysis table to measure the accuracy 45

Table 7.2 : Accuracy of Identification Cross tabulation 45

Table 7.3 : Symmetric measures 46

Table 7.4 : System's recognition over the provided inputs 46